Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

David Berendes Date

Environmental Risk Factors during the Zimbabwe Cholera Outbreak, 2008-2009

By

David Berendes

Master of Science in Public Health

Global Epidemiology

_____

Juan Leon

Committee Chair

_____

Thomas Handzel

Committee Member

_____

Diane Morof

Committee Member

Environmental Risk Factors during the Zimbabwe Cholera Outbreak, 2008–2009

By

David Berendes

B.S., Duke University, 2009

Thesis Committee Chair: Juan Leon PhD, MPH

An abstract of

A thesis submitted to the Faculty of the
Rollins School of Public Health of Emory University

in partial fulfillment of the requirements for the degree of
Master of Science in Public Health
in Global Epidemiology

2011

Abstract

Environmental Risk Factors during the Zimbabwe Cholera Outbreak, 2008–2009

By David Berendes

Cholera is a waterborne disease that causes outbreaks worldwide. Despite advances in safe water distribution networks in the developing world, cholera and other waterborne diseases persist in the environment. In Zimbabwe, gradual economic collapse over the last 10 years culminated in the creation of a complex humanitarian emergency state in 2008, with massive loss of health and water infrastructure. This situation put the country at risk for one of the largest and most severe cholera outbreaks in the past 10 years. The goal of this study was to assess the effect of safe water use, particularly borehole use, at the village level in preventing cholera spread and lowering cholera attack rates in the Chivi and Gokwe North districts during the outbreak. Safe water use, and borehole use in particular, was associated with reduced attack rates and fewer outbreak-affected villages when compared with unsafe water use by villages. Both the reductions seen in attack rates and those seen in cholera prevalence across villages were borderline significant. Missing data was an issue in the analysis, and imputation methods were compared and contrasted in the development of final logistic and multiple linear regression models. Despite the limited availability of timely point-of-use water treatment in the Zimbabwean cholera outbreak, a characteristic inherent to many response efforts, there was suggestive but inconclusive evidence that water quality at the source may reduce cholera morbidity by itself. This paper has important implications not only for field outbreak data methodology, but for water and sanitation promotion as well. While simple imputation methods seem to be the norm in outbreaks in the field, there was value in multiple imputation methods for improving the validity and precision of the model estimates.

Environmental Factors in Boreholes Use Affecting the Cholera Outbreak in Zimbabwe

By

David Berendes

B.S., Duke University, 2009

Thesis Committee Chair: Juan Leon PhD, MPH

A thesis submitted to the Faculty of the

Rollins School of Public Health of Emory University

in partial fulfillment of the requirements for the degree of
Master of Science in Public Health
in Global Epidemiology

2011

Acknowledgements:

I would like to thank everyone in my life who contributed to and supported me through this thesis. To my friends and family at Emory and IERHB: I could not have remained sane without you. To my mother and late father especially: you have been my inspiration in this field, and my hope is to one day live up to the lofty standards of service you have given the world and do it all with the same grace and kindness.

To my field advisor, Tom Handzel: your attitude and work inspired me to pursue water and sanitation work in emergency settings. Thanks for the hours of guidance and mentorship in the face of your demanding schedule. You definitely pushed me to do the best epidemiologic work I could on this with your own work and examples, and I couldn't have done this without you. I hope to emulate your work ethic and learn to balance all aspects of the job as well as you do. You gave me my first step into this world at CDC, I hope to inspire and earn your trust and run with this opportunity. I look forward to continuing to work with you.

To my favorite EIS officer, Diane Morof: without your patience and responses to the constant bombardment of questions I raised, I would never have written this thesis. Thanks for taking large amounts of time out of your own fellowship experience to nurture a fellow newbie. You were a fantastic (and natural) mentor, and I know you will do an outstanding job wherever you end up. I hope to work with you more in the near future as well.

To my faculty advisor, Juan Leon: I name you as a faculty advisor, but know that you also filled the roles of life advisor, cheerleader, moral supporter, coach, and friend. I thank you so very much for the patience and guidance you've given me throughout the last 2 years. Wherever my career in water and sanitation goes, I know that it started in your lab with your support and guidance. You are, without a doubt, the faculty member who has most influenced, contributed to, and guided my academic career, and I hope to work more with you in the future. I can only expect that your name and your influence will grow more and more as you continue to produce fantastic work and do great mentoring. The success of your work and the great people who work around you is a tribute to your diligence and effort. I feel honored to have been in your lab (and expect I may find myself there again in the near future).

Table of Contents

Literature Review:

Cholera is an infectious waterborne disease that affects millions of vulnerable populations, especially in the poorest and least developed countries and regions. It is a disease that can spread quickly through a population, resulting in individuals with dehydration and causing severe morbidity and mortality. Cholera is endemic in some regions and has epidemic potential in others. There have been seven cholera pandemics: the earliest began in 1817 around the Indian subcontinent, while the most recent began in 1961 in Indonesia and spread to Africa in 1970, and South America in 1991 [1, 2].

Between 2006 and 2009, almost one million cases of cholera were reported to the World Health Organization (WHO), though the actual burden is probably thousands to millions more due to incomplete surveillance and misclassification [3-6]. The WHO estimates that only 1% of cholera cases are actually reported and that the total burden is close to 3–5 million cases worldwide yearly, with 120,000 deaths [7, 8]. This estimation yields a case fatality rate (CFR) of 4% [7, 8]. Another study estimated the total global burden of cholera was closer to 11 million cases per year [9]. Of the 826,255 cholera cases reported to WHO between 2006 and 2009, 20,431 deaths were reported, yielding a CFR of about 2.5% [3-6], still well above the emergency containment threshold of 1% CFR [10]. This CFR underscores cholera's role as a major public health disease of international proportions. Nearly half (20 of 41) countries reporting native (non-imported) cholera cases to the WHO in 2009 had CFRs above the 1% threshold. Those 20 countries contained over 78% of all cases reported worldwide [3], indicating a disproportionate burden of cholera on select countries.

Cholera Surveillance

It is difficult to accurately estimate the burden of cholera infection in a population primarily because cholera's presentation varies from asymptomatic to severe. Given this range of symptoms, identification of cholera in a community in overburdened and underdeveloped surveillance systems is difficult. Estimates indicate that from as low as 1% to as high as 33% of infections can be symptomatic, with the majority of clinical estimates ranging from 20–30% [1, 7, 11-13]. Because the majority of cases are mild to moderate in severity, treatment advice is generally not sought and stool culture, a key identification tool in cholera outbreaks, may not be undertaken. Without using cultures to isolate the bacterium, mild to moderate cholera symptoms may easily be confused with those of other diarrheal diseases, making reliable surveillance estimates of cholera cases difficult to attain [7].

While the WHO presents surveillance information from reporting countries, there is great variation in the quality of the passive and active surveillance systems in countries, if those surveillance systems are present at all. In 2005, 95% of reported cases worldwide were in Africa, and in 2009, 98% were in Africa [3, 14]. These numbers, however, exclude the endemic seasonal cholera cases in Bangladesh and India that comprise a significant number of unreported or misreported—as acute water diarrhea (AWD)—cases [3, 7, 14]. While thousands to millions of cases can be misclassified, there are a large proportion of cases simply labeled as "acute watery diarrhea" to avoid social or political stigma that could impact the country's government and tourism, among other problems. In 2009, only 27 imported cases were reported to WHO, representing a probable underestimation of the true number of cholera cases [3]. Adding to misreporting, travelers to these areas may become unknowingly infected with an asymptomatic or mild form of cholera, causing the individual to experience symptoms in a short period while abroad. The short duration and variation in symptoms, if present at all, decreases the likelihood

of reporting the disease, even if the individual is from a nation with formidable surveillance infrastructure. If the individual does in fact 'import' the disease to their native land, there is still a low chance of the infection being recognized, documented, and treated as cholera, rather than one of the multitude of other similar diarrheal diseases or simply "traveler's diarrhea." While cholera is a major disease of epidemic proportions, it is still difficult to comprehensively document its spread due to asymptomatic cases.

Cholera Types and Environmental Features

Cholera is caused by the bacterium *Vibrio cholerae*, a rod-shaped and Gram-negative organism. *Vibrio cholerae* exists in various strains with multiple subgroups and serotypes. The infectious organism can be differentiated by serogroups based on its O antigen of its lipopolysaccharide component of the outer membrane. Serogroup O1 has been the main cause of epidemic cholera, although a newly-described outbreak of "O139 Bengal" in India and Bangladesh in 1992 caused concern of an 8th, concurrent pandemic. Serogroup O139 now coexists with O1 in Bangladesh and India, though it has failed to spread outside of that region in the epidemic proportions that were anticipated [1]. The O1 serogroup is further subdivided into two distinct biotypes: El Tor and classical. Both biotypes are able to be further classified into serotypes Inaba and Ogawa [15, 16]. The seventh, current pandemic is comprised primarily of *V. cholerae* O1 El Tor, while the classical strain was associated with early pandemics. Both strains are capable of yielding severe infection in individuals.

In the environment, cholera organisms have many unique features that allow them to survive. Cholera bacteria can attach to plant, green algae, copepod, crustacean, and/or insect surfaces [17-26], and use these organisms as protection from harsh and otherwise less-than-hospitable aquatic conditions. Additionally, *V. cholerae* build up on biofilms, which provide survivable microenvironments for the organisms [17], and can switch between viable and non-

culturable states in response to nutrient conditions/deprivation [17, 27, 28]. This non-culturable state can be altered or resuscitated into a more infectious state with extreme environmental changes, such as exposure to the human intestinal environment from consumption [17, 27, 29]. The *V. cholerae* organism is considered a facultative human pathogen due to its persistent survival outside the human host, especially in O1 and O139 strains [17, 18, 30-32]. In the natural environment (ecosystems, non-human environments), many of the O1 strains are non-toxin forming [33], suggesting that cholera bacteria must undergo a genetic change that leads to toxin generation in the human host. This change may occur in or be facilitated by contact with intestinal conditions in humans [17, 34, 35].

Clinical Features of Cholera:

In humans, the clinical manifestations of cholera infection vary from isolated loose/watery diarrhea and mild dehydration to vomiting and extensive diarrhea with a characteristic "rice water" appearance. While usually painless, this diarrhea produces rapid dehydration, with volume loss of up to a liter an hour or more in adults with severe cholera [1]. In patients with severe cholera, symptoms of dehydration may present initially as restlessness and extreme thirst, but as the dehydration progresses and shock sets in, patients may become apathetic, develop poor skin turgor and sunken eyes, and have weak to undetectable blood pressures. Death can occur in the first few hours or day if the water and electrolyte loss is significant enough. On a population scale, while mortality from asymptomatic and mild infections (though mostly undocumented) is extremely low, outbreaks associated with *cholera gravis* (severe infection) can yield mortality rates of up to 50% if patients are left untreated [13].

The devastating diarrhea and dehydration characteristic of cholera are caused by the release of the cholera enterotoxin from the ingested bacterium. The toxin binds to the mucosal cells of the intestine, where it activates adenylate cyclase enzyme, causing increased production

of intracellular cyclic adenosine monophosphate (cAMP). The cAMP in turn causes those cells to pump out potassium, sodium, and other important ions. A large loss of water follows these ions out, causing the severe dehydration and "rice water" stool. This water loss contains toxin and free organisms in sufficient quantity to infect others in the area, especially if the excrement or water source it contaminates are handled improperly, ingested, or otherwise consumed or used in food processing during the subsequent exposure period.

Treatment of Cholera and Mortality:

The treatment of cholera requires replacement of adequate fluids and electrolytes to keep the individual alive throughout the symptomatic period, though the methods vary by case. Treatment with oral rehydration salts (ORS) is typically the primary treatment in patients retaining the ability to swallow fluids. This treatment has the benefit of replacing both fluid and electrolyte loss from the diarrhea using glucose to facilitate uptake. Proper treatment with ORS and IV fluids lowers mortality to around 1% [13], though these numbers depend on the size, location, and quality of the distributor or facility.

