Walden University

ScholarWorks

Walden Dissertations and Doctoral Studies

Walden Dissertations and Doctoral Studies Collection

2020

Population-Level Determinants of Cholera Incidence in African countries

Ngoy Nsenga
Walden University

Walden University

College of Health Sciences

This is to certify that the doctoral dissertation by

Ngoy Nsenga

has been found to be complete and satisfactory in all respects,
and that any and all revisions required by
the review committee have been made.

Review Committee
Dr. Aimee Ferraro, Committee Chairperson, Public Health Faculty
Dr. Xianbin Li, Committee Member, Public Health Faculty
Dr. Pelagia Melea, University Reviewer, Public Health Faculty

Chief Academic Officer and Provost
Sue Subocz, Ph.D.

Walden University
2020

Abstract

Population-Level Determinants of Cholera Incidence in African countries

by

Ngoy Nsenga

MD, University of Lubumbashi, 1991

MPH, University of Kinshasa, 2001

Dissertation Submitted in Partial Fulfillment

of the Requirements for the Degree of

Doctor of Philosophy

Public Health

Walden University

June 2020

Abstract

Cholera remains endemic in Africa, with limited access to safe drinking water and inadequate sanitation as 2 of the main drivers of its dissemination. Few studies have examined the impact of health system capacity, even though it plays an important role in prevention, early detection, and response to a cholera outbreak. Grounded in the ecosocial theory of infectious disease dissemination, this quantitative ecological study explored the effect of social vulnerability (as measured by access to safe drinking water and sanitation, rate of open defecation, poverty, income inequality, gender inequality, and adult literacy) and health system (as measured by health financing and density of human resources for health) capacity on incidence of cholera in the 47 countries of the African region of the World Health Organization.

Logistic regression results showed that only access to improved sanitation [$p < .05$; OR = .904; 95% CI: .823 – .992; *N*= 47], rate of open defecation [$p < .05$; OR = .894; 95% CI: .822 – .973; *N*= 47], and health system capacity [$p < .05$; OR = .792; 95% CI: .630 – .995; *N*=47] had a statistically significant association with incidence of cholera. The components of social vulnerability [$p < .05$; OR = 1.080; 95% CI: 1.004 – 1.162; *N*=47] and the interaction between social vulnerability and health system capacity [$p < .05$; OR = 1.004; 95% CI: 1.002 – 1.009; *N*= 47] were also significantly associated with the outcome. These findings can impact social change by guiding the development of effective multisectoral programs for cholera prevention and elimination.

Population-Level Determinants of Cholera Incidence in African countries

by

Ngoy Nsenga

MD, University of Lubumbashi, 1991

MPH, University of Kinshasa, 2001

Dissertation Submitted in Partial Fulfillment

of the Requirements for the Degree of

Doctor of Philosophy

Public Health

Walden University

June 2020

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction to the Study

Cholera remains a major public health problem in developing African countries. As many as three to four million cases of cholera, causing 21,000 to 143,000 deaths occur worldwide every year (Ali et al., 2012; Ali, Nelson, Lopez, & Sack, 2015). Of these cases, the majority are reported in African countries. In 2015, five African countries accounted for 80% of cases of cholera (World Health Organization [WHO], 2016a). Records from WHO show that cholera outbreaks primarily occur in the most impoverished countries of Africa.

Cholera outbreaks in impoverished countries overstretch the already weak health systems and divert and deplete resources, further deepening poverty (Kirigia et al., 2009). For that reason, the Global Task Force on Cholera Control has developed a roadmap aiming to eliminate cholera by 2030. The roadmap calls on countries to implement sound and evidence-based cholera control policies and plans with support from development partners (Global Task Force on Cholera Control, 2017). However, the task force did not define selection criteria to prioritize the countries that will benefit from international support in the development of robust cholera control strategies. Moreover, most cholera control interventions listed by the task force focus on individual-level risk factors, such as access to safe drinking water and improved sanitation. Meanwhile, studies have shown that population-level factors, such as poverty, inequality, and adult literacy, play a significant role in the emergence or dissemination and amplification of cholera outbreaks (Jutla et al., 2013; Root, Rodd, Yunus, & Emch, 2013). Further, the ability of a country to control outbreaks depends on the health systems for early detection, adequate health care services, and response coordination (Gostin & Friedman, 2015).

Therefore, the identification of population-level socioeconomic and environmental factors of vulnerability to cholera can be crucial in the characterization of outbreaks and the development of effective interventions to control cholera in Africa. The purpose of my research was to explore to what extent factors of social vulnerability and health systems capacity predicted the incidence of cholera in the 47 countries of the African region of the WHO.

Background of the Study

Cholera is an infectious disease caused by *Vibrio cholera* bacterium, usually transmitted through consumption or ingestion of contaminated drinking water or food (Phelps, Simonsen, & Jensen, 2019; Sun et al., 2017; Wolfe, Kaur, Yates, Woodin, & Lantagne, 2018). To control cholera outbreaks, WHO (2010a) recommends three interventions: (a) adequate treatment of cholera cases, (b) implementation of appropriate water and sanitation strategies, and (c) community engagement for positive behavioral change. The treatment cost of an individual cholera episode can vary from about $30 to more than $200, a significant financial burden for most African households that live on less than one dollar per person per day (Kirigia et al., 2009; Poulos et al., 2012). Thus, based on the number of cases officially reported by African countries, cholera represents up to $156 million of direct medical costs for the continent on a yearly basis (Kirigia et al., 2009). Cholera vaccine exists and has protective efficacy ranging from 60% to 80% for six months to two years (Bhattacharya et al., 2013; Khatib et al., 2012; Qadri et al., 2015). The vaccine is usually administered in mass vaccination campaigns to prevent cholera outbreaks, such as in humanitarian situation or hot spots of endemic areas. WHO (2018a) recommends a targeted use of cholera vaccine, prioritizing high-risk regions and population groups. One dose of cholera vaccine costs approximately $0.50 (Ali et al., 2013; Luquero et al., 2013; Qadri et al., 2016). In many African countries, even this cost represents a significant barrier to mass immunization campaigns. Thus, for many poor and developing African countries, the human, social, and economic cost of cholera is substantial.

Cholera is now endemic in Africa where at least 20 countries report outbreaks every year (WHO, 2016a). Studies have shown that the recurrence of cholera outbreaks was associated with various socioeconomic and environmental factors of vulnerability or health care capacity (Cerda & Lee, 2013; Filauri, 2010; Root et al., 2013; Talavera & Pérez, 2009). In an attempt to explain the mechanisms of social and environmental vulnerability to infectious diseases, Confalonieri, Wilson, and Najar (2006) proposed a framework for the emergence of infectious diseases that suggested that various factors play a role as the drivers of emergence, dissemination, or amplification of infectious diseases in a community in addition to the social capacity to respond to cholera outbreaks. Researchers applied a similar framework to study the recurrence and dissemination of cholera in the population (Filauri, 2010; Olago, Marshall, & Wandiga, 2007). However, in these later models, researchers did not include the capacity of the health system as one of the elements of the constructs. Nonetheless, health systems play an essential role in breaking the chain of transmission through adequate provision of health services, including prevention, detection, and response to outbreaks of cholera (Coltart & Atkins, 2017). A comprehensive characterization of factors associated with the incidence of cholera could inform the development of effective strategies for cholera control.

Problem Statement

Sub-Saharan Africa accounted for 72% of cholera deaths reported worldwide in 2015, with a recorded highest case fatality rate of 1.3% (WHO, 2016a). However, the real number of cholera cases per year was estimated to be much higher because of low reporting rates, poor disease surveillance systems, and limited laboratory capacity in most African countries. Some studies have suggested that only 5–10% of cholera cases were reported to WHO (Ali et al., 2012; Ali et al., 2015). Conservative estimations suggest that every year, 100,000 to 200,000 cases of cholera occur in Africa (Kirigia et al., 2009; Poulos et al., 2012). Cholera could be eliminated in Africa, provided that appropriate and evidence-based measures target the causes of the outbreaks. Some developing countries in Latin America and Asia have successfully eliminated cholera (Ali et al., 2015). WHO (2018b) attributes the failure to eliminate cholera in Africa to the weakness of health systems. WHO developed a global strategy for the elimination of predictable cholera epidemics by 2030, which targets specific groups and settings. According to WHO (2016b), shortcomings in human resources, health financing, and governance undermine the ability of African countries to prevent, detect, and respond to cholera. However, to what extent health system capacity is associated with the incidence of cholera has not been explored. Also, the effects of the interaction between health system capacity and socioeconomic indicators on the incidence of cholera need further exploration.

Purpose of the Study

The purpose of my research was to explore the relationship between population-level socioeconomic indicators and health system capacity as independent variables and incidence of cholera as dependent variable. My study setting was the 47 countries of the African region of WHO. Based on the ecosocial theory and Confalonieri et al.'s model of the emergence of infectious diseases, I conducted an ecological quantitative enquiry. I classified the independent variables into three groups: (a) drivers of emergence, (b) factors of dissemination or amplification, and (c) factors of health system capacity. The drivers of emergence include access to safe drinking water, open defecation, and improved sanitation. Elements of dissemination are poverty, income and gender inequalities, and adult literacy. The density of human resources for health and health financing composed factors of health system capacity.

Research Questions and Hypotheses

This study explored the following research questions and tested the related hypotheses:

RQ1: To what extent does social vulnerability (as measured by access to safe drinking water, access to improved sanitation, open defecation, poverty, income inequality, gender inequality, and adult literacy) determine the incidence of cholera in African countries?

$H_0$1: Population-level social vulnerability is not associated with incidence of cholera in African countries.

H_a1: Population-level social vulnerability is associated with incidence of cholera in African countries.

RQ2: To what extent health system capacity (as measured by health financing and density of human resources for health) is associated with incidence of cholera in African countries?

$H_0$2: Health system capacity is not associated with incidence of cholera in African countries.

H_a2: Health system capacity is associated with incidence of cholera in African countries.

RQ3: To what extent does the interaction of social vulnerability and health system capacity impact incidence of cholera in African countries?

$H_0$3: The interaction of social vulnerability and health system capacity does not impact incidence of cholera in African countries.

H_a3: The interaction of social vulnerability and health system capacity impacts the incidence of cholera in African countries.

Theoretical Foundation

The ecosocial theory of disease attempts to explain the occurrence and dissemination of diseases in a population. Its tenets suggest that emergence and distribution of diseases in a community are the combined effect of several factors and their interactions. Such population-level causes of diseases include social, economic, political, environmental, and biological parameters of a population (Krieger, 2011). Thus, ecosocial theory moves away from a purely biological explanation of the distribution of diseases in a community. It distinguishes between, on the one hand, the occurrence of a disease in individuals where the significant factors are biological, which is the focus of the biological model of the disease; and on the other hand, the incidence of the disease in the population where the main factors are a multilevel combination of social, economic, political as well as physiological parameters and their interaction (Krieger, 1994, 2004, 2011).

Formulated in 1994, the ecosocial theory of disease is the foundation of social determinants of health. Krieger (1994, 2001) described four constructs of ecosocial theory: (a) embodiment; (b) pathways of the embodiment; (c) cumulative interplay of exposure, susceptibility, and resistance across the life course; (d) and accountability and agency. Each of these four building blocks of ecosocial theory has a specific research perspective. Embodiment encompasses the exploration of the effect of social inequalities and poverty on population health outcomes. Pathways to embodiment focus on national arrangements as determinants of social development, which in turn, generate cumulative interplay. The latter focuses on conditions that are external to the population but determine the population's vulnerability to diseases. The last building block of the ecosocial theory, accountability and agency, tries to draw from embodiment to explain the incidence of diseases and population health and define prevention and control measures (Filauri, 2010).

In line with ecosocial theory, Confalonieri et al. (2006) developed a framework of social vulnerability applied to emerging infectious diseases. This model comprised three components: (a) drivers and mechanism of emergence, (b) factors of dissemination of infection or amplification, and (c) social response or the capacity to control infectious diseases. According to Confalonieri et al.'s model, the three groups of factors determine the dissemination of infectious diseases in a population. Drivers and mechanisms of emergence include characteristics of society and conditions that facilitate the exposure to infection. Dissemination and amplification groups comprise features that enable infection to occur and its amplification in the community. The capacity includes the means and resources the population can use to respond to the diseases.

My study was an adaptation of Confalonieri et al.'s model, which I applied to African countries' vulnerability to cholera outbreaks. Like Confalonieri et al.'s model, my adapted model also comprised three components: (a) drivers of emergence, such as lack of access to safe drinking water or improved sanitation and open defecation; (b) factors of amplification; and (c) health financing and human resources. Drivers of emergence increase risk of contamination. The factors of amplification include population poverty, inequality, illiteracy, and access to information. Health financing and human resources for health represent the parameters of a health system's capacity to control cholera in a country. In summary, ecosocial theory seeks to determine the drivers of occurrence, recurrence, and distribution of diseases in a population. Accordingly, my study encompassed aspects of social causes of diseases (embodiment), social arrangements conducive to the dissemination of disease (pathways to embodiment), and the environment that facilitates the exposure to cholera (cumulative interplay).