ORS has limited effectiveness if it cannot be sustained in sufficient capacity to replenish lost fluids and electrolytes and is ineffective in cases where the individual is so dehydrated that they cannot swallow fluids. In this case, treatment is switched to intravenous polyelectrolyte solution to sustain the treatment until the individual can consume fluid orally again. Severely dehydrated patients may require replacement of 10% of their body weight in the first 2–4 hours of treatment. While consistent and prolonged ORS treatment has been shown to be effective as long as fluids and electrolytes are being replaced at the rate they are lost, the challenge in resource-poor settings throughout the world is to provide consistent fluid treatment and track an individual's fluid loss. Improper or unsustained treatment can lead to complications, such as acute renal failure, electrolyte imbalance, and even miscarriage or premature delivery in pregnant

women [1, 36, 37]. Resources and staff are necessary to sustain the ORS or IV treatment, while use of a 'cholera cot'—essentially a cot with a hole in the middle draining into a bucket to collect the patient's excretions—allows for easy tracking of fluid loss.

Cholera Transmission

The organism is transmitted fecal-orally, via ingestion of water or food contaminated with the bacterium, frequently in the context of poor hygiene or sanitation. Contaminated water sources have been implicated in the spread of cholera since John Snow's investigations in 1855, and have been the most commonly implicated risk factor in recent global outbreaks [2]. Cholera transmission has also been associated with food itself, especially filter-feeding seafood or shellfish. Glass et al. [38] documented transmission by crabs, squid, partially dried fish, bi-valve mollusks, rice, and street food, among others, in their 1991 review, though this list is far from comprehensive. Food can become a vehicle for cholera transmission after contact with other contaminated food, contaminated water, or an infected food handler.

Transmission of cholera has been documented through social and community interactions as well. Cholera transmission in association with migration and residence in refugee camps has been documented since the 1970s. Asymptomatic migrants probably led to the spread of cholera throughout previously unaffected areas of Africa [38, 39]. Additionally, burial practices involving close contact with recent cholera deaths, including cleaning and transporting bodies, have been implicated in outbreaks [38, 40], including a 1994 outbreak in Guinea-Bissau [41]. In general, these funeral practices involve disinfection of the body, but are frequently accompanied by feasts, facilitating the fecal-oral contact with disease-free individuals.

While cholera requires a high infectious dose (10^5-10^8 organisms, varying on host characteristics) compared to shigellosis and other bacterial infections causing diarrhea, rapid spread associated with outbreaks still occurs due to large amounts of highly infectious shedding

in stool of symptomatically and asymptomatically infected individuals. Asymptomatic individuals can shed up to 10^5 vibrios per gram of stool, while mildly infected individuals can shed up to 10^8/g stool, and those with *cholera gravis*, shed on the higher end—10^7–10^9 /g stool, and also experience severe vomiting which can contain similar numbers of infectious vibrios (cited in [15]).

Cholera Outbreaks:

Outbreaks can occur from a single host's actions, whether that is migration of an infected case to a new area or ingestion of environmental *V. cholerae* by an immunogenically naïve individual, allowing the organism to change into an infectious state. During epidemics, cholera organisms can remain hyperinfectious for at least 5 hours after excretion from an infected patient. The persistence of this hyperinfectious state, combined with crowded conditions and poor sanitation and hygiene, can lead to quick uptake by other immunogenically-naïve individuals [15, 42, 43]. Because of the amount of cholera organisms shed by infected individuals and the length of the organism's hyperinfectious period, cholera outbreaks can appear to simulate and be on the scale of person-to-person outbreaks. The organism may spend little time in the environment alone, especially when hyperinfectious or in immunogenically-naïve communities [15, 44]. This characteristic allows for cholera to spread quickly and widely throughout a region or population, especially those with overcrowding and poor or insufficient sanitation and hygiene.

Safe Water:

Currently, 884 million people worldwide do not get drinking water from improved sources (for example, wells and piped water), representing about 16% of the population in the developing world [45]. Worldwide, 87% of people use improved drinking water sources, but in

sub-Saharan Africa, that number is just 60% [45]. "Safe water" describes a range of activities and interventions designed to promote access to clean, potable drinking water in low-resource settings, and often gets grouped with aspects of sanitation and hygiene to form the "WASH" (water, sanitation, and hygiene) sector, representing the inter-connectedness of these arenas. These topics areas are heavily linked to waterborne disease as well [46].

One of the problems associated with providing safe water is the finite quantity of fresh water on Earth. With increases in population, especially in already resource-poor developing countries, water quantity for drinking alone is strained. This limitation is even more pronounced when other water requirements, such as water for hygiene and food production, are taken into account. Minimal standards of water quantity and quality have been established by leading public health organizations, including the WHO and UNICEF, to give providers a target for water supply interventions in resource-poor conditions. These estimates of daily water use have changed over time. Early papers focused mainly on the minimum water supply per person necessary for survival, yielding estimates ranging from 1.8–5 liters (L) per person per day [47].The current SPHERE project standards take into account the range of uses and needs a family has for its household water, including not only water for ingestion, but also water for cleaning and basic hygiene and water for cooking. This estimate ranges from 7.5–15 L per person per day in complex humanitarian emergencies (including up to 20 L per person per day by UNHCR standards). [10].

Water usage is not governed solely by the quantity supplied by the individual water sources, but also by access and proximity to those sources. Water use varies with distance to the water source and climate, with hotter, tropical climates requiring more water consumption per capita [48]. The International Water Management Institute estimates that by 2025, 1.8 billion people in 17 countries will live in areas of physical water scarcity (insufficient water quantity in

the country itself, regardless of water development projects implemented). Another 24 countries will be classified as being in "economic water scarcity" (countries without sufficient economic resources to implement sufficient development projects to provide water to their people) [49] if they have not been already, further adding to the water quantity constraints.

Water source types and quality vary across developing nations, where water is more vulnerable to contamination, which reduces the number of safe and viable sources in any one community. Source types are generally divided into those providing surface water and those providing groundwater. Water source type can significantly affect the quality and safety (microbial contamination, mineral content, and turbidity, among others) of the water provided. Surface water (including rivers, ponds, streams, and even standing water) doesn't require drilling or high economic investment and is more easily accessible, thus it is more commonly used in developing and developed countries. However, surface water that is untreated, as is frequently observed in developing countries, has a higher potential for contamination and is less "safe" as a source than deeper groundwater. In urbanized overcrowded areas, without sewage systems, surface water can be perpetually contaminated as a result of poor sanitation. Surface water quality and quantity are also vulnerable to environmental events, such as storms, which can raise turbidity and bring organic and fecal matter into the water itself [50]. Sewage treatment can substantially reduce pathogen concentrations in the water, by up to 90–99% in some studies [50]. Groundwater sources comprise subterranean water that must be bored or drilled to access. In addition to being more difficult to access, groundwater sources have low potential for bacterial or other contamination due to the protective layers of soil and may be important in outbreaks, where maintaining clean communal water sources can be a concern [50, 51].

Piped water systems have been the norm in developed nations, with developing countries increasing their piped water supplies in recent decades. Piped water systems have traditionally

carried groundwater directly to houses or villages; however, they are not indestructible and have themselves been vulnerable to contamination and outbreaks [48]. "Distribution deficiencies" such as broken water system pipes have caused cross-connections and back-siphonage of the water, and were associated with 51% of distribution system outbreaks in the US from 1971 to 1998 [48, 52]. In developing countries, breaks in the water infrastructure can frequently cause much longer term problems due to inadequate resources for repair and maintenance. These breaks often spur further interest in establishing illegal connections to distribution systems in impoverished areas [48]. Water loss and contamination are major problems with piped water in these countries: water loss is estimated at 25–45% in those countries [53] and contamination has been implicated in many typhoid and cholera outbreaks in the past 25 years [54].

Boreholes, a less costly source of groundwater in developing settings, have been shown to yield better quality water than untreated surface water and may thus reduce disease transmission, but are also subject to seasonal fluctuations in quality because of contaminant seepage in the rainy seasons [55, 56]. A study in Nigeria found that residential areas with high borehole use reported low counts and proportions of cases of waterborne diseases, such as typhoid fever, bacillary dysentery, amoebic dysentery, and cholera [56]. They also noted that inadequate source water supply quantities, including those of boreholes, were a significant problem for residents. Where boreholes were compared to surface and municipal or mining compound water in Bindura, Zimbabwe, the chemical borehole-water profiles matched the quality of the municipal water. However, when bacterial water quality was assessed, borehole water exhibited seasonal fluctuations in fecal coliforms (safe water should be free of coliforms of all types). Municipal water supplies were free of coliforms year-round and river water was continuously contaminated with greater than 1,800 coliforms per 100 mL. Borehole water fell in between these sources in its fluctuations, ranging from being free of coliforms in the drier seasons to up to 240 coliforms per 100 mL in the rainy season. Seepage contamination of borehole

groundwater was suggested to explain this observed seasonality [55]. Use of borehole water when compared with surface water provides protection against waterborne diseases. A review of a cholera outbreak in a Mozambican refugee camp in 1988 found borehole water use was significantly more associated with control (non-disease) status rather than cases (those positive for cholera), who tended to use shallow wells. Cases were 4.5 times as likely as controls to have used shallow wells instead of boreholes (95% CI: 1.0–20.8) [57].

Cholera in Africa:

Cholera cases have been more concentrated in Africa in recent decades, possibly due to a number of socio-economic and environmental factors associated with disease prevalence, incidence, and transmission. Between 1995 and 2005, 632 outbreaks with 484,246 cases were reported, of which two-thirds of those outbreaks were reported in sub-Saharan Africa [2]. These outbreaks comprised 88% of the total number of cases reported worldwide during that time period [2], indicating that sub-Saharan African outbreaks were larger in scale in general and probably more widespread. Throughout sub-Saharan Africa, cholera outbreaks have higher reported death tolls as well: in 2005, the cholera CFR in sub-Saharan Africa was 1.8% in outbreaks, 3 times that of Asia (0.6%) [58]. In 2005 alone, 31 of the 40 sub-Saharan African countries (78%) reported native (in-country, instead of imported) cholera cases, and the incidence of 166 cholera cases per million people was 95 times higher than the next highest incidence in Asia (1.74 cases/million) and 16,600 times higher than that of Latin America (0.01 cases per million) [58]. South Africa reported the largest outbreak in the 1995–2005 period with 103,320 cases,[2], although WHO reported just over 86,000, with only 181 deaths (CFR = 0.21%) for the same outbreak [59]. In comparison, Zimbabwe's outbreak of 2008–2009 was at least as large, with the final June 2009 WHO report listing 98,424 *official* cases (officially documented at health centers or cholera

treatment centers, among others), and was much more severe, with 4,276 deaths (CFR = 4.34%, over 20 times that of South Africa) [60].

The causes and factors contributing to cholera's stronghold in Africa are varied and numerous. A 2002 paper by Naidoo and Patric identified two main contributors: displacement and environmental factors [61]. Large displacement often results from instability in a region; these populations require housing, and generally end up in resource-poor refugee camps. Instability in the region can also create complex humanitarian emergencies (CHEs), multi-faceted humanitarian crises in a country, region, or society that result in total or substantial breakdown of authority due to internal or external conflict, necessitating international response [62, 63]..

Complex Humanitarian Emergencies:

Since 1970, the start of the seventh cholera pandemic, there have been millions displaced across Africa primarily for political and socioeconomic reasons, generating complex emergency situations. Complex emergencies worldwide have been shown to be frequently associated with mass epidemics in the decade from 1995–2004, where 63% of complex emergencies worldwide included at least one epidemic [62]. Reasons for this association include longer duration of complex emergencies (as compared with natural disasters), higher prevalence of CHEs in Africa, where epidemic diseases are more endemic and poverty is more pervasive, and increased malnutrition and population migration because of CHEs, among others [62, 64]. CHEs can result from and lead to disastrous health and safety situations for a population, including administrative and political decay or collapse, intense and prolonged violence, genocide, disease outbreaks, strained resources, population displacement, wars, and high morbidity and mortality in the population [62, 65]. CHEs require a variety of responses in different arenas, including but not limited to water, sanitation, nutrition, communicable diseases, reproductive health, human rights, psychosocial/mental health, security, logistics, and infrastructure reconstruction [65].

Natural disasters can result in complex emergency situations in developing countries when structural damage or populations affected are extremely large. Natural disasters require international help to aid the country's response and can exacerbate CHEs in countries where high governmental instability exists. In addition, disasters, especially when combined with CHEs, strain health ministry resources within the country, exposing weaknesses within the public health infrastructure and leaving populations vulnerable to epidemics in those situations [65]. A significant number of epidemics have been documented during or immediately following CHEs, but the same magnitude of relationship has not been shown between epidemics and natural disasters alone [62]. Between 1995 and 2004, 63% of the largest CHEs had an epidemic concurrent to or just following the event, compared with just 23% of natural disasters [62]. Also, 87% of the CHEs documented during that period had one or more natural disasters in the recent history. Epidemics follow CHEs and follow the combined CHE-natural disaster situations much more frequently than natural disasters alone, underscoring the public health importance of CHEs in communicable disease prevention [62, 66].

CHEs have been environments for outbreaks associated with crowding and lack of safe water especially, with cholera frequently found in these situations[66], especially in Africa. In 1998, 500,000 refugees residing in the Democratic Republic of the Congo (DRC) finally returned to Rwanda after fleeing their country years earlier due to genocide, though they were unable to avoid a 14,000 case-outbreak of cholera in the camps themselves [61, 67]. Similarly, cholera outbreaks also occurred in populations of refugees or displaced persons in Somalia, Ethiopia, Malawi, Sudan, Rwanda, and Mozambique in 1997 [61, 68]. Larger outbreaks in Liberia (2003), Democratic Republic of the Congo (2000–2007, and still ongoing), and Kenya (2005 and 2009) have been prominent more recently. In Liberia, the civil war that had been ongoing since 1989 caused mass displacement of about 300,000 residents of Monrovia as rebels approached that capital. This quickly led to a shortage of clean water, poor sanitation, and mass overcrowding. In

addition, most healthcare facilities in the area were closed. Overall, over 17,000 suspected cholera cases were observed in 1993, though the CFR remained below emergency thresholds (<1%) due in large part to quick facility response in the form of inpatient hospitals, cholera treatment centers, ORS clinics, and outpatient clinics [69].

In the Democratic Republic of the Congo (DRC), Lake Kivu, bordered by the provinces of North and South Kivu, has been a site exhibiting the after-effects of natural disasters and complex humanitarian emergencies. In 1994, the fleeing of 500,000–800,000 Rwandan refugees to Lake Kivu resulted in over 62,000 cases of diarrheal disease, most *V. cholerae* O1. Overall attack rate estimates ranged from 7–16%. Estimates of the overall average CFR in this outbreak ranged from 3–7%, but the weekly CFR peaked at 22% in the middle of the outbreak [70]. These provinces have been subject to occupations, civil war, volcanic eruption, earthquakes, and population displacement in the 2000s as well, with about 5.4 million deaths in the area between 1998 and 2006 [71, 72]. Between 2000 and 2007, 18 large-scale population displacements took place, with 6 overlapping with ongoing cholera epidemics and 4 being followed by cholera outbreaks within 12 weeks, 2 of which were in IDP camps. While statistical time-series simulations in the study by Bompangue et al. [71] indicate that the occurrence of cholera epidemics after displacements or wars was no higher than would be expected by chance, there is still evidence to support the influence of these events in amplifying ongoing cholera outbreaks and contributing to the spread of cholera in the region [71]. Overall, at least 73,000 cases were observed around Goma and Lake Kivu between 2000 and 2007, leaving it as a continuously notable cholera-affected and war-torn region in the world [2, 71].

Kenya has also experienced political unrest, but also has felt the effects of CHEs in surrounding countries. All of these factors have contributed to the spread of cholera within Kenya, including cholera outbreaks in refugee camps in 2005 and 2009. In January–July, 2005,

cholera spread through 5 areas of the country from January to July, totaling about 1000 suspected cases and 25 dead. The largest of these affected areas was in the Kakuma refugee camp near the Sudanese border, where about half of the cases (estimates ranging up to 522) were found and 14 deaths occurred (CFR = 2.7%). Use of unsafe or untreated water, storage of drinking water in an open containers, poor sanitation and hygiene practices, and funeral attendance were all significant risk factors [73]. The movement and turnover of refugees into and out of the camp, as well as their camp conditions, probably added to these risk factors and contributed to the magnitude of the outbreak itself, exhibiting the vulnerability of refugee camps [73, 74]. Risk factors for cholera infection within the camp included sharing latrines among ≥3 households and having newly arrived at Kakuma camp [74]. Analysis of the outbreak data recommended targeting new arrivals at refugee camps for cholera prevention measures to control and prevent future outbreaks [74]. This outbreak, combined with those already mentioned, shows the widespread effects of CHEs on the spread of cholera, even across international borders.

Cholera outbreaks in the last 4 years have overwhelmingly occurred in countries without stable governments. Countries with outbreaks, especially high CFR outbreaks, were ranked higher in instability in the Failed States Index [75], compared with countries without cholera cases or with solely imported (non-native) cholera cases. Further, countries with higher numbers of internally displaced persons (IDPs) or refugees were more likely to have had cholera outbreaks in the past 4 years, or be areas of endemic cholera, indicating that failed states, frequently sites of CHEs, are significantly associated with epidemic cholera [76]. Many of the recent conflicts mentioned above were political in nature, leading to CHE situations themselves, and were corresponding with or leading to a cholera outbreak. Because of these conflicts, the importance of CHE conditions in cholera outbreaks is underscored, especially in developing countries with insufficient water supply and quality in the first place, such as Zimbabwe.

Zimbabwe:

Zimbabwe had the perfect mix of political unrest, instability, and infrastructural deficiencies, especially in water and sanitation, to engender the massive cholera outbreak seen in 2008–2009. It was one of the largest in African history, comprising over 100,000 cases and 4,000 deaths [3, 60, 77]. In 2008, an almost complete collapse of the economy exacerbated already strained water, sanitation, hygiene, and health infrastructure [78, 79]. Combined with a growing political crisis in the country and smaller cholera outbreaks seen annually in the country since 1998, the foundation existed for a disastrous situations [78, 80]. Violence and political tension began prior to the March 29th elections and continued between forces in the ruling party (Zimbabwe African National Union—Patriotic Front (ZANU-PF)) and those of the opposition (primarily Movement for Democratic Change (MDC)) through the run-off elections in late June. Additionally, the Government of Zimbabwe suspended relief work in the country on June 4th, 2008, and effectively crippled the Zimbabwean infrastructure to respond to humanitarian needs [80].

This governmental instability, combined with healthcare deficiencies in the country, resulted in mass population displacement out of the country (with the majority fleeing to South Africa) leaving dire conditions for those who stayed in their homes [78, 80]. Specifically, food security, healthcare, and water supply were the major problems in the country. Food security was mainly a concern because of hyperinflation, rising global food prices, inadequate governance, and decreased crop production within Zimbabwe, mostly due to inaccessible agricultural supplies in-country, adverse climate, and violence against farm workers [80]. Similarly, the economic conditions of Zimbabwe stifled wages for hospital workers and arrested supply lines for hospitals, causing many hospitals to shut down in mid to late 2008 [78, 79]. Harare, the capital of

Zimbabwe, had to close its 2 largest health facilities early on in the outbreak after they were faced with a shortage of supplies and worker strikes [78].

Insufficient water, sanitation, and hygiene conditions existed mainly because of the government's inability to sustain WASH projects and inadequate or halted funding for relief projects [80]. Zimbabwe's water supply and sanitation system, supported by UNICEF and other international organizations, had been one of the more expansive and effective systems in southern Africa before the outbreak. It had combined a mix of community-based and government-funded supply maintenance of existing boreholes and other water supply with the construction of new boreholes and other sources of safe water in rural regions to increase water access [77, 79-82]. The borehole project was designed to switchover the Zimbabwean water supply from 70% surface water between 1980 and 2000 to groundwater sources [83]. Insufficient funding in 2004 and 2007 borehole hand-pump repair efforts, however, left most of these rural communities without sources of "safer" groundwater and forced them to utilize surface water sources [83]. By 2007, the Zimbabwe National Water Authority estimated that there were between 40,000 and 50,000 boreholes in the country, which should have been sources of safe water for the surrounding communities, though at that time, most weren't functioning properly and needed to be repaired [83].

District-Level Water Access:

In rural districts, such as Gokwe North and Chivi, the conditions noted throughout Zimbabwe resulted in limited local water supply, forcing many households to turn to alternative, unsafe sources of drinking water around the time of the cholera outbreak in 2008. A 2008 evaluation of water supply in Gokwe North [81] determined that primary sources of water used included both "protected" and "unprotected" sources. Protected water sources included piped water, protected wells, and boreholes, while unprotected sources mainly comprised rivers,

unprotected wells, streams, and small earth dams. Upkeep of boreholes built as part of the national project's efforts was cut-off due to a lack of funds, resulting in inaccessible borehole water with broken hand-pumps and forcing households to find new, generally unsafe sources of water. Additionally, sanitation services were lacking for 75–85% of households in the province. Overall, at least 50% of boreholes were considered non-functioning before the cholera outbreak, and about 59% of households in the community had no easily accessible water points and instead were driven to use surface water sources in the form of ponds and rivers or at best, supplement surface water use with borehole use and vice versa [81]. While these data are from a case study of a district in Gokwe North, similar issues in resource development and management and access were seen across Gokwe North and in the province of Chivi [84].

Goals and Aims:

There is a need to identify whether functioning boreholes in communities such as Chivi or Gokwe North would have protected against the spread of cholera between villages and restricted the size of the outbreak within each village in Zimbabwe. Safe water reduces the transmission and risk of cholera incidence; however, any source effects observed are frequently overshadowed by water treatment options available in the community. Chivi and Gokwe North, for the most part, water treatment supplies, as in many outbreaks, the supply distribution followed the spread of cholera, rather than preceding it. This provides a unique opportunity to evaluate borehole effectiveness in the outbreak to determine if funding and support for source water repair in outbreak should be a priority. Thus, the goals of our study are to evaluate the effects of borehole water and protected source use on cholera infection in the cholera outbreak in Zimbabwe from 2008–2009. Our specific aims are to determine if differences in safe water access, including specifically borehole access, at the village level were associated with cholera spread between villages. Additionally, and more specifically, we would like to identify the

relationship between safe water use, including borehole use, and cholera attack rate within a village. The effect of safe source water on cholera attack rate has not been addressed through modeling in outbreak settings

Significance:

Our study will define the role of protected groundwater sources in cholera outbreak settings. We will be able to evaluate the effectiveness of functioning borehole projects in the specific provinces of Gokwe North and Chivi against the transmission of cholera in the outbreak setting. This research will also generalize to public health emergencies where groundwater sources are considered for rapid implementation as safe water sources in the midst of the outbreak. We have a unique opportunity to examine the effect of boreholes and protected wells on cholera transmission at the village level with little of the effects of point-of-use treatment. Boreholes can have an important and, as of yet, undefined role to play in preventing the spread of cholera in future outbreaks and complex humanitarian emergencies, and this study may serve to renew the advocacy efforts for safe water projects in developing countries by showing protection against cholera in an outbreak setting, making water source projects a priority with point-of-use treatment at the household level.

Materials and Methods:

Study Population: Our sampling frame consisted of all villages in 2 districts in Zimbabwe affected by the cholera outbreak: Chivi and Gokwe North. The study was conducted from August 20th to September 5th, 2009.

IRB: This survey was part of an outbreak investigation (Epi Aid) qualified as "non-research" and was exempt from the formal Internal Review Board approval process at Centers for Disease Control and Prevention. The protocol was evaluated by the Associate Director of Science for the National Center for Environmental Health (NCEH) for ethical clearance.

Personnel and Partners: Drs. Susan Cookson and Diane Morof from the International Emergency and Refugee Health Branch (IERHB) at the Centers for Disease Control and Prevention (CDC) led the data collection in Chivi, while Merlin and the -German Agro Action (GAA) conducted the survey in Gokwe North.

Selection of Districts: These districts were chosen based on cholera command and control center (C4) listings for total cholera cases and deaths in Zimbabwe. The C4, a collaboration between the World Health Organization (WHO), the Zimbabwean Ministry of Health and Child Welfare, and other partners, was the most reliable and accurate cholera case count database. Districts selected had to 1) have at least 1,500 cholera patients that reported to cholera treatment centers, health facilities, or cholera treatment units within that district (i.e. in the C4 database), 2) have over 100 cholera deaths in the community (i.e. deaths not in a health facility), 3) be geographically separated, 4) be small and accessible enough to allow for complete data collection within 2.5 weeks, and 5) have not been part of a study during the previous cholera outbreak.