Nature of the Study

My research was an ecological, quantitative, multigroup, comparison study, exploring the association between countries' socioeconomic indicators and health system capacity as predictors, and on the incidence of cholera as the outcome. The independent variables included access to safe drinking water, open defecation, improved sanitation, poverty, income and gender inequalities, adult literacy, the density of human resources, and health financing. My independent variables were population-level indicators, continuous at interval and ratio level of measurements, such as proportion, rate, indices, or ratios. My dependent variable was categorical, dichotomous expressed as *high* and *low*; hence, the use of logistic regression tests to assess the association between the independent and dependent variables as well as between the interaction of the predictors and the outcome. The sources of my secondary data were databases and reports from United Nations agencies. The source of data for the independent variable was a study conducted by Ali et al. (2015) estimating the global incidence of cholera by country. For consistency in data, I collected all dependent variables for the year 2015.

Definitions

The definitions of the indicators reported here were drawn from the Indicator Compendium of the World Health Statistics (WHO, 2015).

Adult literacy rate: The proportion of population ages 15 years and over with the ability to read, write, and make simple arithmetic calculations in everyday life (WHO, 2015).

African region of WHO or AFRO region: One of the six regions of the WHO, composed of 47 countries, including all African countries in sub-Saharan Africa, the Indian Ocean, the Atlantic Ocean, and one country in the Arab region, Algeria (WHO, 1997).

Gender inequality: Legal, social, and cultural conditions that determine disadvantageous treatment of people on the sole basis of gender. Gender inequality index is the common indicator of gender inequality (UN Entity for Gender Equality and the Empowerment of Women, 2017)

Health financing: One of the functions of a health system aiming to make funding available and ensure that all individuals have access to effective public health and health care services (WHO, 2015)

Human resources for health: All people whose primary objective of their work is to enhance health. Human resources for health include clinicians who deliver health services or other officers who manage or support the delivery of such health services (WHO, 2015).

Improved drinking water: Drinking water collected from a source protected from outside fecal contamination. Such sources include a pipe into dwelling or tap, stand pipe, borehole, protected dug well, protected spring, and protected rainwater collection (WHO & United Nations Children's Fund [UNICEF], 2017).

Improved sanitation: Facility that hygienically separates human excreta from human contact. Facilities can include flush or pour-flush to piped sewer systems, septic tanks or pit latrines, ventilated improved pit latrines, or pit latrines with slab and composting toilet (WHO & UNICEF, 2017).

Incidence of cholera: The number of cases of cholera during a given period in a specified population (Porta, 2008)

Income inequality: Extent to which the national income is unevenly distributed in a country. Income inequality is commonly expressed in Gini index or Gini coefficient (World Bank, 2014, 2018).

Open defecation: Human feces disposed in an open field, in the bush, or in absence of a latrine (WHO & UNICEF, 2017).

Poverty: People living on or below the poverty line of $1.25 per day as measured at 2005 international prices (WHO, 2015).

Assumptions

I worked under the assumptions that the data obtained from the databases and reports from United Nations agencies were accurate, and the accuracy is consistent across countries. The second assumption was that Ali et al.'s (2015) model was fit enough to compute a realistic incidence of cholera in African countries. Ali et al.'s (2015) study was the most recent and comprehensive estimation of cholera incidence in African countries.

Scope and Delimitations

The geographic scope of my study was the 47 countries of the African region of the WHO. The AFRO region of WHO includes 40 sub-Sahara African countries, four

African islands in the Indian Ocean, two African islands in the Atlantic Ocean, and one country in the Maghreb region. Most AFRO countries have epidemiological and socioeconomic similarities. The other six African countries, mostly in the Maghreb and Horn of Africa, do not belong to the AFRO region of WHO and thus were excluded from the study. Also, non-AFRO countries did not share a similar epidemiologic and socioeconomic profile with AFRO countries. In addition, indigenous cases of cholera are rare in the six non-AFRO countries. While exploring the association between socioeconomic indicators and the incidence of cholera, my study did not include environmental factors, such as rainfall and temperature. At a population level, such factors mostly impact the incidence of cholera through socioeconomic factors, such as access to water, quality of water, or improved sanitation.

Limitations

The ecological design of my study involved some limitations, mainly related to the risk of ecological fallacy. According to Morgenstern (1982), one of the threats to ecological studies is ecological fallacy, which results from making inference on individuals based on population-level data. Thus, the results of my ecological study at the national level will not apply to lower geographic entities or individuals within the countries (Frankfort-Nachmias & Nachmias, 2008). Another common bias in ecological studies is the reversal of the effect of the disease or outcome on the exposure or predictor. However, because of the acute nature of cholera outbreaks, their reverse effect on socioeconomic indicators of a country could be deemed as not significant. However, their cumulative impact on socioeconomic indicators over a long period cannot be ruled out.

Nevertheless, because my study was cross-sectional, I could not assess such effect over a long period. Finally, I used a convenience sample by including in the study all the 47 countries of my sampling frame without applying any probabilistic or randomization method. Therefore, my sample was not necessarily representative of other developing countries in the world. The results cannot be generalized to other developing countries or even other African countries not included in the study.

Significance of the Study

Significance to Theory

Epidemiologists have extensively studied risk factors related to cholera infection in Africa. My study adds to the discussion on social epidemiology of cholera with the inclusion of the health system to the analysis. In the general debate on the social determinants of health the discussion on social causes of diseases predominantly focuses on non-communicable diseases. My study adds to the discussion on the concept of social determinants of infectious diseases—more specifically, social determinants of cholera.

Significance to Practice

Cholera affects millions of people each year, killing thousands, mostly in Africa. Currently, the design and implementation of interventions to control cholera are mainly guided by individual-level risk factors. However, interventions based on individual-level risk factors with no societal ground may be less effective for population-based health outcomes (Link & Phelan, 1995). Also, from the operational and programming point of view, it is much easier to translate population-level determinants of outbreaks into public health programs. Therefore, the characterization of population-level drivers of cholera emergence and dissemination can inform the development of policies and programs to control cholera countrywide. Analysis of the social determinants of cholera outbreaks can inform the characterization of the socioeconomic profile of cholera incidence for African countries. Such population-level risk profiling can then guide the development of cholera elimination programs in Africa and establish a concrete step in the operationalization of the concept of health in all policies (WHO, 1986).

Significance to Social Change

For the last 10 consecutive years, recurrent cholera outbreaks have occurred in at least 20 of 47 countries in the African region of the WHO every year (WHO, 2016a). Many of these impoverished countries divert a significant portion of their resources to respond to recurrent cholera outbreaks (Kirigia et al., 2009). The diversion of a substantial part of health budgets to managing outbreaks of cholera perpetuates social underdevelopment. Characterization of the population-level risk of cholera occurrence in African countries can assist decision-makers in addressing social vulnerability to cholera outbreaks. Control of cholera outbreaks can contribute to the alleviation of infectious diseases induced by underdevelopment and, thus, shape the road to attainment of sustainable development goals (WHO, 2017a). At household and individual levels, cholera episodes can plunge people into poverty, including through catastrophic health expenditures (Kirigia et al., 2009). Addressing the social determinants of cholera can also support and boost poverty alleviation programs in the community.

Summary and Transition

The magnitude and recurrence of cholera outbreaks in African countries call for the exploration of epidemic drivers there. Researchers have characterized individual-level risk factors, often focusing on proximal risk factors such as access to safe drinking water and improved sanitation. Although other previous studies attempted to explore the association between population-level risk factors and cholera incidence, they did not consider health system capacity. Moreover, observation of social indicators based on such risk factors have shown that African countries with similar socioeconomic profiles have significantly different records of cholera outbreaks, hence the need for an inquiry to explore the drivers of cholera outbreaks. I conducted a quantitative and ecological inquiry to investigate the relationship between population-level drivers of countries' vulnerability and capacity and the incidence of cholera in Africa. To that end, I examined to what extent factors of vulnerability and capacity were associated with the incidence of cholera in the 47 countries of the African region of the WHO. My research was grounded in the ecosocial theory of the dissemination of diseases and built around the framework of the emergence of cholera, which I adapted from Confalonieri et al.'s (2006) model of dissemination of infectious diseases.

The next chapters, literature review and research method, include an in-depth discussion about the application of the ecosocial theory and the framework of the emergence of infectious diseases, the study design and sample population, the independent and dependent variables, and statistical analyses used to answer the research questions.

Chapter 2: Literature Review

Cholera remains a prevalent health and social problem in many impoverished African countries, but it has been mostly eliminated in developed countries. In the last 10 years, about 20 African countries have reported numerous cases of cholera each year (WHO, 2016a). In 2015, sub-Saharan Africa accounted for 72% of cholera deaths reported worldwide, with the highest case fatality rate of 1.3% (WHO, 2016a). Recurrence of cholera outbreaks in Africa has been associated with socioeconomic indicators such as limited access to safe drinking water and poor sanitation (Cerda & Lee, 2013; Root et al., 2013). Treatment of each cholera episode can cost between $30 and $200, a significant financial burden for most households in Africa (Kirigia et al., 2009; Poulos et al., 2012).

Cholera is an infectious disease caused by *Vibrio cholera* bacterium and is usually transmitted through ingestion of contaminated drinking water or food (Bompangue et al., 2008; Jutla et al., 2013; Nkoko et al., 2011; Olago, Marshall, & Wandiga, 2007). Poor hygiene and sanitation or unsafe treatment or handling of food and water are primary drivers of cholera dissemination (Bwire et al., 2017; Mengel et al., 2014; Rebaudet et al., 2013). Also, factors such as the level of education and poverty, which often determine social status and access to health services, impact the occurrence of cholera outbreaks (Al-Arydah et al., 2013; Abdussalam, 2015).

Researchers have extensively studied factors that determine the risk of contracting cholera at the individual level (Nkoko et al., 2011; Lilje, et al., 2014; Lilje, et al., 2015; Mintz & Tauxe, 2013; Nguyen, et al., 2014; Nsagha, 2015; Schaetti et al., 2013).

However, little research has been done to explore population-level drivers of the incidence of cholera (Abdussalam, 2015; Filauri, 2010; Root et al., 2013). Moreover, the few studies that have analyzed population-level determinants of the incidence of cholera have not considered the effect of the health system, a critical aspect in the prevention and control of cholera outbreaks (Gostin & Friedman, 2015). The purpose of my research was to investigate the effect of population-level social vulnerability and the capacity of the health systems on the incidence of cholera in the 47 countries of the WHO’s AFRO region. This literature review is organized into six sections. The first section discusses the strategy to search and identify primary sources. The second and third sections respectively confer about the ecosocial theory and the theoretical framework for the emergence of infectious diseases. The fourth section introduces the various variables, or population-level determinants of the incidence of cholera, and the chapter closes with a summary.

Literature Search Strategy

The strategy to locate relevant primary sources consisted of two main strategies: keyword searches and mining bibliographies. In the first approach, I used a series of keywords to locate relevant sources through search engines or within the Walden University library. Thus, in the initial search I browsed through a broad range of studies and selected relevant publications. The second strategy consisted in locating other primary sources from the references or citations from the first search. I searched the following search engines and databases in the first approach: Google Scholar, ProQuest, Medline, CINAHL, Science Direct, PubMed, Cochrane, and Thoreau. To search the databases, I used the keyword *cholera* combined with other concepts related to my variables or their equivalent: *poverty, access to safe water, access to improved sanitation, open defecation, income inequality, gender inequality, health financing,* and *density of health workforce*.

The Google Scholar search engine and Walden library databases were set to retrieve studies published from 2013 onward. However, I also included in my literature review some seminal studies published before 2013 when they were the most relevant to my study topic or research approach. My literature review included as a priority, studies extracted from peer-reviewed journals, original articles, systematic reviews, dissertations, and books. Finally, I selected, as much as possible, studies that met at least one of the following criteria: relevance, theoretical and method similarity, and focus in Africa or other developing countries.

Theoretical Foundation

The premise about the role of the social and environmental context in the emergence and dissemination of diseases generated the ecological approach to epidemiological research of infectious diseases (Diez-Roux, 1998; Ackers, 1998). The ecological approach stems from the ecosocial theory of disease distribution or *web of causation of diseases* theory formulated in the 1960s (Berkman & Kawachi, 2014; Krieger, 1994, 2001). The ecosocial theory of disease explains the occurrence or dissemination of diseases as a result of the interaction between social, economic, political, environmental, and biological parameters of a population (Krieger, 2011). Thus, according to the theory, the distribution of diseases in a population is not driven by a mere sum of the individual risk factors (Krieger (2011). Instead, it originates from the interaction between several factors at the population level, including social, economic, political, environmental, and biological parameters (Frohlich, Corin, & Potvin,2001). Hence, Krieger (1994) distinguished the "causes of cases" from the "causes of incidence" (p. 892). The concept of causes of cases versus causes of incidence suggests that the features of individuals only determine the manifestation of a disease in individuals whereas the population-level factors drive the dissemination of the disease or its frequency in a community.