Selection of Wards and Villages: Within Gokwe North, wards (sub-levels of districts with multiple villages) were selected using the total number of deaths from line lists at the cholera treatment unit (CTU). In Chivi, wards were selected using Cholera Treatment Center data listing of community deaths. All villages were surveyed in Chivi, while sampling was used for ward selection in Gokwe North. In the Copper Queen Central ward of Gokwe North, all 6 villages were surveyed. To get to the target of 10 villages per ward, an additional 4 of 13 small-scale farms listed on facility line lists from Copper Queen Central were selected, however only one of those farms was visited due to time limitations. For Goredema, Chireya IIa, Chireya IIIb wards in Gokwe North, 10 villages were randomly selected. To ensure that these villages existed and to determine their ward, the nursing staff of the CTU (Goredema) or the ward councilor (Chireya IIa and Chireya IIIb) were consulted. Villages that were unknown, misplaced, or incorrectly designated as villages were replaced with other randomly selected villages. Once the 10 villages were selected, the villages were split into two groups based on their geographical location. The teams were also given the names of two additional villages that should be surveyed in the event that the initial 10 villages did not provide sufficient data and there was sufficient time. This process was carried out in Chireya IIIb and Chireya IIa where more than 10 villages were visited.

Selection of Key Informants in each Village: Key informants (e.g. village health workers, ward council members, school teachers, religious leaders, traditional healers, and village kraal heads (village leader roles) in the village) were identified in each village. Ideally, 3 key informants per village were chosen to collaboratively fill out the survey, but often the village head, village health worker(s), and/or the school master were not present. In those cases, the one or two key informant types that were available, and in some cases peer educators, village secretaries, and choloroquine-holders, were surveyed.

Survey Instrument: The community survey consisted of 27 questions, divided into 8 sections: Village Characteristics, Population Changes, Security Issues, Communication, Water Questions, Healthcare Services, History of the Recent Cholera Outbreak (since October 2008), and Health Education. Village characteristics included ward and village population counts and GPS coordinates for the village, location of the nearest road, and various healthcare centers. The Population Changes section assessed if new settlement occurred in the past year. The Security Issues section aimed to evaluate village safety, while the Communication section described community mobilization and availability of radios and cell phones. Water Questions documented the village's main source(s) of drinking water, safe water usage, and the number, type, and status of safe boreholes serving the village. Healthcare services focused on estimated HIV/AIDS prevalence, while the History and Recent Cholera Outbreak section attempted to determine if a previous cholera outbreak had taken place and what kind of access to health services was available during the recent cholera outbreak. Finally, the Health Education section assessed community knowledge and practices related to homemade salt-sugar solutions (SSS).

Survey Methods: The survey questionnaire was created in English, translated to Shona, the local language, and then blind back-translated to English by a separate translator to assure accuracy. A total of 10 in-country interviewers underwent a 4-day training on the survey instrument, which included pilot testing of the survey instrument. Minimal modifications were made to the survey after the pilot test. In the main survey, verbal consent was obtained from all respondents and documented on the survey. Interviewers then conducted the surveys, generally in sheltered areas in a central location of the village. The surveys were completed on paper, and Epi Info 3.5.1 (CDC, Atlanta, GA) databases were created for survey entry. The Chivi data entry was completed by Drs. Diane Morof and Susan Cookson; however, the Gokwe North data was entered into Epi Info by a hired data entry team (Great Minds), with a second data entry required and completed by Dr. Diane Morof. A complete audit of 10% of the questionnaires was conducted

and differences in data entry were corrected. Less than 2% of data had errors in entry. All missing values were double checked with original surveys for accuracy. Responses were analyzed by village using SAS 9.2 (SAS Corporation, Cary, NC).

Analysis:

Case Definition: Classification of villages was based on line listings from a concurrent case/control patient survey conducted in the same villages, cholera treatment centers (CTCs), hospitals, and other healthcare centers included in the sample by the same research team. Cholera patients for line listings were defined as anyone greater than 5 years of age living in the respective district for at least a week prior to diarrheal onset who had 3 or more episodes of diarrhea in a 24 hour period between October 1, 2008 and July 28, 2009. Patients were not tested for *Vibrio cholerae* due to resource restrictions. Cholera attack rate (AR) was calculated per village by dividing the total number of cases from the line listings from the community by total village population (or imputed village population, see below). Villages were classified as (1) a 'case' if any C4-confirmed cholera cases were present and reported in the village or (2) a 'control' if cholera cases were not reported in the village during the outbreak between October 1, 2008 and July 28, 2009.

Imputation methods: Enumerators collected data for both the number of households and the total population in each village. The total population was missing for 37 (30.6%) of the villages; however, the total number of households were available for 120 (99.2%) of the villages following an initial survey assessment. In villages lacking population figures (necessary to calculate cholera attack rates for multiple linear regression techniques), we used simple and multiple imputation methods to generate population estimates. 'Simple imputation' methods imputed the unknown population from the number of households in the village multiplied by the average population per household from the villages with known populations in our sample. Multiple

imputation methods were based on the Markov Chain Monte Carlo (MCMC) method estimation of the village population and used a compilation of repeated drawings of replacements for the missing values [85].

Logistic Modeling: We used logistic regression to identify the best model for cholera case/control status of villages based on the use of borehole or surface water for drinking. In this analysis, we identified village population, the type of respondent, and the presence of a previous cholera outbreak in the year prior to the current outbreak as potential confounders for our model. We hypothesized that the quality of information may have varied by the type of respondent answering the survey, so dummy variables were constructed for 5 of the 6 categories of respondent: village health worker, village leader, religious leader, home-based care provider, and school headmaster ("no respondent listed" was used as the reference category), though only the village leader category was used because of stratification. Previous cholera outbreaks in the last year may have provided village members with residual immunity to cholera infection [11, 15, 86-88]. Therefore, the presence of previous outbreaks in the village was included as a potential confounder in the analysis. Population was included to allow for standardization across villages, as there was a large range of population counts across all villages. Interaction terms for all confounders with the borehole variable were constructed for the initial analysis. We used the "All-Subsets" analysis approach proposed by Kleinbaum and Klein [89] to evaluate interaction and confounding in constructing the model. We assessed confounding by sequentially removing variables or combinations of variables and comparing the resulting odds ratio (OR) with the OR of the model containing all variables (the "full" model). Bounds for model acceptance were estimates within 10% of the full model odds ratio (OR). Reduced and full models were judged by accuracy of the model OR and precision of the OR's 95% confidence interval.

Linear Modeling: We used multiple linear regression (MLR) techniques to model the cholera attack rate (AR) by village. To generate attack rates, we divided the total number of cholera cases (obtained by village line listings) by the population—estimated by simple imputation methods, MCMC multiple imputation methods or no imputation methods (only using complete observations)—and fit water source (environmental risk) parameters to the imputed ARs. We also controlled for village respondent type and previous cholera outbreaks (within one year), which may have conferred residual immunity and consequently affected the cholera attack rate within individual villages [11, 15, 86-88]. All model tests and estimates were evaluated at an alpha of 0.05.

In our MLR analysis, we combined water source variables to avoid over-stratification of source variables. The following water source types were considered: borehole (safe water source), other safe water (use of a safe water source besides borehole water, usually a protected hand-dug well), unsafe water (use of a single unsafe water source for the village: streams, unprotected hand-dug wells, scoop holes, rivers, and other similar surface water sources), and multiple unsafe water (concurrent use of 2 or more unsafe water sources during the outbreak). We built a full model around these 4 predictors—while controlling for previous cholera outbreaks in the village and respondent type (see above)—using each of the imputation techniques noted: no imputation, simple imputation, and multiple (MCMC) imputation. Sample size increased from 73 (no imputation model) to 102 (simple imputation model) to 120 (MCMC multiple imputation model) across imputation types. Effect modification and confounding were both assessed, in that order, by constructing variables for their potential interactions with the exposure water source variables and by inserting them into the model. Interaction was assessed using partial T tests/F tests [90], while confounding was assessed using methods described in Kleinbaum and Klein [89].

Results:

The goal of this study was to assess any differences in safe water source access that may have contributed to differences in cholera spread and attack rate across villages in the Chivi and Gokwe North districts. A total of 121 villages were included in the sample: 72 in Chivi and 49 in Gokwe North, totaling over 7,500 households. To ensure that our study districts—Gokwe North and Chivi—were comparable in population and households, we assessed population and cholera indicators in the overall study population and within each district's population (Table 1). Overall, villages varied widely in both population and cholera attack rates and case fatality rates, though the differences between village populations, number of households, average attack rates, and case fatality rates were not statistically significantly different between Chivi and Gokwe North. Overall village populations ranged from 34 to over 1600. Of the 121 villages surveyed in Chivi and Gokwe North districts, 105 villages (86.8%) had cholera cases present during the outbreak and 13 villages (10.7%) had no cholera cases (the remaining 3 villages were missing cholera status due to inconsistencies in the line listings and thus were left out of the analyses). Among the 105 villages with cholera, attack rates ranged from 2-578 persons per 1000, and 53 villages (50.5%) reported at least one death due to cholera. 51.4% of villages in Chivi reported at least one cholera death compared with 35.4% in Gokwe North. A total of 108 deaths were reported overall, 61.1% of which were in Chivi. Overall, because there were no significant differences in population and cholera statistics between the districts observed, the districts will be combined for further analysis.

Water:

To understand our study population's potential exposures to cholera, we examined the distribution of water source used and cholera prevalence and attack rate by source type (Table 2). Data from villages listing multiple 'main' sources of water were omitted from this analysis as

there was no reliable way to confidently characterize a single water source type from those surveys. Overall, only 37.0% villages reported using a safe water source as a 'main source' of water before the outbreak. Taps, as well as water trucking and safe water tankers, were not reported as major sources of drinking water for these rural villages; boreholes and protected wells served as the only reported sources of safe water in the village. Access to safe water sources was associated with about a 50% reduction in overall attack rate compared with unsafe water sources, though this difference was not detected as statistically significant at the 0.05 level. This difference was mostly due to the effect of boreholes, as protected wells exhibited an overall attack rate closer to that of unsafe sources. Villages with boreholes reported the lowest attack rates of all the drinking water source types, while villages using streams reported the highest. Villages employing safe water sources before the outbreak were more likely to have avoided the cholera outbreak compared with those using unsafe water sources (81% of safe water source villages reported cholera compared with 93% of unsafe water source villages), and this difference was borderline significant ($p = .053$). Suspicion of misclassification of protected wells led to reassignment of water source divisions into boreholes (safe water), wells—both protected and unprotected (unsafe), and all other (surface water sources). In this organization, boreholes exhibited the lowest attack rates of all sources, followed by well sources, and then surface water sources (data not shown), though this difference was not significant. This is further support for the use of the classification of boreholes as the major source of safe water (see below). Overall, though, there is evidence to suggest that type of water source affected cholera prevalence and attack rate through Chivi and Gokwe North.

Of the originally-classified safe water sources, boreholes (76% among the safe water sources) were the most commonly reported source of safe water for villages during the outbreak, as had been indicated in background investigation prior to the study's inception (data not shown). To further describe the frequency and distribution of boreholes across the region, data on

borehole availability, functionality, and use with regards to the cholera outbreak were collected (Table 3). Villages with primarily borehole-supplied water reported the lowest overall cholera proportions (82%) relative to villages with other water source types in the outbreak. However, borehole distribution and functionality was far from ubiquitous in the outbreak setting. Sixty-nine villages (57.0%) reported no boreholes serving the village. The 52 villages with boreholes were distributed unevenly by district between Chivi and Gokwe, with Chivi comprising 85% of the boreholes in the region and reporting many more villages with functioning boreholes during the outbreak. Of the 52 villages with boreholes present, 17 (32.7%) reported zero borehole hand-pumps working during the outbreak and 24 (46.2%) reported at least one borehole hand-pump as broken or not working during the outbreak (data not shown). In total, 31 out of the 79 boreholes sources (39.2%) were nonfunctional during the outbreak. Overall, there was evidence to suggest that use of boreholes and/or safe water sources for drinking water was associated with a meaningfully, though not quite statistically significant, lower attack rate compared with all other water source types used in villages during the outbreak, especially surface water sources and surface water overall.