Consequently, researchers have analyzed the causality of the incidence or prevalence of diseases in populations as opposed to exploring etiological risk factors in individuals. The differentiation between the causes of cases and causes of incidence is also at the center of the controversy that still surrounds ecosocial theory. Other authors have likened ecosocial theory to a reductionist and shortsighted approach to epidemiology (De Camargo, Ortega, & Coeli, 2013; Poole & Rothman, 1998; Rothman, Adami & Trichopoulos, 1998; Vandenbroucke, 1998). Nevertheless, and despite the controversy, some studies have shown that population-level indicators are statistically associated with incidence of several infectious diseases (Ackers, 1998; Filauri, 2010; Pinzon-Rondon et al., 2014). Far from researchers' disputes, and from the operational point of view, public health measures aiming to control outbreaks are often population-oriented. Hence, the application of ecosocial theory to explore factors associated with the incidence of diseases in various settings, at national, sub-national, or even household levels has expanded in recent years (Jones, Betson, & Pfeiffer, 2017). Based on ecosocial theory, Confalonieri et al. (2006) developed a model of social and environmental vulnerability to emerging infectious diseases, including cholera. They posited that the emergence of infectious diseases results from the breach of the equilibrium between social, environmental, and microbiological factors.

Conversely, Filauri (2010) used the political-ecology framework derived from ecosocial theory to investigate the effect of state capacity on the incidence of neglected tropical diseases, including cholera, in 33 African countries. Filauri's (2010) results indicated a statistically significant correlation between state capacity as measured by *control of corruption, external debt stocks,* gross domestic product (GDP), government effectiveness, foreign direct investment, political stability, regulatory quality, and secondary-school enrollment and incidence of cholera. Leckebusch and Abdussalam (2015) also applied the socioecological approach to explore meteorological and socioeconomic factors as determinants of the spatiotemporal variability of the burden of cholera in 36 states of Nigeria. Their results showed a positive association between the combined effects of rainfall, seasonal temperature, poverty, and population density as predictors and incidence of cholera as the outcome. Leckebusch and Abdussalam's (2015) results also showed a negative association between the combined effect of access to piped water and adult literacy and the incidence of cholera. Further, Root et al. (2013) used a similar ecosocial approach to investigate the impact of socioeconomic and demographic factors on the severity of a cholera outbreak at the household level in Bangladesh. They found that factors of socioeconomic status, such as the household assets, years of education of the head of the household, access to latrines, and ownership of agriculture land were associated with the burden of cholera at the household level.

Based on ecosocial theory, other researchers have formulated tentative explanations of the emergence and dissemination of diseases in populations (Birkmann, 2006). For example, Confalonieri et al. (2006) developed a model of the emergence of infectious diseases in developing countries. Thus, in line with Confalonieri et al.'s model, ecosocial theory would explain the occurrence, dissemination, and persistence of cholera in African countries as resulting from the effects of population-level social and environmental determinants and the capacity of the health system. The level of social and environmental features in a country would, thus, determine to what extent such factors impact incidence of cholera in Africa. However, the incidence and prevalence of cholera in the population could also depend on the capacity of the health system to swiftly break the chain of contamination through adequate disease control programs.

In summary, the level of the incidence of cholera in a country could depend on the socially vulnerability of its population and the capacity of its health system to control the outbreak. Social vulnerability could increase the exposure to the germs and the susceptibility of a community to cholera outbreaks (Sugimoto et al., 2014). The country's health system, on the other hand, could control the magnitude of outbreaks through equitable access to health care, early detection of an outbreak, and timely action for response. Thus, health system capacity could determine the length and magnitude of the outbreak. The Confalonieri et al. model also suggested that the impact of social vulnerability and the health system results from the effect of each parameter as well as the effect of their interaction.

The social factors included elements such as poverty, economic and gender inequalities, literacy rate, access to safe water, improved sanitation, and open defecation. The health financing and human resources for health constituted the main features of the capacity of the health system. Also, in addition to the individual effects of the drivers of the emergence and occurrence of the disease in a community, the interaction between the variables needs to be considered. Indeed, while factors of social vulnerability, such as poverty and poor access to safe drinking, could increase the exposure to cholera, the effect of other factors, such as low rates of literacy and access to health services, could further compound the vulnerability of the population to cholera (Sugimoto et al., 2014). For instance, poverty could further limit access to health services, thus likely increasing the exposure of the community due to delayed or ineffective treatment and isolation of cholera patients. The analysis of the impact of social, environmental, and health system characteristics on the incidence of cholera was the focus of my inquiry. In the next section, I discuss the parameters of social vulnerability and health system that impact the dissemination of cholera in Africa.

Literature Review

Factors Impacting Cholera Outbreaks

Two groups of population-level predictors were thought to determine cholera transmission and dissemination in the community and, consequently, the level of the incidence of cholera in a country. On the one hand, the social vulnerability group of predictors included the drivers of emergence, the factors of amplification, and on the other side, the capacity of the health system to control cholera outbreaks. The drivers of emergence comprised the rates of access to safe drinking water, access to improved sanitation, and open defecation. On the other hand, the rate of poverty, income inequality, gender inequality, and adult literacy constituted the factors of amplification. Extended dissemination of cholera could also depend on the effect of factors that amplify the dissemination of the disease in a community (Anbarci, Escaleras, & Register, 2012). For instance, inequality and illiteracy could negatively impact access to health care services and health information, which could thwart timely control of the outbreak. Also, a weak health system would fail to rapidly detect and respond to an outbreak, which can jeopardize prevention and control measures, leading to prolonged dissemination of the disease (Mengel et al., 2014). Moreover, poor management of biological waste in health facility settings would further amplify the outbreak through nosocomial transmission of the disease (WHO, 2018b). In the next section, I discuss the variables that determine the occurrence and recurrence of cholera outbreaks in Africa, drivers of emergence, factors of dissemination, and the capacity to control.

Drivers of Emergence

Safe drinking water. Since the work of John Snow in the 19th century, access to safe drinking water remains one of the most effective ways to prevent cholera infection and dissemination. Indeed, John Snow’s study in London in 1856 established the association between the use of unsafe and contaminated water with the incidence of cholera in the community (Snow, 1856). John Snow and other subsequent studies revealed the primordial role of fecal contamination of water sources in the occurrence of cholera (Bain et al., 2014; Kwesiga et al., 2018). As a result, epidemiologists and public health practitioners view consumption of contaminated water as the primary driver of the emergence or occurrence of cases of cholera in a community (Reidl & Klose, 2002; Taylor, 2015), while the practice of open defecation or an inadequate disposal of human feces represent the primary cause of contamination of water sources. Researchers even use *access to safe water* as a proxy-indicator of at-risk population (Ali et al., 2015). For that purpose, the WHO and UNICEF (2012) defined a safe water source as one that is adequately protected from fecal contamination. WHO and UNICEF definition also classifies water sources in *improved* and *unimproved*. Improved sources of water refer to sources protected from outside fecal contamination such as pipe, borehole, and protected spring while unimproved sources include all unprotected well and surface water.

In line with the WHO classification of water sources, a systematic review of 319 studies published between 1990 and August 2013 assessed the fecal contamination of different types of sources of drinking water in middle and low-income countries. The study found that unimproved sources of drinking water had a higher odds ratio of fecal contamination than improved sources (Bain et al., 2014). On the other hand, WHO and UNICEF report indicated that African countries have the highest rate of use of unimproved sources of water in the world (WHO & UNICEF, 2017). Therefore, the African population had a higher risk of cholera contamination from water sources and, African countries have a higher likelihood of occurrence and persistence of cholera outbreaks. Further, the results of the systematic review also showed that notwithstanding the type of water sources, improved or unimproved, the water sources were more likely to be contaminated in low-income countries than in middle and high-income countries; in rural areas than in urban areas. Bain et al., (2014) conducted a systematic review of 319 studies conducted in middle and low-income countries, between 1990 and August 2013. Their analysis included 96,737 of water samples tested for fecal contamination. The meta-analysis results indicated that unimproved sources of drinking water had a higher odds ratio of fecal contamination than improved sources. Thus, in the African settings, especially in rural areas, even improved sources of drinking water were likely fecal-contaminated. Corroborating the review results, Kwesiga et al. (2018) in Uganda found that water collected from the public pipes during a cholera outbreak in a rural district in 2015 had high fecal contamination (Kwesiga et al., 2018). The above results suggested that in Africa, contaminated water sources were likely the main driver of cholera contamination. They also suggested that African countries or communities that ranked poorly in the management of human waste had likely higher incidence of cholera regardless of the proportion of the population that had access to improved water sources.

Open defecation. Contaminated water supply is an indication that human feces has been in contact with the water source. Thus, contamination of water sources and its ultimate consequence, the cholera infection, is only a result of poor management of human waste or human feces. For example, Cowman et al. (2017) found that the incidence of cholera at the district level in Kenya was associated with the rate of open defecation. They conducted univariate and multivariate regression analyses, using data from the disease surveillance system and the environmental census. The results showed a positive association between the percentage of households practicing open defecation in the district and the incidence of cholera (Cowman et al., 2017). In contrast, Kenyan counties that eliminated open defecation were found to have a lower prevalence of diarrhea diseases (Njuguna, 2016). Statistics produced by the WHO and UNICEF (2017) indicate that open defecation significantly declined worldwide between 2000 and 2015, except in sub-Sahara Africa where it increased by about 7% (WHO & UNICEF, 2017). Therefore, the risk of cholera dissemination related to open defecation did not only remain significant for the last 15 years in Africa, but it likely increased. In such a context, improved sanitation, which is discussed in the next session, could represent a critical protection and prevention factor against the spread of cholera in the community.

Improved sanitation. Failure to adequately protect water sources or treat drinking water is the major risk of fecal contamination of drinking water. Further, treated water or food only remains safe as long as it is preserved from recontamination during transport, storage, or consumption. Therefore, improved sanitation represents one of the critical services for the prevention of cholera. A randomized controlled trial conducted in Bangladesh from 2013 to 2014 assessed the effect of hands washing on the risk of cholera. The researchers randomly assigned households of pregnant women from several villages in rural areas of Bangladesh to either intervention or control group. The interventions consisted of water treatment at the point of use, access to improved latrines, safe disposal of feces, and handwashing with water and soap. Also, the intervention group received weekly visits from health promoters for 6 months. The results showed that the prevalence of diarrhea in the past 7 days was lower in the intervention group compared to the control group (Luby et al., 2018). The results also showed that hand washing with soap reduced the risk of cholera by half even during a cholera outbreak (George et al., 2016). Several other studies have shown an association between limited access to safe drinking and poor sanitation with the occurrence of cholera in a country (Bompangue et al., 2008; Cerda, 2013; Rebaudet et al., 2013). However, similar randomized controlled trials in Kenya did not find any effect of sanitation and handwashing interventions on diarrhea prevalence (Null et al., (2018). But, authors of the Kenyan study cautioned against a flat interpretation of their results because of very low adherence to the intervention in Kenya. They reported that adherence to sanitation and handwashing intervention was only 19% in Year 2 in Kenya compared to 90% in Bangladesh.

Factors of Dissemination or Amplification

Several African countries are endemic to cholera despite having different profiles in access to safe drinking water and improved sanitation. Phelps et al. (2016) observed the same paradox between Denmark in the 19th century and Haiti in 2010-11. Phelps and coauthors noted that during the cholera outbreaks in Denmark in 1853 -1857 and in Haiti in 2010-11, the two countries had different profiles on access to safe water and improved sanitation. However, the two outbreaks had similar reported attack rates and reproductive number. They concluded that factors other than access to safe water and improved sanitation played a role in the dissemination of the disease during the two outbreaks.

Poverty rate. One of the most common characteristics of developing countries, particularly in Africa, is the dyad of a high prevalence of communicable diseases and rampant poverty of the population. Poverty and infectious diseases reinforce each other. The former constitutes a factor of vulnerability for the community while the later can trigger loss of income which further deepens poverty. Poverty, in turn exacerbate the social vulnerability of the population to infectious diseases such as cholera (Anbarci et al., 2012; Confalonieri et al., 2006; Cohen et al., 2003). Further, high rate of poverty in a country can represent a barrier to the access to health services for a significant proportion of the population, which can hinder the implementation of and adherence to interventions for cholera prevention and control (Asiedu et al., 2015; Cohen et al., 2003). Consequently, countries with a higher percentage of people living in poverty would also likely have a higher burden of infectious diseases (Eisenstein, 2016). The evidence also showed that cholera was associated with attributes of poverty such as lower income at national or household level. For instance, Bwire et al. (2017) conducted a prospective study to characterize the epidemiologic and socioeconomic features associated with the cholera outbreak in fishing communities of Hoina district in Uganda between 2011 and 2015. The results showed that households that registered cases of cholera had an income three times lower than the average. Other studies also confirmed a positive correlation between the incidence of cholera and the absolute poverty rate (Leckebusch & Abdussalam, 2015; Matsuda et al., 2008; Snowden, 2008). In conclusion, poverty could amplify the dissemination of cholera in a community through limited access to health care, health information or inadequate adherence to public health interventions.

National income. The implementation of public health programs including services for health promotion and prevention and control of outbreaks is resourcedependent. Consequently, a country's population health outcomes often reflect its national wealth and vice versa (Pop, Van Ingen, & Van Oorschot, 2013). As a corollary, the level of expenditure on health could be a proxy measure of the capacity of a country's health system to deliver services, promote health, and prevent diseases (Kim & Lane, 2013). Thus, limited capacity due to lower expenditure on health could result in high incidence and prevalence of infectious diseases. Studies have shown an association between the level of national income and the incidence of infectious diseases, including cholera. For instance, a cross-sectional survey in the Caribbean, North and Latin America indicated a strong and positive association between the Gross national product per capita and the incidence of cholera (Ackers, 1998). Another cross-sectional survey assessing the relationship between the Gross national income and the incidence of cholera worldwide showed that the percentage of countries with higher incidence of cholera was three times higher in the group of low-income countries compared to the middle and high-income countries (Perez, 2009). Thus, it appears that, at population-level, national income is another non-water factor that impacts the dissemination of infectious diseases such as cholera in the community.

Income inequality. Although the level of income of a country indicates the ability for a country to provide health, the distribution of wealth is often unequally distributed across different population groups within a country (World Bank, 2012). Income inequality exposes the most disadvantaged groups to the impact of poverty on health outcomes including the incidence of infectious diseases. In fact, Pickett and Wilkinson (2015) established an epidemiological causal effect of income inequality on poor population health outcomes. They applied the causality criteria of temporality, biological plausibility, consistency, and lack of alternative explanations to the causal effect of income inequality. The authors reviewed nearly 300 peer-viewed studies that also included ecological inquiries applying various research approaches, from cross-sectional to cohort and time-series and, in different geographical and temporal settings. The results indicated that the association between income inequality and health outcomes was statistically significant regardless of the methodological approach or the geographic settings. Their results were consistent with findings from Murray and Chen (1993) study which also indicated that the distribution of national income, together with efficiency and effectiveness of health expenditures, modulated the relationship between country's income and population health outcomes. Besides, Pop et al. (2013) assessing whether the reduction of inequality, particularly in developed countries led to improved population health, found that high income inequality was associated with lower life expectancy. Their data set covered a 10-year period and 140 countries. Gross domestic product (GDP) was the predictor and life expectancy, the study outcome. The cross-sectional and the longitudinal analyses yielded a statistically significant of both static and dynamic partial correlation, particularly in the group of low- and middle-developed countries. In the group of high-developed countries, the relationship between income inequality and life expectancy was non-significant.

However, other studies found no relationship between inequality and health outcomes. (Beckfield, 2004; Babones, 2008; King et al., 2010; Pop, van Ingen et al., 2013). For instance, Rajan, Kennedy, and King (2013) cross-sectional study assessing the effect of income inequality on under-five mortality at the state level in India, found no association between the Gini index and under-five mortality rate after controlling for literacy. They ran a linear regression to test the association between the Gini index and under-five mortality rate across the 35 states of India. Nevertheless, Pickett and Wilkinson (2015) contended the results of studies that showed a negative or inexistent association between income inequality and health outcomes. They attributed such findings to the use of inappropriate scales of measurement of inequality or health variables, as well as to insufficient follow-up period.

Researchers have suggested several hypotheses to explain the effect of income inequality on population health outcomes. Some authors suggested that high income inequality takes away the availability of resources from the majority, leaving a critical resources gaps for the prevention of diseases or access to health care services (Van Deurzen, Van Oorschot, & Van Ingen, 2014). Others suggested that income inequality has a direct effect on well-being and the incidence of diseases (Pullan, Freeman, Gething, & Brooker, 2014). Pickett and Wilkinson (2015) suggested that income inequality reinforce the effects of other known and unknown determinants, including psychosocial deprivation and social distance or even other disparities. Concerning the incidence of cholera in the African context, most probably both mechanisms weigh in the dissemination of the disease in the community. Deprivation of resources, as well as the effect of inequality on the well-being, could reinforce social vulnerability of communities.

Gender inequality. There was no indication that biological features and gender were a factor of vulnerability to infectious diseases such as cholera. Also, several studies during various cholera outbreaks in Africa did not find any gender biased incidence (McCrickard et al., 2017; Saha et al., 2017; Sauvageot et al., 2016). Rancourt (2013) assessed the gender difference in the burden of cholera in Sierra Leone during a large-scale cholera outbreak in 2012. The results indicated that there was no significant difference in the incidence of cholera by sex when adjusted for the size of each sex in the population.

But, gender inequality or the social prejudice that women and girls face in a society can be detrimental to the health outcomes of this section of the population. Branisa (2013) suggested that societies that deprive women and girls of their autonomy, bargaining power, and the freedom to participate in social life fully, were more likely to have higher mortality rates. Further, Guerra-Silveira and Abad-Franch (2013) conducted a meta-analysis study to test the physiological and behavioral prediction models of the exposure to and incidence of ten infectious diseases: American leishmaniosis, schistosomiasis, pulmonary tuberculosis, leprosy, typhoid fever, leptospirosis, meningococcal meningitis, hepatitis A, and severe dengue fever. The results of the estimate of the male/female incidence rate ratios indicated that gender-based characteristics were determinants of the male/female incidence ratio for those infectious diseases. Sen and Östlin (2008) even argued that gender inequality was one of the most “influential” of the social determinants of health (p. xii). Moreover, a study showed a positive association between gender inequality and another health outcome, child mortality rates (Brinda, Rajkumar, & Enemark, 2015).

In African culture, men and women have specific social roles assigned to them, which often create social prejudice for women to gain access to social services such as health care. Also, the duties assigned to women such as water and food fetching, often expose them to parasites and other vectors of diseases. The social prejudice of gender inequality was also found to be strongly associated with lower female education attainment and female literacy rate (Branisa, 2013), both are crucial to access to health information and health services. Thus, gender inequality can also contribute to the dissemination of cholera in the community through deprivation of access to health care services as well as the increased social vulnerability of a significant proportion of the population.

Population literacy. In public health, successful implementation of risk-reducing interventions always requires an adequate level of community participation or collaboration. To that end, public awareness or access to information is critical to cholera prevention and control (Ramesh, Blanchet, Ensink, & Roberts, 2015; Taylor, Kahawita, Cairncross, & Ensink, 2015). Failure to comprehend health information can result in poor adherence of the community to the measures for the prevention and control of diseases, which contribute to the dissemination of the infectious diseases in the community. Therefore, the rate of literacy in a country, which approximates the proportion of the population that can gain access to health information, could impact the dissemination of diseases such as cholera in a country.

In fact, studies have found a relationship between countries' literacy rates and the incidence of cholera. For instance, Ackers (1998) collected surveillance data from North and South America countries to explore a correlation between population-level demographic and socioeconomic indices and the incidence of cholera. The results of the cross-sectional study found among others that countries that had a female literacy rate above 90% also had lower cholera cumulative incidences. Further, Leckebusch and Abdussalam (2015) found an association between adult literacy rate and the incidence of cholera at the state level in Nigeria. They also conducted a cross-sectional study to assess a predictive model of meteorological and socioeconomic factors on the temporal and spatial variation of the incidence of cholera at the state level. Their data set covered 12 years, from 2000 to 2011, on cholera cases and deaths collected at the state level in all 36 states of Nigeria. The results of the stepwise multivariate logistic regression showed a statistically significant, but negative relationship between the rate of adult literacy and the incidence of cholera at the state level. Furthermore, Cowman (2015) in Kenya and Ali et al. (2017) in India found similar results showing a statistically significant association between adult literacy and the incidence of cholera. In Kenya, counties with higher female literacy rates were found to have lower incidence of cholera; while in India, districts with higher literacy rates had three times lower incidence of cholera. Thus, it appears that, regardless of the geographic settings, the adult literacy rate is a strong predictor of the incidence of cholera.

Health System Capacity

The health system is the cornerstone of a country's ability to provide health care services, prevent diseases and promote health. The WHO (2000, 2010b) defines the health system as "all the organizations, institutions, resources and people whose primary purpose is to promote, restore, and improve population health" (p. vi). The health system model of the WHO is composed of six building blocks that determine its performance: health leadership and governance, health financing, human workforce, health information, access to essential medicines and vaccines, and health service delivery. Among the six building blocks, the WHO's panel for the reform agenda for global health security which reviewed WHO management of outbreaks, recommended explicitly among others that for effective control of outbreaks, each country should strengthen in priority its human health workforce and health financing (Gostin & Friedman, 2015; Elston et al., 2016; Piot, Coltart & Atkins, 2017).

Though the relationship between the health system capacity and the incidence of cholera is yet to be explored, Filauri' s study provided an analytical basis for inference reasoning. For her dissertation, Filauri studied the relationship between state capacity and the incidence of neglected tropical diseases including cholera in 33 African countries from East and West Africa. The framework included state capacity as the independent variable and the incidence of neglected tropical diseases as the dependent variable. State capacity was defined as "the basic services provided by the state to its people, characterized by eight attributes: (a) human capital, (b) instrumental rationality, (c) coherence, (d) resilience, (e) autonomy, (f) fiscal resources, (g) research and responsiveness, and (h) legitimacy" (p. 11). The results of the multivariable logistic regression showed a negative relationship between control of corruption, government effectiveness, political stability and school enrolment and the incidence of cholera in 33 African countries, meaning the incidence of cholera decreases as country's state capacity increases.

According to Murray and Evans (2003), health system financing is one of the functions that determine the capacity of a health system to reach its goal. According to Kruk and Freedman (2008), the level of health financing is one of the proxy indicators of health system effectiveness, health governance and performance; and the density of human resources for health is a determinant of services availability and delivery. In the next section, I discuss the two critical factors of the health system capacity: health financing and human resources density.

Density of human resources for health. Most cholera cases and large outbreaks often occur in disadvantaged communities such as slums and rural areas. In Africa, the disadvantaged geographic areas such as slums and remote rural areas are hardly covered by skilled health personnel. The presence of adequate human resources in rural and other underprivileged communities is crucial for timely detection and control of outbreaks. The information about the relationship between the density of the health workforce and the incidence of cholera in Africa is scanty. However, studies have found a significant association between the health workforce density and other health outcomes such as the rate of the immunization coverage, and child and maternal mortality rates (Anand & Bärnighausen, 2004; Chen et al., 2004; Mitchell et al., 2008; Nguyen, Mirzoev & Le, 2016). Anand and Bärnighausen (2004) conducted a multiple regression which showed that, controlled for national income and adult literacy, the density of workforce for health was associated with maternal and infant mortality, the higher the density, the lower the infant and maternal mortality rates. Mitchell et al. (2008) also found a significant association between the density of human workforce and immunization rate at the provincial level in Turkey, independently of female illiteracy and the GDP per capita. But, other studies found no association between the density of human resources for health and some health outcomes. For instance, though Kruk et al. (2009) found an association between the density of health workers and coverage of measles immunization and the use of skilled birth attendants, they did not find any association between the aggregated health workforce (nurses and doctors) and antenatal care and cesarean section. Also, Castillo-Laborde (2011) did not find any association between density of health workforce and the disability-adjusted life years.

Regardless of the research findings, African countries in 2015 only had on average as low as 2.7 physicians and 12.4 nursing and midwifery personnel per 10,000 population (WHO, 2015). The data suggest that African countries can hardly meet population needs for the prevention and control of outbreaks.

Health financing. In developing countries of the AFRO region of the WHO, another critical constraint to the prevention and control of cholera is the chronic lack of funding for the health system. According to the WHO’s 2015 world health statistics, African countries have in average, total expenditure on health per capita of only $105, of which, only half comes from the government (WHO, 2015). Further, more than 60% of the private expenditure on health comes directly from the individuals seeking health care services in the form of direct out-of-pocket spending (WHO, 2015). Meanwhile, the WHO estimates that a minimum of $80 as government health expenditure of per capita is required for an effective and efficient health system (Evans, Tandon, Murray & Lauer, 2001). For instance, a study of the 17 countries members of the Organization for Economic Co-operation and Development has found that health expenditure was strongly correlated with health outcomes such as infant mortality and life expectancy (Kim & Lane, 2013). The study showed that infant mortality decreased as the government expenditure on health increased. In parallel, life expectancy at birth was lower in countries with lower government' health expenditure.

Although the association between health system financing and the incidence of cholera in Africa is yet to be explored, cholera prevention and control requires an adequate level of funding. Indeed, the treatment of a cholera episode varies from $30 to $200 (Kirigia et al., 2009; Poulos et al., 2012). That cost represents a high risk of catastrophic health expenditure for many households in Africa where many live with less than two dollars per day. Catastrophic expenditure on health further exacerbates access to health care services for the most underprivileged and hinders adherence to public health measures meant to prevent and control cholera. Thus, the level of out-of-pocket payment directly affects the attitude of the community towards public health services, as shown by research evidence. Xu et al. (2003) explored the causes of catastrophic health care payments in 59 countries worldwide. The results revealed a positive association between the levels of out-of-pocket payments and the proportion of households facing catastrophic health expenditure. The same study also showed that households facing catastrophic health expenditure avoid seeking health care services altogether. The same results also indicated that the proportion of households facing catastrophic health expenditure was associated with the percentage of people living below the poverty line and with the portion of the country's GDP allocated to health expenditure.

In summary, the prevention and control of cholera are dependent on the availability of funding. Insufficient allocation of resources is detrimental to cholera prevention and control. In Africa, heavy reliance on out-of-pocket expenditure and the limited government expenditure on health likely contribute to the poor management of cholera outbreaks and the recurrence of large-scale outbreaks.