Cholera and Water Source Analysis in Study Villages:

As an initial assessment of the effect of the water source exposure on cholera spread and burden in villages in Chivi and Gokwe North, we compared the proportion of study villages reporting cholera cases and their average attack rates across water source types: all protected sources, borehole water only, and unsafe water. We evaluated cholera presence using a Cochran-Mantel-Haenszel (CMH) chi-square analysis and cholera attack rates by an Analysis of Variance (ANOVA) test, estimating missing population estimates with simple imputation techniques. In both analyses villages with multiple listed “main sources” of water where all listed sources were of one status type (safe or unsafe alone) were included in this analysis, while (4) villages with

'mixed' (safe and unsafe) water source types were omitted from this analysis and all further analyses in this paper (Table 4). Villages using boreholes exhibited decreased odds of cholera prevalence and decreased AR compared with surface water sources, though all analyses were not significant at the 0.05 level. Additionally, villages using boreholes for drinking water saw a 67% reduction in the odds of having a case of cholera compared with villages using unsafe water (95% CI: 0.09–1.26, $p = 0.094$, data not shown), while villages using safe water in general saw a 69% decrease (95% CI: 0.09–1.06, $p = 0.053$, data not shown). AR analysis revealed that borehole-use villages with cholera had lower attack rates compared with well and surface water source categories, respectively, though neither were significant at the 0.05 level. When aggregated, cholera-affected villages using boreholes and protected wells reported a decrease of 23.5 cases per 1000 persons, on average, and those using just boreholes reported a 27.0 cases per 1000 reduction on average compared with unsafe water source villages (data not shown). All AR analyses were borderline significant ($0.10 < p < 0.13$), although it is important to note, as before, that cholera prevalence across villages during the outbreak was relatively high for all water source types ($\geq 80\%$). It was also interesting to note that cholera prevalence and attack rate do not follow the ordinal nature of the water source division. While use of unsafe water was expected to be associated with the highest proportion of case villages and the highest attack rates, followed by use of well water, and the borehole use, villages using well water actually reported the largest proportion of case villages and highest attack rates. Overall, though, we noted a marked yet not quite statistically significant reduction in cholera prevalence and attack rate by safe water, and specifically borehole, use compared with other unsafe water sources.

Cholera Modeling:

Logistic and multiple linear regression models were constructed to generate hypotheses for potential risk factors for cholera outcomes during this outbreak. Case/control villages were fit

in 2 logistic regression models below: one comparing safe (borehole and protected hand-dug well) vs. unsafe (surface and unprotected hand-dug well) water source use and one comparing borehole only vs. unsafe water use. The models evaluated interaction and confounding due to respondent type and previous cholera outbreaks in the village in the last year prior to the current one using simple and multiple imputation techniques (Tables 5 and 6). The final model parameters included the exposure variable of interest—safe or borehole water—and controlled for village population, respondent type, and the presence of a previous cholera outbreak in the village in the last year. Final logistic regression model analysis suggested that use of safe water sources as the main source of village drinking water (Table 5) was borderline significantly protective ($0.058 \leq p \leq 0.100$ for adjusted model estimates), indicating about a 70% reduction (on average across models) in the odds of cholera affecting a village using safe water sources compared to those using unsafe water sources. Presence of a previous cholera outbreak in the village was associated with about a 50% reduction in the odds of cholera compared with villages without a previous outbreak (results not significant) across all model types. When disaggregated from all safe water source use, the use of borehole water by villages (Table 6) was associated with a borderline significant ($0.062 \leq p \leq 0.162$ for adjusted model estimates) reduction in the odds of cholera of about 70-75% (across adjusted models) in villages using borehole water compared with villages using unsafe water sources. In both cases, we saw a marked and borderline significant (overall $0.058 \leq p \leq 0.162$ for adjusted model estimates) decrease in the odds of cholera in a village based on water source type. These data suggest that borehole use has a protective effect from cholera in the villages above that of safe water use alone, and that both reduce the odds of cholera; however, our study lacks sufficient power to detect these differences as significant at the 0.05 level.

We also evaluated cholera attack rates in outbreak-affected villages across water source type. To define the relationship between water source type and AR in a village, we fit ARs in a

multiple linear regression (MLR) model controlling for previous cholera outbreak and respondent type with water source divided into 4 variables comprising borehole use, other safe water use, unsafe water use from a single main source, and unsafe water use from multiple main sources, along with all relevant interactions. Using imputation techniques in the data for the model, no meaningful interaction nor confounding was detected among the variables in the model using the methods noted by Kleinbaum and Klein [89] to select variables in the model while keeping it hierarchically well formulated. We used the 'full' model without interaction terms (i.e. leaving all confounders in the model) for ease of comparison of results in differing imputation techniques. The final model assessment of collinearity by conditional index analysis revealed no significant collinearity issues in the model (all conditional indices (CI) < 30). Thus, the final model included 1) borehole, other safe water, unsafe, and multiple unsafe drinking water source variables as exposure variables of interest, 2) respondent type variable (defined as village leader/other), and 3) an indicator variable for the village having had a previous cholera outbreak in the last year (yes/no) (Table 7). Main water source was not significantly associated with cholera attack rate in one-way, unadjusted, and simple imputation-adjusted models when divided 4 ways, with the exception of the use of multiple unsafe sources of water, which was significant at the 0.05 level. Analysis of water source divided by borehole use, well use, and surface water use (as in Table 4) was marginally significant for wells as a risk factor for elevated attack rate ($p = 0.095$ in simple imputation adjusted model, data not shown). The multiple unsafe water source variable was not significant in the MCMC imputation analysis. Additionally, despite the loss of power and sample size due to sparseness and missing values in this data collected from an outbreak setting, the simple imputation model explained the data better than the mean of the outcome ($p = 0.023$), though only 14% of the variation in attack rate was explained by the model for water source type controlling for respondent type and previous cholera outbreaks. Across unadjusted and simple-imputation models, multiple sources of unsafe water were a significant risk factor for extreme

increases in attack rate above and beyond the effects of the other levels of the water source variable; however, water source type overall in each village does not seem to be significantly related to changes in attack rate in the study villages across all linear models constructed.

Discussion:

We evaluated borehole use on cholera spread and the magnitude of the outbreak on the village level in the Chivi and Gokwe North districts of Zimbabwe during a major outbreak. Survey results indicated that villages with a safe water sources were at a lower risk of cholera ($0.058 \leq p \leq 0.100$ across adjusted/unadjusted models). Furthermore, we found evidence to suggest that boreholes were the main source of safe water associated with reduced cholera spread across villages, though again this was not significant ($0.062 \leq p \leq 0.162$). In examining cholera attack rates, type of water source used by villages did not significantly reduce the attack rate in unadjusted and adjusted models. Finally, in comparing imputation techniques used in analyzing missing data frequently found in outbreak response and Epi-Aid studies, we noted differences in estimates that seemed to vary according to the model analyzed.

Villages using safe water sources were less likely to have one or more cholera cases, compared with villages using unsafe water sources and controlling for the size of the village. This result, though important, was only borderline statistically significant. In analyzing water sources at the village level, we are unable to assess individual household use, and it is likely that not all members of the village used the sources reported, whether safe or unsafe in nature, which would dilute the protective effects associated with safe water sources. Those in Gokwe North, in a study by Muguvu and Mutengu [81], reported frequent supplementation of borehole water with other, usually unsafe, source water due to the inaccessibility of sources. Despite this supplementation, in our study, the population using the reported protected sources was sufficient to buffer the community against cholera. While lacking power to detect this at the 0.05 level, this finding suggests an important role for water source maintenance in preventing outbreaks. Maintenance of safe water sources makes large, public, easily contaminated surface water sources, such as lakes or streams, a less available option in those villages. This is important as

these large surface water sources have been the subject of other outbreaks [71, 91]. The spread of cholera was aided by continued use of large water sources during multiple large outbreaks in Malawi and the DRC, among others [71, 91]. Additionally, the use of unprotected hand-dug wells has proved sufficient to be a source of contamination and further cholera outbreaks in point-source propagated spread patterns around the world [92]. Both of these scenarios could have appeared in these villages in Zimbabwe in 2008–2009 and affecting the spread of cholera throughout the region.

The ability of a large proportion of safe water-use villages to avoid the cholera outbreak is probably attributable to a lack of vibrios contaminating protected water sources, though we lack the microbiologic assays to detect this. It is plausible that infectious cholera bacteria did not penetrate these sources, particularly boreholes. Protected wells may have been slightly more susceptible to cholera bacteria despite their designation as a safe water source, since villages using 'protected' wells actually showed elevated attack rates over most other water source types, including boreholes. Additionally, these designations could have been misplaced, as we were unable to observe the sources of water noted for each village and verify their microbiologic quality and classification.

Among villages using safe water sources, borehole-supplied villages had the smallest proportion of cholera across villages, and experienced a reduction in cholera prevalence that approached significance when compared with villages with unsafe water sources, and even those with protected wells. These results imply that maintenance of borehole hand-pumps in Zimbabwe could have been an important component in protecting villages from cholera. These results have been supported by another outbreak investigation [57] that identified boreholes as a protective exposure when compared to shallow, unprotected wells. In an analysis conducted in Malawi, boreholes were significant ($p = 0.04$) with a very wide confidence interval when

compared with shallow well use in the outbreak (1.0–20.8). The Moren et al. study also supports the hypothesis that boreholes are the safest water source, when measured in terms of protection from cholera contamination through the soil. This data also indicates similar deficiencies in wells in outbreak or CHE settings as we see in our study (Table 4). In the Malawian refugee camp, none of the camp boreholes tested positive for cholera, while positive results were found for 50% of unprotected and 25% of protected wells [57]. Similar to our results from Gokwe North and Chivi, the protected wells in Malawi did not remain cholera free. This result may be explained by insufficient lining and sealing of the protected wells and/or better protection against cholera organism conferred through boreholes, possibly due to deeper drilling. In both Moren et al. and our study, we found evidence to suggest a protective effect of borehole water use in a cholera outbreak without the implementation of water treatment at point-of-use or the source. These findings support the use of boreholes as a safe water sources at the individual and village level.

In our study, the borderline protective effect of borehole water at the community level may also be due to water quantity as well as water quality. Borehole water may provide a more reliable source of water in quantities sufficient not only for drinking, but also for hygiene. In a 1991 meta-analysis, Esrey et al. [93] observed water quantity to be just as important, if not more important, as water quality in protecting against disease. They describe the effect of increased water supply interventions, which allow for increased household water use and therefore increased sanitation and hygienic practices [93]. Cairncross [94] noted that water source availability acts as a major link between water quantity and hygiene. Increasing the accessibility of a reliable water source to households in a community increases water quantity per family in that community significantly. With that increase, the family is able to practice better hygiene. Hygiene practices, including handwashing with soap, have been identified as significant protective factors in cholera outbreaks [95]. Over half of those surveyed by Maguvu and Mutengu [81] in Gokwe North in 2004 had to walk 5–10 km to get any source of water, as they

did not have a source of their own near them. Another 59% lacked any sort of sanitation coverage. Villages reporting functioning boreholes in our study had more reliable water overall, including water for basic hygiene and sanitation, than those with unsafe or surface water. Although undocumented in our survey, it is possible the villages with functioning boreholes may have had better hygiene or sanitation practices

Overall, cholera attack rates did not vary significantly across water source types (in multiple linear models; however, in bivariate analyses, villages using unsafe water sources exhibited twice the cholera attack of those using safe water sources (Table 2). Though attack rate risk factor modeling for cholera has been sparse in the literature, our study indicates that, among villages with cholera, safe water source use was associated with fewer cases of cholera in the village. This further emphasizes the role of borehole and hand-pump maintenance in outbreak settings. While we were unable to determine the exact environmental source of cholera in case villages, our results suggest that multiple unsafe water sources provided an opportunity for increased host-pathogen interaction and increased infection in the community. Provided they were functional, protected, and maintained, boreholes could have continued to provide safe water in villages with cholera despite the outbreak. The problem in case villages, then, would have been due to the lack of boreholes to supply the entire community and prevent supplementation by unsafe sources. This deficiency may have been caused by crowding or overpopulation in areas served by boreholes. As Fernandez et al. [77] note in their spatial analysis of the 2008–2009 Zimbabwe cholera outbreak in Harare, high density population and low socioeconomic status were important risk factors for cholera in the city. While we were able to control for village population, we have no data on the spatial distribution of people or water resources, which could have contributed greatly to spread of the disease in the height of the outbreak. Lower SES was also associated with poorer sources of water in Harare, and these risk factors may extend to the

rural communities as well. Insufficient safe water sources in communities can be supplemented with further borehole projects similar to those that were abandoned in the early to mid-2000s.