Conclusion

The review of the literature showed that, although several types of research explored environmental and socioeconomic as predictors of the incidence of cholera, the health system capacity was yet to be included in the analysis. The review also suggested that the emergence and persistence of cholera outbreaks in African countries could result from the interaction of multiple factors, within and outside the health sector at the population level. Several determinants could play various roles in the dynamic of the occurrence, dissemination, or control of outbreaks of cholera. The determinants included environmental factors such as access to safe drinking water and improved sanitation, and the rate of open defecation which were grouped as the drivers of occurrence of cases of cholera in the community. Besides, other socioeconomic indices such as poverty, adult literacy rates, income, and gender inequalities were likely to amplify the dissemination of cholera in the community. But, the impact of the capacity of the health systems to control cholera outbreak on its incidence in the community was yet to be determined. And yet, to what extent the capacity of the health system to prevent, detect and control cholera outbreaks, affects the occurrence and recurrence of cholera outbreaks was one of the critical questions for operational purposes. My research explored the extent to which socioeconomic drivers of cholera, combined with the capacity of the health system could predict the incidence of cholera in Africa.

Chapter 3: Research Method

The purpose of my study was to explore the population-level determinants of dissemination and amplification of cholera in the African region of the WHO. The results of this study will contribute to a better understanding of the dynamic of cholera outbreaks in African countries for tailored interventions to eliminate cholera in Africa. This chapter includes four major sections discussing the research design, study methodology, data analysis plan, and threats to validity. The first section highlights the key features and rationale of my study design, including the presentation of the variables and their connection to the study questions. The second section presents the study methodology: (a) the study population, (b) the sampling procedures and sample size, as well as (c) the effect size and the study power. The data analysis plan describes and provides a rationale for the use of specific statistical tests. The section on threats to validity will discuss internal and external validity as well as the ethical aspects related to the process and procedures of my research. Finally, I conclude this chapter with a summary of the concepts discussed in this chapter as well as an introduction of the next.

Research Design and Rationale

My study explored to what extent population-level socioeconomic indicators were associated with the incidence of cholera in Africa. My units of analysis consisted of geographic entities instead of individuals, and my sampling frame was the 47 member states of the African region of the WHO. Therefore, I conducted an ecological study with countries' socioeconomic indicators as independent variables and the incidence of cholera as the dependent variable. An ecological inquiry is often used to explore the association between population-level characteristics of geographic entities and population health outcomes, usually to compare prevalence or incidence of diseases between geographic areas (Levin, 2006).

One of the advantages of ecological studies is that they can generate hypotheses for practical application at the population level, even though they are subject to potential ecological fallacy (Frankfort-Nachmias & Nachmias, 2008). Also, for feasibility and ethical reasons, I could not conduct an experimental study or assign and subject countries to various levels of socioeconomic performance to assess the impact on the incidence of cholera. Therefore, an observational study remained the best option to assess the association between socioeconomic indicators and incidence of cholera in African countries. Also, the ecological approach fits with the use of population-based data, comparing countries rather than individuals.

In my conceptual model, the independent variables were classified into two main groups: social vulnerability and health system capacity. Social vulnerability was further divided into drivers of emergence and factors of dissemination or amplification. All independent variables were expressed in terms of rate, ratio, or index. Thus, all independent variables were continuous. The dependent variable, the incidence of cholera, was also collected as a continuous variable. However, because of little variation in the incidence of cholera among the African countries, the dependent variable was coded as categorical at two levels: high and low. Hence, logistic regression was the appropriate statistical test for my research questions and hypotheses. The data analysis plan is discussed in more depth in the next section.

Methodology

The study population was the 47 member countries of the African region of the WHO. For administrative purposes, the WHO groups its member states into six regions: (a) Africa, (b) the Americas, (c) Southeast Asia, (d) Europe, (e) Eastern Mediterranean, and (f) Western Pacific. The African region of WHO has 47 countries that, with the exception of Algeria, are all sub-Saharan countries. In line with the World Bank income classification, WHO also groups countries into four categories of income level: (a) low, (b) lower-middle, (c) upper-middle, and (d) high. Of the 47 countries in the African region of the WHO, only one is classified as a high-income country (Seychelles) and seven are upper-middle-income countries (World Bank, 2017). The rest of the countries are classified as either lower-middle-income (13 out of 47) or low-income countries (26 out of 47). Among the 47 AFRO countries of the WHO, at least 23 countries (48.9%) are endemic to cholera or have reported a cholera outbreak for at least three consecutive years in the 5 years preceding 2015. Among the 23 endemic countries, 15 countries (65%) are in the group of low-income countries, and eight (35%) are classified lower-middle-income countries. No country in the upper-middle or high income categories is endemic to cholera in the Africa region.

Sampling and Sampling Procedures

The sampling frame, the African region of the WHO, comprises 47 countries which are my units of analysis. Thus, all 47 countries belonging to the African region of WHO are eligible for the study. I decided to include all 47 countries in the study because of a relatively limited sampling frame. I used convenience sampling, a nonprobability sample design. To assess the study power, I first determined the achieved power given the sample size of 47 countries, using the G*power application. The results showed that for a large size effect (0.3), the study had a power of 99%, and 96% and 57% of power, respectively, for medium (.15) and small effect size of .02. I computed power under the assumption that the probability of *high incidence* of cholera in the absence of independent variables was negligible, as low as 0.1. The proportion of variance between independent variables (R^2) was estimated at 0.01 (Filauri, 2010). Second, I also conducted the sensitivity analysis to compute the required effect size, given α, power, and the sample size. With α set at 0.05 and a total sample size of 47, the results showed that a large sample size of 0.2 was required at 95% of the study power. Further, because of the relatively small and fixed size of my sample, I also conducted compromise power analyses to determine the probability of Type I and Type II errors by computing the critical value, and the values of α and β, given my fixed sample size of 47 units and the odds ratio of 0.15. The analysis yielded a critical z of –1.84, the α value of 0.03, which means that with power of 97%, I have 3% chance of Type I error of rejecting the null hypothesis while it is true. The computed α was relatively lower than the usually accepted value of α set at 0.05.

Instrumentation and Operationalization of Constructs

The conceptual model of my study had social vulnerability and health system capacity as the two groups of predictors of the incidence of cholera. Social vulnerability included two subcomponents: drivers of emergence and factors of dissemination. The drivers of emergence included the following parameters: (a) access to safe drinking water, (b) access to improved sanitation, and (c) rate of open defecation. Factors of dissemination or amplification included (a) poverty, (b) income inequality, (c) gender inequality, and (d) adult literacy. Health system financing and density of human resources for health constituted the parameters of health system capacity. At least one specific indicator measured each of the parameters, either parameters of social vulnerability or parameters of the health system capacity. Table 1 presents the indicators for each parameter.

Table 1

Indicators of the Independent Variables

Component	Subcomponent	Variable	Indicator
Social vulnerability	Drivers of emergence	Access to safe drinking water	% of population with access to safe drinking water
		Access to improved sanitation	% of population with access to improved sanitation facilities
		Open defecation	Proportion of the population practicing open defecation (open defecation rate)
	Factors of dissemination/ Amplification	Poverty	Proportion of people living below the poverty line ($1.90 per day)
		Income inequality	Income Gini index
		Gender inequality	Gender inequality index
		Literacy	Literacy rate among adults ages ≥ 15 years
Health system capacity		Health financing	Total expenditure on health as a percentage of GDP
			Proportion of general government expenditure on health as a percentage of total expenditure on health
			Proportion of out-of-pocket payment as percentage of total expenditure on health
		Density of human resources for health	Density of physicians per 10,000 population
			Density of nursing and midwifery personnel per 10,000 population

Sources of Data

Data were collected data from various databases and reports from international organizations, including WHO, UNICEF, and the United Nations Development Program (UNDP). The incidence of cholera in African countries was extracted from an original study. Ali et al. (2015) computed the global burden of cholera. All data were collected for 2015, the same period as the computed cholera incidence. A data entry form was also developed in Microsoft Excel as a workbook in table style to combine the data sets from different databases, then exported to an SPSS data entry table for analysis.

I opted to use data on incidence of cholera from an original study because of the limitation of the surveillance system in most African countries. Indeed, the WHO cholera database only records cholera cases as reported by countries to WHO on a voluntary basis, and from their respective surveillance systems. But, because of structural and operational weaknesses of the disease surveillance systems in Africa, the number of cases of cholera as reported by countries is reported to be significantly underestimated (Ali et al., 2015). It is estimated that the surveillance systems in Africa only captures ten to fifteen percent of cases of cholera (Ali et al., 2012). Also, even in relatively advanced countries, the surveillance systems still depend on several other factors such as the utilization rate of health facilities and the laboratory capacity to confirm cholera (Ali et al., 2012; Bompangue et al., 2011; Sauvageot et al., 2016). Ali et al. (2015) estimated the global burden of cholera in endemic and non-endemic countries. They used the data from the WHO's Annual Cholera Global Surveillance Summaries from 2008 to 2012, as reported in WHO's weekly epidemiological reports to build a model which provided an estimated burden of cholera worldwide. First, they determined the population at risk by using the proportion of the population without access to improved sanitation as a proxy indicator. Then, they used the total people at risk, factoring in a constant to account for reporting performance, to estimate the total annual number of cases of cholera, and derived the incidence rate. Table 2 shows the source of data for each variable.

Table 2

Sources of Data for Each Variable

Variable	Source of data/**d**atabase	Organization
RQ1		
Incidence of cholera	Ali, Nelson, Lopez, & Sack. (2015). Updated global burden of cholera in endemic countries. *PLoS Neglected Tropical Diseases*	
Poverty rate	Human Development Report	UNDP (2017)
Access to safe drinking water	Progress on drinking water, sanitation, and hygiene	UNICEF/WHO (2017)
Access to improved sanitation	Progress on drinking water, sanitation, and hygiene	UNICEF/WHO (2017)
Open defecation	Progress on drinking water, sanitation, and hygiene	UNICEF/WHO (2017)
Income inequality	Human Development Report	UNDP (2017)
Gender inequality	Human Development Report	UNDP (2017)
Adult literacy rate	Human Development Report	UNDP (2017)
RQ2		
Health financing	World Health Statistics 2015	WHO (2015)
Human resources density	World Health Statistics 2015	WHO (2015)

Data Analysis Plan

I used SPSS software Version 24 (IBM Corp, 2016) to run descriptive and inferential analyses. I did not conduct any data cleaning since my source databases already contained processed data. However, prior to conducting the analysis, data were screened for outliers, missing data and examined whether the assumptions underlying logistic regression were met, specifically, the assumption of linearity, outliers, and multicollinearity.

It was anticipated to encounter missing data on some variables because of poor records and incomplete reporting by many African countries, which could result in incomplete socioeconomic records in the databases of international organizations. There exist several approaches to imputing missing data, ranging from simple exclusion of missing data to mean imputation, and regression substitution (Wang, Sedransk, & Jinn, 1992). Discarding records with missing data could lead to selection biases and a reduced sample size which, in turn, could lead to an overestimated standard error (Little & Rubin, 2014). Because poor countries can have challenges in collecting information, and the reporting to international organizations, I assumed that missing data on socioeconomic and health information on African countries were not at random. Also, because of the existing significant gaps in terms of socioeconomic development among African countries, mean imputation could unduly distort the distribution for the concerned variables, leading to underestimated standard deviation (Sterne et al., 2009). I, therefore, for possible missing data, anticipated to conduct regression substitution which considers the performance of each country on other socioeconomic and health indicators.

Descriptive statistics included frequency, measures of central tendency, and measures of dispersion for numeric variables measured at least at an ordinal level (Frankfort-Nachmias and Nachmias, 2008). The algorithm for selecting the appropriate statistical tests to run the inferential analyses depends on the type of variables, whether categorical or numeric and their level of measurement (Field, 2010). The inferential analyses comprised of bivariate and multivariable analyses to test the following research questions and hypotheses:

RQ1: To what extent does social vulnerability (as measured by access to safe drinking water, access to improved sanitation, open defecation, poverty, income inequality, gender inequality, and adult literacy) determine the incidence of cholera in African countries?

$H_0$1: Population-level social vulnerability is not associated with incidence of cholera in African countries.

H_a1: Population-level social vulnerability is associated with incidence of cholera in African countries.

RQ2: To what extent health system capacity (as measured by health financing and density of human resources for health) is associated with incidence of cholera in African countries?

$H_0$2: Health system capacity is not associated with incidence of cholera in African countries.

H_a2: Health system capacity is associated with incidence of cholera in African countries.

RQ3: To what extent does the interaction of social vulnerability and health system capacity impact incidence of cholera in African countries?

$H_0$3: The interaction of social vulnerability and health system capacity does not impact incidence of cholera in African countries.

H_a3: The interaction of social vulnerability and health system capacity impacts the incidence of cholera in African countries.

The bivariate analyses assessed the relationship between each predictor and the outcome independently of the other variables. My dependent variable, the incidence of cholera, was a categorical variable with two levels, *high* and *low* incidence of cholera. The category *low* included countries with cholera incidence between 0 to 2 cases per 1,000 population, and the category *high* incidence included countries with incidence above 2 cases per 1,000 population. The two strata of incidence of cholera were based on the proportion of the population at risk of cholera (Ali et al., 2015). All the independent variables were quantitative indicators expressed as numeric values measured at least at the ordinal level. Multivariable analyses consisted of reducing the predictors into the two components of social vulnerability, and health system capacity, and then assessing the association between the computed composite variables with the incidence of cholera. Tables 3 and 4 respectively present the statistical tests for each independent variable and the summary of the statistical tests by research question.

Table 3

Summary of Data Type and Level of Measurement and Statistical Tests

Variable names	Types of variables (independent/dependent)	Level of measurement	Statistical test for the study outcome
RQ1: Social vulnerability binomial analysis			
Incidence of cholera	Dependent	Categorical/2 levels	
Poverty rate	Independent	Interval	Logistic regression
Access to water	Independent	Interval	Logistic regression
Access to sanitation	Independent	Interval	Logistic regression
Open defecation rate	Independent	Interval	Logistic regression
Income inequality rate	Independent	Interval	Logistic regression
Gender inequality rate	Independent	Interval	Logistic regression
Adult literacy rate	Independent	Interval	Logistic regression
Social vulnerability (composite variable)	Independent	Interval	Logistic regression
Multivariable analysis			
All independent variables	Independent		Logistic regression
RQ2: Health system capacity			
Binomial analysis			
Total health expenditure	Interval		Logistic regression
Out-of-pocket	Interval		Logistic regression
Density of health workforce	Interval		Logistic regression
Health system capacity (composite variable)	Interval		Logistic regression
Multivariable analysis			
All independent variables			Logistic regression
RQ3:			
Multivariable analysis			Logistic regression
Social vulnerability x health system capacity			Logistic regression

Table 4

Type of Analysis and Statistical Test by Research Question

Research question	Type of analysis	Statistical test
RQ1	Binomial	Logistic regression
RQ2	Binomial	Logistic regression
RQ3	Binomial	Logistic regression

The logistic regression report includes statistics to assess the following: (1) how well the model fits the data; (2) the contribution of the predictors to the occurrence of high incidence of cholera in Africa and its level of significance; (3) the strength of the association between the predictors and the outcome and its level of significance.

The statistics assessing the model comprise the log-likelihood and the deviance (Field, 2013). The larger the value of the log-likelihood, the more observations remain unexplained by the model. The deviance and the likelihood ratio served to assess whether the model improves the prediction of the level of the incidence of cholera as compared to the baseline of non-inclusion of predictors in the model. The Cox and Snell's statistic served to gauge the substantive significance of the model and whether the model fitness has improved as a result of including the predictors.

Second, the Wald statistic for each predictor and its level of significance provided information on whether the coefficient of that specific predictor was significantly different from zero. Thus, the Wald statistic indicated the level of contribution of a particular predictor and whether that contribution to a high incidence of cholera was statistically significant (Field, 2013; Hosmer, Lemeshow, & Sturdivant, 2013).

Third, the odds ratio and its confidence intervals indicate how much the probability of the occurrence of a high incidence of cholera change, as an effect of change that occurs in the value of the predictor (Hosmer et al., 2013). Thus, the odds ratio was one of the crucial statistics in interpreting the strength of the association between the predictor and the incidence of cholera.

Threats to Validity

For valid inferences, researchers must take measures to ensure that the changes observed in the dependent variable are indeed the effect of the dependent variable and not attributable to alternative explanations. The researcher, therefore, must take measures to control factors that may jeopardize the internal and external validity of the study (Frankfort-Nachmias & Nachmias, 2008). Internal validity refers to whether the inferences and conclusions drawn from the study results are not erroneous or misleading. Threats to internal validity comprise factors that are external (extrinsic) as well as factors that are internal (intrinsic) to the study process. Extrinsic factors are related to the selection of the study participants and their assignment into groups. Intrinsic factors concern change in the units of analysis or the instrument of measurement, occurring during the research operations. Such factors include the history or time-lapse during the experiment, the biological change in participants or maturation, the loss to follow up or experimental mortality, and the instrumentation or change in the instrument during the experiment (Frankfort-Nachmias & Nachmias, 2008). On the other hand, external validity refers to the process of selecting a sample for results that are generalizable to the population.

One of the significant threats to the validity of my study is related to the sampling method. The lack of random selection of the units of analysis may have introduced selection bias. However, the sample size of 47 countries represented the total population of countries that belong to the regional office for Africa of the WHO. Also, I did not use any instrument to collect data, therefore, no threats to validity related to the instrument of measurement or instrumentation is expected to have occurred. Finally, the threat of loss to follow up or study mortality did not apply to this study since I only analyzed secondary data.

Ethical Considerations

Only population-level secondary data, already in the public domain, were used. The data did not include any information at the individual level. Thus, there were no direct ethical concerns about individuals in study-targeted countries. Nonetheless, I obtained approval from the Walden Institutional Review Board (approval number: 06-28-19-0415624) before proceeding to data collection and analysis. I also sought and obtained, through email, the approval to use the data on the incidence rates of cholera in Africa from Ali et al.'s original study.

Summary

I conducted an ecological study exploring the association between socioeconomic and health system features and the incidence of cholera in African countries members of the Regional Office for Africa of the WHO. I used secondary data collected from databases of international organizations including the WHO, United Nations Children Fund, and The United Nations Development Program. My sample includes all the 47 Member States of the Africa region of the WHO. The convenience sampling method was applied to select units of analysis to be included in the study. The central research questions assessed to what extent social vulnerability and health system capacity impacted the incidence of cholera in Africa. The two main predictors, social vulnerability and health system capacity, were composite variables composed of various indicators.

The dependent variable in the study was the incidence of cholera. I conducted logistic regression to assess the relationship between the predictors and the outcome. The next chapter presents the results of the study, which include the report of the logistic regression statistical tests assessing the fitness of the model, the significance of predictors' contribution to the model, and the strength of the association between the predictors and the incidence of cholera.

Chapter 4: Results

This study was conducted to assess the relationship between population-level socioeconomic indicators, elements of the health system, and incidence of cholera in the 47 countries of the African region of the WHO. The main research question for this study posited that population-level factors of social vulnerability and parameters of health system capacity impact the incidence of cholera in African countries. The elements of social vulnerability include access to safe drinking water, access to improved sanitation, the rate of open defecation, adult literacy, income inequality, and gender inequality. The parameters of health system capacity include health financing and human resources for health. Health financing was defined by the proportion of GDP allocated to health expenditure and the proportion of out-of-pocket payment from the national health expenditure. The density of physicians and the density of nursing and midwifery staff represented the health system capacity in human resources for health. The following research questions were explored:

RQ1: To what extent does social vulnerability (as measured by access to safe drinking water, access to improved sanitation, open defecation, poverty, income inequality, gender inequality, and adult literacy) determine the incidence of cholera in African countries?

$H_0 1$: Population-level social vulnerability is not associated with incidence of cholera in African countries.

$H_a 1$: Population-level social vulnerability is associated with incidence of cholera in African countries.

RQ2: To what extent health system capacity (as measured by health financing and density of human resources for health) is associated with incidence of cholera in African countries?

$H_0$2: Health system capacity is not associated with incidence of cholera in African countries.

H_a2: Health system capacity is associated with incidence of cholera in African countries.

RQ3: To what extent does the interaction of social vulnerability and health system capacity impact incidence of cholera in African countries?

$H_0$3: The interaction of social vulnerability and health system capacity does not impact incidence of cholera in African countries.

H_a3: The interaction of social vulnerability and health system capacity impacts the incidence of cholera in African countries.

This research was an ecological study using secondary data. The data of the outcomes, the incidence of cholera, were drawn from an original study that estimated the global burden of cholera (Ali et al., 2015). It was a categorical variable with two levels of incidence of cholera: high and low. All the independent variables were continuous, mostly measured at ordinal, interval, and ratio levels. Multivariable logistic regression was used to assess the relationship between the independent variables and the outcome. This chapter presents the main outputs and results obtained from the data analysis and comprises two sections. The first section describes the data collection process, a short recap about the sample and the sampling process, and the descriptive characteristics. This section also presents the descriptive statistics, the results of the assessment of the assumption for the logistic regression, the results from the multivariable analyses, as well as the evaluation of the model fitness. The chapter also includes the testing of the research questions and hypotheses. For each research question and hypothesis, the section will present the exact statistics and their associated probability values and the confidence intervals around the statistics when relevant.

Data Collection

Secondary data were obtained from existing reports from the UNDP, WHO, and UNICEF. Data on income inequality (Gini index), gender inequality, rate of the population living below the poverty line, and adult literacy rate were collected from the UNDP's (2015) Human Development Report. Data for the rate of open defecation, access to safe drinking water, and access to improved sanitation were collected from the joint WHO and UNICEF (2017) progress report on drinking water, sanitation, and hygiene. Data on health system financing and human resources for health were collected from the WHO's (2015) world health statistics. The incidence of cholera in African countries was extracted from an original study by Ali et al. (2015). However, Ali et al.'s estimation of the incidence of cholera did not include six countries of the WHO's African region (Algeria, Botswana, Equatorial Guinea, Mauritius, Seychelles, and South Africa) because they were cholera-free at the time of the study. According to a WHO (2016a) report, these six countries remain cholera-free. Given the small size of my sample and the fact that my dependent variable only categorizes countries with a high or low incidence of cholera, I decided to include these countries in my analysis and assign them to the group of countries with a low incidence of cholera. Thus, 23 countries were classified as high incidence of cholera, and 24 countries were classified as low incidence. My sample frame was relatively limited and contained the 47 countries of the African region of the WHO. I used convenience sampling by including all the units in my study without applying any exclusion criteria.

Missing Data

Analysis of the pattern of missing data showed that six cases of 47 (12.8%) had missing data clustered in only one variable (gender inequality). Thus only 7.7% of variables had missing values and, overall, less than 1% (.98 %; six of 612) of values were missing. Because the analysis results suggest a random rather than systematic pattern of missing data, I conducted multiple imputations to fill in the six missing values (IBM, 2017).

Study Results

The descriptive analysis shows that countries of the African region of the WHO have very low average access to improved sanitation (36.6% of the total population), a high rate of people practicing open defecation (23.7%), and relatively high income inequality (Gini index of 44.2767). Also, on average, the proportion of total expenditure on health still represents a negligible percentage of the GDP (1.9%)

Table 5

Descriptive Statistics of the Predictors Included in the Analysis (N = 47)

Predictor	Mean	SD	Median	Min*	Max*
Proportion of population with access to safe drinking water	74.2340	15.00755	77	48	100
Proportion of population with access to Improved sanitation	36.5957	23.17200	30	7	98
Rate of open defecation	23.7021	21.03611	17.0	0	76
Rate of people living below the poverty line	47.5432	20.03950	43.50	6.1	87.70
Gini index rate	44.2767	9.48861	43.0	27.6	65.80
Gender inequality index	.560389	.0878802	0.561	0.380	0.695
Adult literacy rate	65.8870	20.00749	70.80	19.10	95.30
Total expenditure on health as proportion of GDP	1.9021	1.26551	1.60	0.4	5.60
Proportion of general government expenditure as a percentage of total health expenditure	33.7617	19.42634	28.60	7.40	97.0
Proportion of out of pocket as a percentage of total expenditure on health	34.9106	18.88845	36.10	2.50	74.80
Density of physician	2.3022	3.68272	0.90	0.10	20.0
Density of nursing and midwifery	11.1030	12.34366	6.70	1.40	51.10

Assumptions

In line with the prerequisites of a logistic regression analysis (Field, 2013), the following assumptions were assessed: the existence of outliers, linearity, and multicollinearity.

Linearity. Logistic regression assumes a linear relationship between the continuous independent variables and the logit of the dependent variable (Field, 2013). The linearity of the relationship between the continuous variables and the logit of the dependent variable was assessed using the Box-Tidwell (1962) procedure. The Bonferroni correction method was also applied by dividing the level of significance α of

0.05 by the number of the terms in the models. The result indicated a statistical significance of 0.004 (Field, 2013). The results of the linearity assessment showed that all dependent variables were linearly related to the logit of the incidence of cholera. Also, none of the coefficients of the regression of the interaction of each independent variable with its natural log and the dependent variable was statistically significant. Also, none of the correlation coefficients had a p-value lower than the Bonferroni corrected significance value of 0.004. These results indicated that the assumption of linearity was met for all the independent variables.

Multicollinearity. Multicollinearity occurs when two or more independent variables are highly correlated (Field, 2013). I assessed whether multicollinearity existed in my data set by two approaches: (a) the Pearson correlation coefficient, and (b) variance inflation factor together with the tolerance statistics. Scanning the correlation matrix of the predictor variables, I sought to identify a Pearson correlation coefficient between two independent variables equal to or above the cutting point of .80. The correlation matrix showed there were no highly correlated independent variables at the cutting point of .80. However, the matrix showed the existence of a statistically significant ($p < .01$) correlation of .7, indicative of some level of relationship between the following independent variables: access to improved sanitation and the density of physician; proportion of general government expenditure as percentage of total health expenditure and the density of nursing and midwifery; and the density of physician and the density of nursing and midwifery. But the analysis of tolerance and the variance inflation factor showed that no independent variable had a variance inflation factor greater than 10 or a tolerance below .1. However, four variables (total expenditure on health as a proportion of GDP, proportion of general government expenditure as a percentage of total health expenditure, density of physician, and density of nursing and midwifery) have a tolerance below .2, indicative of the existence of potential multicollinearity. Nevertheless, the four variables were included in the analysis for two reasons. First, the tolerance level was higher than the cutting point of .1 indicatives of serious multicollinearity (Field, 2013). Also, although O'Connell and Ann (2005) recommended the removal of highly correlated variables from the model or an increase in the sample size, Midi, Sarkar, and Rana (2010) presented other alternatives in case the independent variables are too important to be replaced and the sample size extension is not feasible, which is the case with my study. Alternatives include the transformation of the independent variables by centering them (using z-score, for instance) or computation of composite variables by running a factor or principal component analysis and using the resulting components as predictors (Midi et al., 2010). I decided to compute composite variables by running the principal component analysis as suggested by Midi et al. (2010). It represented the double advantage of reducing the independent variables into composite variables of social vulnerability and health system capacity while addressing potential multicollinearity.