One of the strengths of our study is that it is one of the first to evaluate the effects of untreated safe water in an outbreak setting using retrospective village level surveillance. Often, studies in outbreaks evaluate water sources with point-of-use water treatments that mask any source water effects. Our study not only provides evidence that source maintenance in villages can limit the spread of cholera, but it also tests a new potential methodology for investigating outbreaks. Normal outbreak investigations involve case-control studies identifying common exposures in the affected population. Our village-level surveys allow a much larger population to be reached more efficiently. The fact that we observe a borderline significant reduction in cholera attack rates and village case/control status suggests the survey could be applied reliably at this level, though we need further evaluation.

There are several limitations to the methodology as well. A lack of household-level data means that our survey assumed the use of a single main source of water across a large community. This probably oversimplified the true nature of water use in the community and prevented us from documenting other water sources used, potentially including unsafe source use that would have reduced the effects seen in the data, possibly to the point of the observed non-significance.

While this methodology accounted for limited time for response and research, adequate care was not taken in completing the community level surveys as thoroughly once the case-control portion of the study was fully completed. We lacked sufficient time for a lengthy pilot test and were forced to deploy and collect information as quickly as possible from the available sources. Overall, incoming data was unclear in some places and sparse in others. A comparison of the variables from the main water source and percent of households with safe water revealed

vastly inconsistent data. This discrepancy led us to drop the variable assessing the percentage of households using safe water sources. Furthermore, a whole district (Kadoma) had to be dropped from our analysis due to uncertainty in water source variables, as it was an urban setting with less easily-identifiable water source data. Additional time may be required for future assessments using this methodology. Sero-surveys and microbiologic assays at the household level would help to make generalizations and conclusions at the village level more reliable.

It was especially difficult to gather village population and household estimates from village leaders, which resulted in large amounts of missing data and decreased power. To compare traditional 'field' methods in a controlled model, we used imputation methods to model cholera prevalence and attack rates in the villages. We are one of the first studies to compare no-imputation, simple imputation, and multiple imputation models for attack rate in an outbreak setting. In CHE settings, time and logistical constraints on rapid assessments frequently require quick estimation of village or camp populations. This estimation is generally crude: conducted by simple imputation with the average number of persons per household used, or some other standard (4 or 5 persons per household uniformly across a population, for example).

In comparing no-imputation, simple imputation, and multiple imputation models (Tables 5–7), we noted several key differences in estimates and standard errors. In logistic models (Tables 5 and 6), village population (imputed and unimputed values) was included as an independent variable for control purposes. Imputation of exposure variables in this manner is the traditional use of multiple imputation [96]. In these models, we observed little difference in estimates and standard errors and confidence intervals between unadjusted, simple, and multiple imputation models. As was expected, imputation methods—both simple and multiple—led to more precise confidence intervals by increasing the sample size. The larger the number of observations used, the more estimates and confidence intervals converged across model types.

We do note, however, that no-imputation adjusted models differed meaningfully from the imputation models in both analyses.

In our multiple linear regression models (Table 7), the village population was a component of the attack rate calculation, and therefore a dependent variable. This approach to imputation has not been tested in the literature. While it was difficult to judge the effects of multiple imputation on model estimates due to the lack of power and significance, we do note that p-values did not change markedly from adjusted no-imputation to simple imputation models. Simple imputation modeling, as expected, added markedly to the power and available sample size for the model. When moving to multiple imputation modeling, we observed a marked increase in p-values across estimates, especially in the variable representing multiple unsafe water sources, which, as noted, was associated with elevated cholera attack rates in all models except the multiple imputations MCMC model. This increase in p-values may be due to the initially sparse observations under this variable. While simple imputation methods did not correct the lack of observations with appropriate standard errors, multiple imputation methods yielded standard errors accounting for the imputation itself. It is possible that our model lacked sufficient power to detect differences in attack rates between villages, and therefore may not have been the ideal model to examine simple vs. multiple imputation methods. However, we aim to encourage and see a value in future research in this field with better methods in the literature.

The 2008–2009 cholera outbreak in Zimbabwe allowed us to evaluate the effects of isolated source water on cholera incidence and prevalence without confounding due to chlorine or other point-of-use treatments. While it has been standard outbreak response practice to provide these treatments to the population as soon as possible to ensure safe water for all at the household level, ‘safe’ water sources, particularly boreholes, showed borderline protective effects on cholera spread and burden at the village level when analyzed in aggregate without these point-of-use

treatments. These results provide support for the implementation of new or maintenance of sustained safe water source programs before and even during cholera outbreaks. In areas with endemic cholera or frequent outbreaks, these programs would lessen the population's morbidity and mortality relative to former outbreaks. In the face of economic and political hardships, priority must be given to existing community- or government-based borehole sustainability programs. This investment gives the population a chance at avoiding future waterborne disease outbreaks that may arise in CHE states.

In addition to taking a novel research approach to safe water in emergencies, this paper is the first to evaluate imputation methods for missing data in logistic and linear modeling of a cholera outbreak. We acknowledge that our analysis is largely incomplete and must therefore be viewed as a starting point for this approach, however, we do offer some conclusions from this work and to guide future research. In working with outbreak data, we must caution against the sole use of simple imputation analysis or conversion across seemingly related factors (such as number of households and population) to reach broad conclusions without first accounting for over-precise variances and standard errors. Multiple imputation methods have an important role in disease outbreaks due to the propensity for missing or incomplete data in this setting. Hastily constructed model estimates can impact resource allocation and ultimately human lives in the ongoing outbreak, thus, if undertaken, model construction must ensure validity across all biostatistical and epidemiologic arenas. We would encourage a collaborative effort by biostatisticians, epidemiologists, and other public health practitioners to re-evaluate imputation conversions from previous outbreaks in order to ensure best method practices and preparedness for future international emergencies.

Future Directions and Public Health Implications:

- Boreholes were a common main source of water due to the borehole development campaigns of the late 1990s and early 2000s that were abandoned in the mid-2000s. They were also the most common safe water source in the villages reporting People in these areas are dependent on boreholes for their water supply, especially safe water; these projects must be restarted and reinvigorated to replenish their supplies of safe water to prevent further diarrheal disease.

- Boreholes used as the main source of water in villages, among other safe water sources, were associated with reduced cholera prevalence across those villages compared with non-borehole-using villages. This result, however, was only borderline significant. Still, there is evidence to suggest that source water type may affect the spread of cholera from village to village during an outbreak, even in the absence of point-of-use and household treatment. Future studies should evaluate the effects of source water type on cholera prevalence at the community level.

- Borehole-use villages were also associated with reduced attack rates compared with villages not using boreholes. This result was also only borderline significant. This study provides evidence to suggest that source water type may affect the spread of cholera not only between villages but within villages, reducing the attack rate markedly. Still, future studies must evaluate this prospect through larger, more controlled, and more rigorous studies.

- General safe water source use by villages was associated with reductions in cholera prevalence and attack rate similar to those seen in only borehole-use villages that approached significance. Villages using safe water reported an average attack rate that was half that of villages using unsafe water sources in the cholera outbreak. Again these results were only borderline significant. These results have more widespread implications than those noted above, as safe source water, generally an easier component to regulate, may protect against cholera in villages.

- Missing values in the village population, as well as other variables, forced us to exclude over one-third of the data in our no-imputation models. This modeling yielded the lowest odds ratios (farthest from the null) of all models in cholera attack rate and prevalence analyses, but the widest confidence intervals. When assessing outbreak data, no-imputation modeling, though statistically valid, fails to utilize all the data and thus wastes the efforts of those collecting it, making it less practical in outbreaks and emergencies.

- Simple imputation modeling allowed the missing village population values to be estimated by multiplying the average persons per household by the number of households per village for the missing observations. This process resulted in a marked increase in the quantity of usable data, though 10-15 observations were still unusable. Logistic regression odd ratios and multiple linear regression estimates were similar to those above, though generally slightly closer to the null value and with narrower confidence intervals. While common and convenient, this method is not necessarily statistically valid and thus may yield inaccurate or imprecise estimates in models.

- Multiple imputation modeling allowed all missing village population and other missing data values to be approximated by the values of the other variables in that model. This technique also penalized the model estimates by elevating the standard errors to account for the imputing of the values, rather than their direct observation. While this method may dilute the p-values and confidence intervals in the data, it is the most statistically valid and conservative method and does allow for the use of all data with reliable imputation of missing values.

- We acknowledge that multiple imputation modeling of missing values in an outbreak setting is not the norm today; however, we would encourage future outbreak responders to consider multiple imputation techniques when estimating data in the outbreak. This technique allows one to more accurately assess the situation and incoming data, and to actually increase the precision of the estimates over those from simple imputation, the more prevalent technique in the field.

Tables and Figures:

Table 1: Village Demographics and Attack Rate by District

Descriptor	Overall		Chivi		Gokwe North	
	N	Mean (Std. Dev.)	N	Mean (Std. Dev.)	N	Mean (Std. Dev.)
Total Village Population*	120	338.09 (260.63)	71	384.46 (278.60)	49	270.90 (217.76)
Total Number of Households	120	62.84 (44.32)	71	74.55 (47.36)	49	45.88 (33.21)
Attack Rate†	103[1]	39.45 (67.89)	62	28.83 (40.89)	41	55.52 (93.62)
Case Fatality Rate†	104[1]	139.10 (214.65)	63	145.61 (219.46)	41	129.11 (209.33)

*The village population for villages with the total number of households reported but missing population data (N = 36) was estimated by simple imputation: multiplying the total households by 5.55, the average number of persons per household for the remaining 84 villages with population estimates.

†Estimates are cholera cases per 1000 persons.

[1] Villages may be missing population/household data yet have data for cases of cholera, thus AR and CFR numbers may not match

Table 2: Main source of pre-outbreak drinking water and cholera distribution for all villages

	Type of Water Source	Villages Using Source no. (%) (n = 100)	Villages with Cholera No. (%)	Mean Attack Rate by Source per 1000 (95% CI)*
Unsafe		63 (63.0)	69 (93.2)	47.9 (28.2, 67.5)
	Hand dug well	13 (13.0)	13 (100.0)	28.5 (15.3, 41.6)
	River	24 (24.0)	23 (95.8)	26.8 (11.9, 41.7)
	Sand abstraction	16 (16.0)	14 (87.5)	40.0 (15.2, 64.7)
	Scoop holes	8 (8.0)	7 (87.5)	50.8 (22.5, 79.1)
	Stream	2 (2.0)	2 (100.0)	141.3 (N/A)
Safe		37 (37.0)	30 (81.1)	24.4 (12.8, 35.9)
	Borehole	28 (28.0)	23 (82.1)	20.8 (10.0, 31.6)
	Protected well	9 (9.0)	7 (77.8)	35.9 (0.0, 78.3)

*Represents AR for villages with cholera only

Table 3: Borehole Functionality and Distribution Across Districts

District	Total Villages Reporting Sources	Villages with Borehole Water as Main Source (%)	Villages with 1+ Functioning Borehole Pre-Outbreak (%)	Total Number of Boreholes Prior to Outbreak	Total Number of Boreholes Functioning During Outbreak (%)	Borehole-use Villages with Cholera (%)	Attack Rate* in Borehole-Use Villages with Cholera
Chivi	72	25 (34.7)	44 (61.1)	70	41 (58.6)	21 (84.0)	17.6
Gokwe North	49	3 (6.1)	8 (16.3)	9	7 (77.8)	2 (66.7)	54.7
Overall	121	28 (23.1)	52 (43.0)	79	48 (60.8)	23 (82.1)	20.8