Outliers. The outlier analysis showed only one case had standardized residuals of 2.233, greater than two standard deviations. Because no case had standardized residuals higher than 2.5 standard deviations, I included all the units in the logistic regression analysis (Field, 2013).

Baseline Analysis

Baseline analysis indicates that without including any independent variable, the best guess is to assume that all countries have a low incidence of cholera as the model correctly classifies 53.7% of cases as a low incidence of cholera.

Model fit. The model fit analysis show that the model is statistically significant ($p < .05$) with a Chi-square of 26.080. The Hosmer and Lemeshow test indicated a Chi-square of 4.595 but not statistically significant ($p = .800$), which indicates that the model is not poorly fitting in predicting the categories of the dependent variable, incidence of cholera. Also, the model summary indicates that the Cox and Snell and the Nagelkerke R^2 are respectively .471 and .629. Thus, the model explains the variation in the dependent variable ranging from 47.1% to 62.9%. The model's overall percentage accuracy in classification (PAC), or the proportion of correctly classified countries, is 78.0 %. Thus, the addition of the independent variables improved the overall model prediction by 24.3% from the baseline. The model also has a sensitivity of 73.7% and specificity of 81.8% in predicting "high incidence" of cholera.

Further, the model had a positive predictive value of 77.78% and a negative predictive value of 78.26%, indicating that the model correctly predicted about 78% of countries with a high incidence of cholera and the same proportion of countries with a low incidence of cholera. The Receive Operating characteristic curve (ROC) was run to assess the ability of the model to discriminate between countries with and without a high incidence of cholera. The results show that the area under the ROC is .903, with a 95% confidence interval ranging from .799 to 1.0 as indicated in Figure 1 below. The results are another indication of excellent discrimination (Hosmer et al., 2013) or the ability to correctly classify.

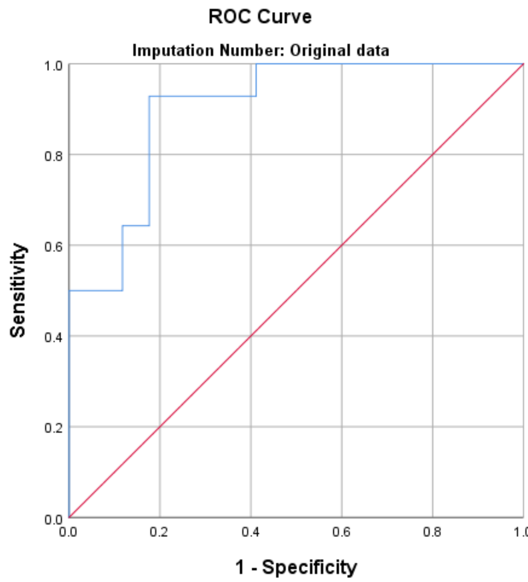

Figure 1. Receive operating characteristic curve.

Contribution of variables in the model. Of all the independent variables, only access to improved sanitation and open defecation had a statistically significant Wald coefficient of respectively 3.040 and 5.072 (p =.034 and p = .010). Table 6 presents the main statistics from the logistic regression.

Table 6

Results of the Logistic Regression (N = 47)

Variable	B	S.E.	Wald	df	Sig	Exp(B)	95% C.I. Exp(B)	
							Lower	Upper
Access to safe drinking water	-.074	.047	.880	1	.112	.929	.847	1.017
Access to improved sanitation	-.101	.048	3.040	1	.034	.904	.823	.992
Rate of open defecation	-.112	.043	5.072	1	.010	.894	.822	.973
Rate of people living below poverty line	.068	.040	2.143	1	.087	1.070	.990	1.157
Gini index rate	.089	.079	2.381	1	.259	1.093	.937	1.275
Gender inequality index	-18.092	11.311	2.686	1	.110	.000	.000	62.642
Adult literacy rate	-.073	.050	1.071	1	.150	.930	.842	1.027
Total expenditure on health as proportion of GDP	.451	.732	.153	1	.538	1.570	.374	6.590
Proportion of general government expenditure as percentage of total health expenditure	.048	.061	.179	1	.430	1.050	.931	1.184
Proportion of out of pocket as percentage of total expenditure on health	.012	.035	.651	1	.743	1.012	.944	1.084
Density of physician	-.987	.720	.214	1	.170	.373	.091	1.530
Density of nursing and midwifery	.159	.134	.001	1	.235	1.172	.901	1.525
Constant	16.318	9.019	.705	1	.071	1818.76		

Transformation of Variables to Test Hypotheses

To test the hypotheses, I had to reduce and convert the twelve variables into only the two components stated in the hypotheses: Social vulnerability and health system capacity. As described in Chapter 3, social vulnerability was composed of the following seven independent variables: access to safe drinking water, access to improved sanitation, open defecation, poverty, income inequality, gender inequality, and adult literacy. Health system capacity included five predictors: (a) proportion of GDP allocated to health expenditure, (b) total government expenditure, (c) out of pocket payment, (d) density of physicians, and (e) density of nurse and midwifery. Two approaches to variables reduction could be used. The first approach is the z-score transformation, and the second is the principal components analysis (PCA). The z-score transformation consists of converting all variables into z-score because z-scores are independent of the unit of analysis. Z-scores transformation can be used to convert data with different units into variables that have the same scale and to sum the score without altering the actual value of each variable (Song, Lin, Ward, & Fine, 2013).

Thus, the z-score of all predictors belonging to either social vulnerability or health system capacity composites could be added to make the values of each one of the two composite variables. However, simply summing z-scores can mask the correlation that exists between variables that potentially belong to a group or a component. Also, summing z-scores does not consider the weight of each variable in each component variable (Song et al., 2013). Therefore, I conducted the principal component analysis to reduce the twelve independent variables in components. The following section presents a summary of data reduction using the principal component analysis.

Principal component analysis or PCA is one of the approaches to extracting underlying dimensions of a data set (Field, 2013). Thus, the principal component analysis establishes linear components that exist within the data set. It also establishes how much various variables contribute to a particular underlying component. The extraction of principal components reduces the number of several numbers of variables to a manageable number of principal components or composite variables, which can then be used to run statistical tests.

The preliminary analysis of the principal components of my data set indicated that the Kaiser-Meyer-Olkin measure was .758, which satisfactorily verified the sampling adequacy for the principal component analysis or PCA (Hutcheson & Sofronie, 1999). The analysis of the variance explained, as well as the scree plot, showed that three components had Eigenvalues greater than the Kaiser's criterion of 1. The first component explained 45.846% of the total variance; the second component explained 15.601%, and the third component explained 9.633% of the total variance. Cumulatively, the three components accounted for 70.046% of the total variance. Thus, based on the scree plot and the Eigenvalue cut-off value, Kaiser's criterion of 1, three components were retained (Field, 2013). Table 7 and Table 8 respectively show the total variance explained, and the factor loading greater than .3 after oblique rotation (direct oblimin).

Table 7

Total Variance Explained (N = 47)

Component	Initial Eigenvalue			Extraction sums of squared loadings			Rotation sums of squared loadings total
	Total	% of variance	Cumulative %	Total	% of variance	Cumulative %	Total
1	5.694	47.454	47.454	5.694	47.454	47.454	3.834
2	1.553	12.942	60.396	1.553	12.942	60.396	3.375
3	1.158	9.651	70.046	1.158	9.651	70.046	3.881

Table 8

Factor Loading (N = 47)

Variable	Component		
	1	2	3
Density of physician	.844		–.512
Rate of people living under the poverty line	–.822		
Density of nurse and midwifery	.733	.548	–.555
Access to safe drinking water	.626	.495	–.362
Total expenditure of health as % of GDP		.912	–.328
Proportion of out-of-pocket payment		–.868	
Proportion of general government expenditure on health as a percentage of total expenditure on health	.629	.726	–.447
Adult literacy rate	.385	.445	–.829
Access to improved sanitation	.668	.332	–.773
Open defecation rate			.768
Gender inequality	–.596	–.417	.719
Income inequality (Gini Index)		.432	–.619

Following the extraction, weighted values were computed for each component based on the eigenvalue of the variables loaded on each factor. Table 9 presents the loading after rotation and the weighted loading on each of the three extracted components. Tables 9 also shows the variables that compose each one of the three extracted components.

Table 9

Weighted Loading and Variables on Each Extracted Component (N = 47)

Component	Variable	Label	Loading	Weighted loading
1	Density of physicians	SocVuln_A	.844	0.235
	Rate of people living under the poverty line		–.822	–0.223
	Density of nurse and midwifery		.733	0.178
	Access to safe drinking water		.626	0.130
2	Total expenditure of health as % of GDP	HealthSystCap	.912	0.332
	Proportion of out-of-pocket payment		–.868	0.301
	Proportion of general government expenditure on health as a percentage of total expenditure on health		.726	0.210
3	Adult Literacy rate	SocVuln_B	–.829	–0.185
	Access to improved sanitation		–.773	–0.161
	Open defecation rate		.768	0.159
	Gender Inequality		.719	0.139
	Income inequality (Gini Index)		–.619	–0.103

Components 1 and 3 represent social vulnerability, respectively labeled as “SocVuln_A” and SocVuln_B. Component 2 summarizes the health system capacity, labeled as “HealthSystCap.”

Testing hypotheses. Logistic regression tests were conducted on the composite variables or components generated by the principal component analysis. The section below presents the results of hypotheses testing by running the logistic regression on components generated by the principal component analysis

Following the extraction of the principal components, a binomial logistic regression was conducted to test the three hypotheses, assessing the association between the predictors SocVuln_A, SocVuln_B, and HealthSystCap and the incidence of cholera.

RQ1: To what extent is population-level social vulnerability (as measured by access to safe drinking water, access to improved sanitation, open defecation, poverty, income inequality, gender inequality, and adult literacy) associated with incidence of cholera in African countries?

> $H_0 1$: Population-level social vulnerability is not associated with incidence of cholera in African countries.

To test the first null hypothesis, the first model included the two components of social vulnerability, respectively, SocVuln_A and SocVuln_B. The model was not statistically significant ($\chi^2(2) = 3.018$, $p = .221$). The model explained 8.3% of the variance (Nagelkerke $R^2 = .083$). The model correctly classifies 61.7% of the cases, with a sensitivity of 42.9%, a specificity of 76.9%. The model also had a positive predictive value (PPV) of 60% and a negative predictive value (NPV) of 62.5%. The two components of social vulnerability had odds ratio or Exp(B) of respectively 1.326 for SocVuln_A (95% CI: .947 – 1.856) and 1.000 for SocVuln_B (95% CI: .932 – 1.035). Individually, none of the two components of social vulnerability had a statistically significant association with the incidence of cholera. SocVuln_A had a Wald statistic of 2.7 ($p = .100$), while SocVuln_B had a Wald statistic of .461. However, the results also indicated that the effect of the interaction of the two components of social vulnerability (SocVuln_A and SocVuln_B) was statistically significantly associated with the incidence of cholera (Wald = 4.283; $p < .05$; OR = 1.080; 95% CI: 1.004 – 1.162). In logistic regression, the statistical inferences decision cannot be based on the main effect alone while the interaction indicates an association (Jaccard, (2001; Frost, 2019; Norton, Wang, & Ai, 2004;). Therefore, based on the presence of a statistically significant association between the interaction of the two components of social vulnerability and the outcome, the first null hypothesis was rejected. Although marginal, the positive coefficient (B = .077) indicates a positive association between the interaction of the two components of social vulnerability and the incidence of cholera. Table 10 summarizes the statistics of the equation testing the first hypothesis.

Table 10

Statistics Output of Logistic Regression Testing H_0*1 (N = 47)*

Variable	B	S.E.	Wald	df	Sig	Exp(B)	95% C.I. for EXP(B)	
							Lower	Upper
SocVuln_A	.282	.172	2.700	1	.100	1.326	.947	1.856
SocVuln_B	-.018	.027	.461	1	.497	.982	.932	1.035
SocVuln_A x SocVuln_B	.077	.037	4.283	1	.038	1.080	1.004	1.162
Constant	-4.265	2.487	2.940	1	.086	.014		

RQ2: To what extent is health system capacity (as measured by health financing, and density of human resources for health) associated with incidence of cholera in African countries?

$H_0$2: Health system capacity is not associated with incidence of cholera in African countries.