*Reported as cases per 1000 persons, reported among villages with cholera

Table 4: Cholera Spread and Attack Rate Analysis by Village Main Water Source Type

Main Water Source	Proportion (%) of Villages with Cholera	CMH X^2-statistic, p-value	OR (95% CI)	Average AR (per 1000 persons)*	Between-groups F-statistic, p-value	Overall F, p-value
Surface Water (reference)	47/52 (90.4%)	—	—	38.88	—	1.98, 0.144
Well Water	29/31 (93.6%)	0.25, 0.618	1.54 (0.28,8.48)	58.90	1.18, 0.282	
Boreholes	23/28 (82.1%)	1.12, 0.291	0.49 (0.13, 1.86)	20.83	2.73, 0.104	

*Among villages with cholera

Table 5: Logistic Regression Model Cholera Distribution by Village on Safe/Unsafe Water Use

	Unadjusted		**Adjusted** No Imputation (n= 74)**		**Adjusted** Simple Imputation (n= 109)**		**Adjusted** Multiple Imputation (n= 120)**	
Variable	**OR***	**95% CI**	**OR***	**95% CI**	**OR***	**95% CI**	**OR***	**95% CI**
Safe Main Water Source (n = 107)	0.311	(0.09, 1.06)	0.252	(0.05, 1.30)	0.294	(0.08, 1.04)	0.309	(0.09, 1.05)
Had Previous Cholera Outbreak (n = 117)	0.480	(0.05, 4.65)	N/A[†]	-	0.513	(0.05, 5.41)	0.564	(0.06, 5.71)
-2 log(Likelihood)			43.116		71.507			
Likelihood Ratio (statistic, p-value)			3.214	0.523	4.075	0.40		

*Villages with cholera during the outbreak are coded as 1, while villages without cholera are coded as 0

**also adjusted for respondent type (village leader/other, dichotomous) and village population (continuous)

[†]OR estimates are unavailable for the 'previous cholera' variable in this model as there are only 2 villages with previous cholera outbreaks in this subset of the data

Table 6: Logistic Regression Model Cholera Distribution by Village on Borehole/Unsafe Water Use

	Unadjusted		**Adjusted** No Imputation (n= 71)**		**Adjusted**Simple Imputation (n= 105)**		**Adjusted** Multiple Imputation (n=111)**	
Variable	**OR***	**95% CI**	**OR***	**95% CI**	**OR***	**95% CI**	**OR***	**95% CI**
Borehole Main Water Source (n = 102)	0.333	(0.09, 1.26)	0.239	(0.03, 1.78)	0.260	(0.06, 1.07)	0.303	(0.07, 1.26)
Had Previous Cholera Outbreak (n = 117)	0.480	(0.05, 4.65)	N/A[†]	-	0.539	(0.05, 6.07)	0.647	(0.06, 7.55)
-2 log(Likelihood)			32.474		59.967			
Likelihood Ratio (statistic, p-value)			2.939	0.57	5.049	0.28		

*Villages with cholera during the outbreak are coded as 1, while villages without cholera are coded as 0

**Also adjusted for respondent type (village leader/other, dichotomous) and village population (continuous)

[†]OR estimates are unavailable for the 'previous cholera' variable in this model as there are only 2 villages with previous cholera outbreaks in this subset of the data

Table 7: Risk Factors Associated with Cholera Attack Rate by Village

Variable	Unadjusted, No Imputation Beta	P-value	Adjusted*, No Imputation (n= 73) Beta	P-value	Adjusted, Simple Imputation (n= 102) Beta	P-value	Adjusted, Multiple Imputation (n= 120) Beta	P-value
Intercept			0.0212	0.500	0.0212	0.428	0.0224	0.856
Main Water Source: (n = 81)								
Borehole	-0.0283	0.219	0.0066	0.873	0.0111	0.727	0.0177	0.908
Other Safe Water	-0.0113	0.690	0.0257	0.571	0.0234	0.531	0.0183	0.913
Unsafe Water (single source)	-0.0047	0.781	0.0261	0.456	0.0239	0.413	0.0568	0.610
Unsafe Water (multiple sources)	0.0681	0.005	0.0911	0.025	0.0919	0.008	0.0823	0.599
Previous Cholera Outbreak (n = 81)	-0.0270	0.622	-0.0212	0.709	-0.0057	0.867	0.0397	0.900
Overall F Test (statistic, p-value)			1.60,	0.160	2.58,	0.023		
R-Squared Value			0.127		0.140			

*Adjusted for respondent type (village leader/other)

Appendix A: IRB Clearance

March 3, 2011

School of Public Health

Emory University

Atlanta, Ga. 30322

RE: Determination: No IRB Review Required
Environmental Factors in Borehole Use Affecting the Cholera Outbreak in Zimbabwe (working title)

PI: Berendes, David

Dear Mr. Berendes:

Thank you for requesting a determination from our office about the above-referenced project. Based on our review of the materials you provided, we have determined that it does not require IRB review because it does not meet the definition(s) of "research" involving "human subjects" or the definition of "clinical investigation" as set forth in Emory policies and procedures and federal rules, if applicable. Specifically, in this project, you are conduct secondary data analysis on data collected by the Centers for Disease Control and Prevention. This data was collected in September 2009 in the context of a cholera outbreak investigation in three districts in Zimbabwe. This data is de-identified and you will have no link to any identifiers. Please note that in future projects if you are not positive whether IRB review and approval is necessary please contact the IRB of review prior to initiating any research related activities.

This determination could be affected by substantive changes in the study design, subject populations, or identifiability of data. If the project changes in any substantive way, please contact our office for clarification.

Thank you for consulting the IRB.

Sincerely,

Emily Sanders, BS

Education and QA Analyst

This letter has been digitally signed

References:

1. **Sack, D.A., et al., *Cholera.* Lancet, 2004. 363(9404): p. 223-33.**
2. **Griffith, D.C., L.A. Kelly-Hope, and M.A. Miller, *Review of reported cholera outbreaks worldwide, 1995-2005.* Am J Trop Med Hyg, 2006. 75(5): p. 973-7.**
3. **WHO, *Cholera, 2009*, in *Weekly epidemiological record*, WHO, Editor. 2010, WHO: Geneva. p. 293-308.**
4. **WHO, *Cholera, 2006*, in *Weekly epidemiological record*, WHO, Editor. 2007, WHO: Geneva. p. 273-284.**
5. **WHO, *Cholera, 2007*, in *Weekly epidemiological record*, WHO, Editor. 2008, WHO: Geneva. p. 269-284.**
6. **WHO, *Cholera, global surveillance summary, 2008*, in *Weekly epidemiological record*, WHO, Editor. 2009, WHO: Geneva. p. 309-324.**
7. **Zuckerman, J.N., L. Rombo, and A. Fisch, *The true burden and risk of cholera: implications for prevention and control.* Lancet Infect Dis, 2007. 7(8): p. 521-30.**
8. **WHO, *Cholera unveiled*, G.T.F.o.C. Control, Editor. 2003, WHO.**
9. **Lanata, C., W Mendoza, and RE Black, *Improving Diarrhoea Estimates*, W.H.O.C.a.A.H.D.M.a.E. Team, Editor. 2002, World Health Organization: Geneva, Switzerland.**
10. **The Sphere Project, *Humanitarian Charter and Minimum Standards in Disaster Response*, I. SCHR, VOICE, and ICVA, Editor. 2004, Oxfam Publishing: Oxford.**
11. **Glass, R.I., et al., *Endemic cholera in rural Bangladesh, 1966-1980.* Am J Epidemiol, 1982. 116(6): p. 959-70.**
12. **Khan, M. and M. Shahidullah, *Cholera due to the E1 Tor biotype equals the classical biotype in severity and attack rates.* J Trop Med Hyg, 1980. 83(1): p. 35-9.**
13. **WHO, *Cholera: Fact sheet no. 107*, in *Cholera*, WHO, Editor. 2000.**
14. **WHO, *Cholera, 2005*, W.E.R. 2006, Editor. 2005. p. 297-307.**
15. **Nelson, E.J., et al., *Cholera transmission: the host, pathogen and bacteriophage dynamic.* Nat Rev Microbiol, 2009. 7(10): p. 693-702.**
16. ***Vibrio cholerae and Cholera: Molecular to Global Perspectives*, ed. I.K. Wachsmuth, Blake, P.A. and Olsvik, O (eds). 1994, Washington D.C.: ASM.**
17. **Reidl, J. and K.E. Klose, *Vibrio cholerae and cholera: out of the water and into the host.* FEMS Microbiol Rev, 2002. 26(2): p. 125-39.**
18. **Colwell, R.R., *Global climate and infectious disease: the cholera paradigm.* Science, 1996. 274(5295): p. 2025-31.**
19. **Huq, A., et al., *Detection of Vibrio cholerae O1 in the aquentic environment by fluorescent-monoclonal antibody and culture methods.* Appl Environ Microbiol, 1990. 56(2370-3).**
20. **Tamplin, M.L., et al., *Attachment of Vibrio cholerae serogroup O1 to zooplankton and phytoplankton of Bangladesh waters.* Appl Environ Microbiol, 1990. 56(6): p. 1977-80.**
21. **Shukla, B.N., D.V. Singh, and S.C. Sanyal, *Attachment of non-culturable toxigenic Vibrio cholerae O1 and non-O1 and Aeromonas spp. to the aquatic arthropod Gerris spinolae and plants in the River Ganga, Varanasi.* FEMS Immunol Med Microbiol, 1995. 12(2): p. 113-20.**
22. **Halpern, M., et al., *Chironomid egg masses as a natural reservoir of Vibrio cholerae non-O1 and non-O139 in freshwater habitats.* Microb Ecol, 2004. 47(4): p. 341-9.**
23. **Huq, A., et al., *Ecological relationships between Vibrio cholerae and planktonic crustacean copepods.* Appl. Environ. Microbiol. Ecol., 1983. 47: p. 341-349.**
24. **Abd, H., A. Weintraub, and G. Sandstrom, *Intracellular survival and replication of Vibrio cholerae O139 in aquatic free-living amoebae.* Environ Microbiol, 2005. 7(7): p. 1003-8.**

25. **Islam, M.S., et al., *Biofilm acts as a microenvironment for plankton-associated Vibrio cholerae in the aquatic environment of Bangladesh.* Microbiol Immunol, 2007. 51(4): p. 369-79.**
26. **Alam, M., et al., *Viable but nonculturable Vibrio cholerae O1 in biofilms in the aquatic environment and their role in cholera transmission.* Proc Natl Acad Sci U S A, 2007. 104(45): p. 17801-6.**
27. **Colwell, R., PR Brayton, DJ Grimes, DR Roszak, SA Huq, and LM Palmer, *Viable, but non-culturable Vibrio cholerae and related pathogens in the environment: implications for the release of genetically engieneered microorganisms.* Bio/Technology, 1985. 3: p. 817-20.**
28. **Xu, H., N Roberts, FL Singleton, RW Atwell, DJ Grimes, and RR Colwell, *Survival and viability of non-culturable Escherichia coli and Vibrio cholerae in the estuarine and marine environment.* Microb. Ecol., 1982. 8: p. 313-23.**
29. **Faruque, S.M., M.J. Albert, and J.J. Mekalanos, *Epidemiology, genetics, and ecology of toxigenic Vibrio cholerae.* Microbiol Mol Biol Rev, 1998. 62(4): p. 1301-14.**
30. **Colwell, R.R., J. Kaper, and S.W. Joseph, *Vibrio cholerae, Vibrio parahaemolyticus, and other vibrios: occurrence and distribution in Chesapeake Bay.* Science, 1977. 198(4315): p. 394-6.**
31. **Garay, E., A. Arnau, and C. Amaro, *Incidence of Vibrio cholerae and related vibrios in a coastal lagoon and seawater influenced by lake discharges along an annual cycle.* Appl Environ Microbiol, 1985. 50(2): p. 426-30.**
32. **Islam, M.S., B.S. Drasar, and R.B. Sack, *The aquatic flora and fauna as reservoirs of Vibrio cholerae: a review.* J Diarrhoeal Dis Res, 1994. 12(2): p. 87-96.**
33. **Colwell, R., and WM Spira, *The ecology of Vibrio cholerae*, in *Cholera*, D. Barua, and WB Greenough III, Editor. 1992, Plenum: New York. p. 107-27.**
34. **Karaolis, D.K., et al., *A Vibrio cholerae pathogenicity island associated with epidemic and pandemic strains.* Proc Natl Acad Sci U S A, 1998. 95(6): p. 3134-9.**
35. **Waldor, M.K. and J.J. Mekalanos, *Lysogenic conversion by a filamentous phage encoding cholera toxin.* Science, 1996. 272(5270): p. 1910-4.**
36. **Butler, T., M. Arnold, and M. Islam, *Depletion of hepatic glycogen in the hypoglycaemia of fatal childhood diarrhoeal illnesses.* Trans R Soc Trop Med Hyg, 1989. 83(6): p. 839-43.**
37. **Carpenter, C.J., A Mondal, RB Sack, PE Dans, SA Wells, *Clinical studies in Asiatic cholera.* Bull Johns Hopkins Hosp, 1966. 118: p. 174-96.**
38. **Glass, R., Claeson M, Blake PA, Waldman RJ, and Pierce NF, *Cholera in Africa: lessons on transmission and control for Latin America* Lancet, 1991. 338: p. 791-5.**
39. **Stock, R., *Cholera in Africa: diffusion of the disease 1970-1975 with particular emphasis on West Africa*, in *Special Report 3*. 1976, International African Institute: London.**
40. **Shapiro, R.L., et al., *Transmission of epidemic Vibrio cholerae O1 in rural western Kenya associated with drinking water from Lake Victoria: an environmental reservoir for cholera?* Am J Trop Med Hyg, 1999. 60(2): p. 271-6.**
41. **Gunnlaugsson, G., Einarsdottir J, Angulo FJ, Mentambanar SA, Passa A, and Tauxe RV, *Funerals during the 1994 cholera epidemic in Guinea-Bissau, West Africa: the need for disinfection of bodies of persons dying of cholera.* Epidemiol Infect, 1998. 120(1): p. 7-15.**
42. **Nelson, E.J., et al., *Transmission of Vibrio cholerae is antagonized by lytic phage and entry into the aquatic environment.* PLoS Pathog, 2008. 4(10): p. e1000187.**
43. **Merrell, D.S., et al., *Host-induced epidemic spread of the cholera bacterium.* Nature, 2002. 417(6889): p. 642-5.**
44. **Hartley, D.M., J.G. Morris, Jr., and D.L. Smith, *Hyperinfectivity: a critical element in the ability of V. cholerae to cause epidemics?* PLoS Med, 2006. 3(1): p. e7.**

45. WHO and UNICEF, *Progress on Sanitation and Drinking-Water: 2010 Update*, in *JMP Reports*, WHO/UNICEF, Editor. 2010, World Health Organization: Geneva.
46. WSSCC. *Water Supply and Sanitation Collaborative Council*. 2011.
47. Gleick, P.H., *Basic water requirements for human activities: Meeting basic needs.* Water International, 1996. 21(2): p. 83-92.
48. Moe, C.L. and R.D. Rheingans, *Global challenges in water, sanitation and health.* J Water Health, 2006. 4 Suppl 1: p. 41-57.
49. Seckler, D.D.A., D Molden, R deSilva, and R Barker, *World water demand and supply, 1990-2025: Scenarios and issues*, in *International Water Management Institute Report 19*, I.W.M. Institute, Editor. 1998, International Water Management Institute: Colombo, Sri Lanka.
50. Medema, G., S Shaw, M Waite, M Snozzi, A Morreau, and W Grabow, *Catchment chracteristics and source water quality*, in *Assessing Microbial Safety on Drinking Water. Improving Approaches and Method*, W.a. OECD, Editor. 2003, IWA Publishing: London. p. 111-158.
51. Cabral, J.P., *Water microbiology. Bacterial pathogens and water.* Int J Environ Res Public Health, 2010. 7(10): p. 3657-703.
52. Craun, G.F., and RL Calderon, *Waterborne disease outbreaks caused by distribution system deficiencies.* Amer. Water Works Assoc., 2001. 93(9): p. 64-75.
53. WHO, *Leakage Management and Control-A Best Training Manual*, W.H. Organization, Editor. 2001, World Health Organization: Geneva, Switzerland.
54. Renkevich, V., et al., *Multi-City Water Distribution System Assessment*. 1998, USAID/CDC: Kazakhstan.
55. Nyati, H., *Evaluation of the microbial quality of water supplies to municipal, mining and squatter communities in the Bindura urban area of Zimbabwe.* Water Sci Technol, 2004. 50(1): p. 99-103.
56. Oguntoke, O., O.J. Aboderin, and A.M. Bankole, *Association of water-borne diseases morbidity pattern and water quality in parts of Ibadan City, Nigeria.* Tanzan J Health Res, 2009. 11(4): p. 189-95.
57. Moren, A., et al., *Practical field epidemiology to investigate a cholera outbreak in a Mozambican refugee camp in Malawi, 1988.* J Trop Med Hyg, 1991. 94(1): p. 1-7.
58. Gaffga, N.H., R.V. Tauxe, and E.D. Mintz, *Cholera: a new homeland in Africa?* Am J Trop Med Hyg, 2007. 77(4): p. 705-13.
59. WHO (2001) *2001-Cholera in South Africa*. Global Alert and Response (GAR).
60. WHO (2009) *Cholera in Zimbabwe - update 4*. Global Alert and Response (GAR).
61. Naidoo, A. and K. Patric, *Cholera: a continuous epidemic in Africa.* J R Soc Promot Health, 2002. 122(2): p. 89-94.
62. Spiegel, P.B., et al., *Occurrence and overlap of natural disasters, complex emergencies and epidemics during the past decade (1995-2004).* Confl Health, 2007. 1: p. 2.
63. Office for the Coordination of Humanitarian Affairs, *OCHA Orientation Handbook on Complex Emergencies*, U.N. OCHA, Editor. 1999, United Nations.
64. Loretti, A. and Y. Tegegn, *Disasters in Africa: old and new hazards and growing vulnerability.* World Health Stat Q, 1996. 49(3-4): p. 179-84.
65. Burkle, F.M., *Lessons learned from and future expectations of complex emergencies.* West J Med, 2000. 172(1): p. 33-8.
66. Watson, J.T., M. Gayer, and M.A. Connolly, *Epidemics after natural disasters.* Emerg Infect Dis, 2007. 13(1): p. 1-5.
67. CDC, *Cholera Outbreak among Rwandan Refugees--Democratic Republic of Congo, April 1997.* MMWR, 1998. 47(19): p. 389-391.

68. World Health Organization, *Cholera in 1997*, in *Weekly Epidemiol Rec. 1997*, WHO, Editor. 1997.
69. CDC, *Cholera Epidemic After Increased Civil Conflict---Monrovia, Liberia, June--September 2003.* MMWR, 2003. 52(45): p. 1093-5.
70. Goma Epidemiology Group, *Public health impact of Rwandan refugee crisis: what happened in Goma, Zaire, in July, 1994?* Lancet, 1995. 345: p. 339-44.
71. Bompangue, D., et al., *Cholera epidemics, war and disasters around Goma and Lake Kivu: an eight-year survey.* PLoS Negl Trop Dis, 2009. 3(5): p. e436.
72. Coghlan, B., et al., *Mortality in the Democratic Republic of Congo: a nationwide survey.* Lancet, 2006. 367(9504): p. 44-51.
73. Mugoya, I., et al., *Rapid spread of Vibrio cholerae O1 throughout Kenya, 2005.* Am J Trop Med Hyg, 2008. 78(3): p. 527-33.
74. Shultz, A., et al., *Cholera outbreak in Kenyan refugee camp: risk factors for illness and importance of sanitation.* Am J Trop Med Hyg, 2009. 80(4): p. 640-5.
75. The Fund for Peace. *The Fund for Peace: The Failed States Index*. The Failed States Index 2010; Available from: http://www.fundforpeace.org/web/index.php?option=com_content&task=view&id=99&Itemid=140.
76. Berendes, D.a.T.H., *Cholera and Failed States*. 2010, CDC/Emory University: Atlanta.
77. Fernandez, M.A.L., PR Mason, H Gray, A Bauernfeind, JF Fesselet, and P Maes, *Descriptive spatial analysis of the cholera epidemic 2008-2009 in Harare, Zimbabwe: a secondary data analysis.* Transactions of the Royal Society of Tropical Medicine and Hygiene, 2011. 105: p. 38-45.
78. Rosborough, S., *World Update: Cholera Crisis in Zimbabwe.* Disaster Medicine and Public Health Preparedness, 2008. 3(1): p. 11-12.
79. Medecins Sans Frontieres, *Beyond Cholera: Zimbabwe's worsening crisis*, MSF, Editor. 2009, MSF.
80. USAID, *Zimbabwe-Complex Emergency: Situation Report #2, Fiscal Year 2009*, in *Situation Report*, USAID, Editor. 2009, Bureau for Democracy, Conflict, and Humanitarian Assistance, Office of US Foreign Disaster Assistance.
81. Maguvu, E., and S Mutengu, *An Investigation into the Factors Limiting Effective Water Supply in Rural Areas of Zimbabwe: A Case of Zhomba in Gokwe North District.* Journal of Sustainable Development in Africa, 2008. 10(1): p. 120-139.
82. Derman, B., Ferguson A, and Gonese F, *Decentralization, Devolution, and Development: Reflections on the Water Reform Process in Zimbabwe*, in *BASIS*, B.B.A.a.S.I.M. Systems, Editor. 2000, Center for Applied Social Sciences, University of Zimbabwe.
83. Tsiko, S. (2008) *Zimbabwe: Water Shortage Stunts Growth*. Black Star News.
84. Nemarundwe, N., and M Mutamba, *Action planning and adaptive management of natural resources in semiarid environments: Experiences from Chivi District, Zimbabwe*, in *Coping Amidst Chaos: Studies on Adaptive Collaborative Management from Zimbabwe*, A. Mandondo, Prabhu R, and Matose F, Editor. 2008, Center for International Forestry Research.
85. Rubin, D., *Multiple Imputation for Nonresponse in Surveys*. 1987, New York: John Wiley and Sons, Inc.
86. Kaper, J.B., J.G. Morris, Jr., and M.M. Levine, *Cholera.* Clin Microbiol Rev, 1995. 8(1): p. 48-86.
87. Levine, M.M., et al., *Immunity of cholera in man: relative role of antibacterial versus antitoxic immunity.* Trans R Soc Trop Med Hyg, 1979. 73(1): p. 3-9.

88. **Clemens, J., et al., *Biotype as determinant of natural immunising effect of cholera.* Lancet, 1991. 337: p. 883-4.**
89. **Kleinbaum, D., and M Klein, *Logistic Regression: A Self-Learning Text*. Third ed. Statistics for Biology and Health, ed. S. Science. 2010, New York: Springer.**
90. **Kleinbaum, D., Kupper LL, Nizam A, and KE Muller, *Applied Regression Analysis and Multivariable Methods, 4th Edition*. 2008, Belmont: Thomson: Brooks/Cole.**
91. **Swerdlow, D.L., et al., *Epidemic cholera among refugees in Malawi, Africa: treatment and transmission.* Epidemiol Infect, 1997. 118(3): p. 207-14.**
92. **Das, A., et al., *An outbreak of cholera associated with an unprotected well in Parbatia, Orissa, Eastern India.* J Health Popul Nutr, 2009. 27(5): p. 646-51.**
93. **Esrey, S.A., et al., *Effects of improved water supply and sanitation on ascariasis, diarrhoea, dracunculiasis, hookworm infection, schistosomiasis, and trachoma.* Bull World Health Organ, 1991. 69(5): p. 609-21.**
94. **Cairncross, S., *More water: better health.* People Planet, 1997. 6(3): p. 10-1.**
95. **Sasaki, S., Suzuki H, Igarashi K, Tambatamba B, and Mulengu P, *Spatial Analysis of Risk Factor of Cholera Outbreak for 2003-2004 in a Peri-urban area of Lusaka, Zambia.* Am J Trop Med Hyg, 2008. 79(3): p. 414-21.**
96. **Little, R., and DB Rubin, *Statistical Analysis with Missing Data*. 1987, New York: Wiley.**