The health system variables extracted through PCA include the proportion of GDP allocated to health expenditure, the proportion of out-of-pocket payment, and the proportion of total health expenditure incurred by the government. The results of the logistic regression to assess the relationship between health system capacity and the incidence of cholera show that the model, including the two components of social vulnerability and the health system capacity, was statistically significant ($\chi^2(3) = 8.036$; $p < .05$). The model explained 21% of the variance (Nagelkerke R^2). The model had a percentage accuracy in classification (PAC) of 70.2, a sensitivity of 61.9%, a specificity of 76.9, a PPV of 68.42%, and an NPV of 71.43%. The health system had a statistically significant association with the incidence of cholera (Wald = 3.999; $p < .05$). The odds ratio was .792 (95% CI: .630 – .995). However, the relationship between health system as the predictor and the incidence was negative (B = -.233), meaning that when health system capacity increases by one unit, the logit of the incidence of cholera decreases by .233. Nevertheless, the second null hypothesis was rejected.

Table 11

Statistics Output of Logistic Regression Testing H_0*2 (N = 47)*

Variable	B	S.E.	Wald	df	Sig	Exp(B)	95% C.I. for EXP(B)	
							Lower	Upper
HealthSystCap	-.233	.117	3.999	1	.037	.792	.630	.995
Constant	.721	3.511	.042	1	.837	2.056		

RQ3: To what extent does the interaction of social vulnerability and health system capacity impact incidence of cholera in African countries?

*H*0 3: The interaction of social vulnerability and health system capacity does not impact incidence of cholera in African countries.

The effect of the interaction between the two components of social vulnerability and the health system on the incidence of cholera was also tested using logistic regression. The results indicated that the model was statistically significant ($\chi^2(5)$ = 21.669; p < .01). The model explained 49.4% of the variance (Nagelkerke R^2), which is a 41% improvement from the baseline. The model with the interaction between social vulnerability and health system capacity also had an overall higher percentage accuracy in classification (PAC) than the baseline model, from 55.3 to 80.9%, a sensitivity of 76.2%, and a specificity of 84.6%. The model had a positive predictive value of 80% and a negative predictive value of 81.41%. Although weak (with a coefficient B of only .002), the association between the interaction of the two predictors and the outcome was positive and statistically significant (Wald = 6.132; *p* < .05; odds ratio or Exp(B) = 1.004; 95% CI: 1.002 – 1.009). Therefore, the third null hypothesis was also rejected. Table 12 presents the statistics of the models assessing the effect of predictors extracted through PCA and their interaction on the incidence of cholera.

Table 12

Statistics of the Logistic Regression Testing the Effect of the Interaction of Predictors Extracted Through PCA (N = 47)

Variable	B	S.E.	Wald	Sig.	Exp(B)	95% C.I. EXP(B)	
						Lower	Upper
SocVuln_A	.089	.264	.114	.736	1.093	.651	1.834
SocVuln_B	-.644	.486	1.758	.185	.525	.203	1.361
HealthSystCap	-.233	.117	3.999	.037	.792	.630	.995
SocVuln_A x SocVuln_B	.077	.037	4.283	.038	1.080	1.004	1.162
HealthSystCap X SocVuln_A	.082	.049	2.775	.096	1.085	.986	1.195
HealthSystCap X SocVuln_B	-.027	.014	3.812	.051	.973	.947	1.000
HealthSystCap X SocVuln_A X SocVuln_B	.002	.001	6.132	.013	1.004	1.002	1.009
Constant	3.146	5.052	.388	.533	23.242		

Table 13 summarizes the statistics testing the three hypotheses based on composite variables computed from the principal component analysis.

Table 13

Statistics of Models With Composite Variables (N = 47)

Statistics	RQ_1 & H_{o1}	RQ_2 & H_{o2}	RQ_3 & H_{o3}
Model Chi-Square (χ^2)	3.018	8.036	21.669
Model significance (*p*)	.221	<.05	< .01
Nagelkerke R^2	.083	.210	.494
Model PAC* (%)	61.7	70.2	80.9
Model sensitivity (%)	42.9	61.9	76.2
Model specificity (%)	76.9	76.9	84.6
Positive predictive value (%)	60	68.42	80
Negative predictive value (%)	62.5	71.43	81.41
Odds Ratio or Exp(B)	1.080	.792	1.004
95% CI Exp(B)	1.004 – 1.162	.630– .995	1.002 – 1.009
Wald Statistic	4.283	3.999	6.132
Wald Significance (*p*)	<.05	< .05	< .05
Decision on H_o	H_{o1} Rejected	H_{o2} Rejected	H_{o3} Rejected

Summary

Binomial logistic regression was conducted to assess the relationship between the predictors and the outcome, the incidence of cholera. Two series of model analyses were conducted. The first series of models assessed the relationship between the twelve primary variables and the incidence of cholera. Variables in the model included: access to safe drinking water, access to improved sanitation, the rate of open defecation, income inequality, gender inequality, the proportion of people living below the poverty line, adult literacy rate, the percentage of GDP allocated to the health expenditure, the proportion of government expenditure allocated to health, the proportion of health expenditure that is incurred by the population as out of pocket payment, the density of physicians, and the density of nursing and midwifery.

The assessment for outliers showed that there was no outlier among the cases. Further, the assessment of the assumptions of linearity and multicollinearity indicated that these assumptions were met for the logistic regression. Furthermore, the Bonferroni correction method, as well as the Box-Tidwell procedure, showed a statistical significance at .004. They indicated that none of the coefficients of the regression between the interaction of each independent variable and its natural log as the predictor and the dependent variable was statistically significant ($p < 0.004$). The results of the logistic regression showed that the model was statistically significant $\chi2(12) = 26.080$, $p < .05$. The model explained 62.9% of the variance and correctly classified 78.0% of cases. The model had a sensitivity of 73.7% and specificity of 81.8%, a positive predictive value of 77.78%, and a negative predictive value of 78.26%. The model also had an excellent discriminating capability. The area under the curve of the Receive Operating characteristic (ROC) was .903 (95%CI: .799 – 1.0).

Of the 12 predictors, only two had a statistically significant relationship ($p < 0.05$) with the incidence of cholera: access to improved sanitation and the rate of open defecation. The relationship between these two predictors and the incidence of cholera were negative, suggesting that the increase in access to improved sanitation or open defecation will decrease the odds in the incidence of cholera. However, a negative relationship between open defecation and the incidence of cholera does not scientifically make sense. Although the relationship between the two independent variables and the incidence of cholera were statistically significant, the odds ratios were relatively marginal. Countries with a high rate of access to improved sanitation or a high percentage of open defecation have respectively .8 and .9 times higher odds to have a high incidence of cholera. The second series of analyses were conducted after reducing the number of variables by computing composite variables social vulnerability and health system capacity to allow testing of the three hypotheses and research questions.

The reduction of primary variables into composite variables was conducted through the principal component analysis (PCA). The Principal component analysis aimed to establish the linear components that existed within the data set and how much the primary variables contributed to a particular underlying component, thus ultimately reducing the number of variables to a manageable number of principal components.

The principal component analysis (PCA) extracted three main components from the twelve variables. Two components loaded variables on social vulnerability, while one component mainly loaded on the health system capacity. The logistic regression tests assessing the association between the three components from the PCA and the incidence of cholera indicated that the main effects of the two components loading on social vulnerability were not statistically significant. However, the interaction between the two social vulnerability components was statistically associated with the incidence of cholera. The third component, which loaded on the health system, was also statistically associated with the incidence of cholera. Further, the interaction of the two social vulnerability and health system components were statistically associated with the incidence of cholera. Thus, the three null hypotheses were all rejected and alternative hypotheses accepted. The next chapter discusses the interpretation of these results, the limitations of the study, and recommendations for future research.

Chapter 5: Discussion, Conclusions, Recommendations

I explored the factors associated with the incidence of cholera in the 47 countries of the African region of the WHO. Although cholera has been eliminated in most of the world, it remains a significant public health problem in the African region of the WHO. My study was anchored in the ecosocial theory of the distribution of diseases, which posits that the causes of incidence, as opposed to the cause of cases, are population-level factors that drive the emergence, dissemination, or amplification of infectious diseases in a community. The reports on cholera tend to indicate that cholera is endemic in countries that score poorly on socioeconomic indicators. However, not all countries with poor performance on socioeconomic parameters are endemic to cholera. Likewise, some countries that have relatively better socioeconomic performance in Africa continuously report cholera outbreaks. Therefore, I hypothesized that socioeconomic factors were drivers of incidence of cholera in the African region, while health system capacity determined the ability to control cholera outbreaks rapidly. I categorized socioeconomic drivers of emergence and amplification, which impact the occurrence and recurrence of cholera outbreaks, as social vulnerability, and parameters of the health system constituted health system capacity. Social vulnerability included the rates of access to safe drinking water, access to improved sanitation, open defecation, adult literacy, and poverty as well as the indices of income inequality and gender inequality. The parameters of health system capacity included density of physicians and nurses and health system financing. All predictors were continuous variables. The outcome was the incidence of cholera, expressed as a categorical variable with two levels: high and low. Logistic regression was conducted to assess the association between the independent and the dependent variables. The logistic regression was also conducted to evaluate the relationship between the interaction of the predictors and the outcome. I performed two sets of analyses, the first by running a logistic regression on all the 12 variables. To test my hypotheses, I conducted the second analysis, which consisted of first reducing the variables into manageable components. The principal component analysis was performed and yielded the three principal components. Then, logistic regression was used to assess the association between the principal components and the dependent variable.

The results indicated that the model, which included all 12 independent variables, was statistically fit in predicting the categories of the outcome. However, only two predictors, rate of access to improved sanitation and rate of open defecation, had a statistically significant association with incidence of cholera. The second model, which included the three components obtained from the CPA, was also fit for the data. Although the main effect of social vulnerability was not associated with incidence of cholera, the interaction of social vulnerability components was statistically associated with incidence of cholera. Likewise, the health system capacity was also associated with the outcome. The discrepancy from the results obtained with the two analysis approaches is discussed in this chapter.

Interpretation of Findings

Baseline Analysis

The results of this study showed that of the 12 predictors, only two were associated with incidence of cholera: rate of access to improved sanitation and rate of open defecation. Thus, access to safe drinking water, rate of people living below the poverty line, income inequality (Gini index), gender inequality, and adult literacy were not associated with incidence of cholera. There was also no association between the proportion of GDP allocated to health expenditure, proportion of health expenditure covered by the government, density of physicians, and density of nursing and midwifery and incidence of cholera. These results seem to contradict several previous studies that found an association between incidence of cholera and most of these predictors. In fact, several studies have shown that numerous socioeconomic indicators were associated with incidence of cholera. For instance, Cerda and Lee (2013) found that rates of access to safe drinking water and access to improved sanitation were associated with incidence of cholera in the Americas. Root et al. (2013) found a correlation between income and incidence of cholera at the household level in Bangladesh. Those results were corroborated by Leckebusch and Abdussalam (2015) in Nigeria and Bwire et al. (2017) in Uganda, who, respectively, found that absolute poverty and low-income at household level were predictors of incidence of cholera. From a study in South America, Caribbean, North America, and Latin America, Ackers (1998) found a negative association between female literacy rate and cumulative incidence of cholera. Similar results showing a negative relationship between the rate of adult literacy and the incidence of cholera were also found in Nigeria (Leckebusch & Abdussalam, 2015), in Kenya (Cowman, 2015), and in India (Ali et al., 2017).

However, my results are also consistent with findings from other previous studies. For instance, Beckfield (2004), Babones (2008), King et al. (2010), and Pop et al. (2013)

found no relationship between income inequality and population health outcomes, such as life expectancy, child mortality, or child immunization. Nevertheless, the indication of no correlation between the rate of access to safe drinking water and incidence of cholera seems to defy even the etiology and epidemiology of cholera. Indeed, it is now established that transmission of cholera results from the consumption of contaminated drinking water or food (Bain et al., 2014; Bwire et al., 2017; Jutla et al., 2013; Kwesiga et al., 2018; Snow, 1856). The seemingly contradicting findings may have resulted first from the differences in study designs. Most of the studies mentioned earlier were not ecological studies but surveys, with data collected at the individual level, not at the group or population level. While the assessment of the relationship between socioeconomic indicators and the occurrence of cases of cholera may show an association at the individual level, such findings may not necessarily be corroborated when applying the analysis with data collected at the population level in an ecological study (Frankfort-Nachmias & Nachmias, 2008). Nevertheless, Pop et al.’s (2013) ecological study found an association between income inequality and other health outcomes, although the analysis did not specifically target the incidence of cholera as one of the outcomes. Second, the effect size may differ depending on the analysis being conducted at the individual or population level. The difference in effect size may be significant enough to be obscured or revealed depending on the study design.

Further, because effect size is affected by sample size (Ellis, 2010), the relatively limited number of units of analysis may have had impacted the capacity of this study to detect a relatively small effect size. Indeed, the power analysis indicated that my sample only had 57% power of detecting a small effect size of .02. The impact of the sample size on the detection of the effect size may explain the discrepancy in study findings.

My results have also indicated that the relationships between the two predictors associated with the incidence of cholera, the rate of access to improved sanitation, and the rate of open defecation, were negative. The negative relationship suggests that an increase in access to improved sanitation or open defecation will decrease the odds in the incidence of cholera. A negative association between the rate of access to improved sanitation is in line with the epidemiology of cholera. However, a negative relationship between the rate of open defecation and the incidence of cholera defies the rational, although the odds ratio of .894 (95% CI: .822 – .973) seems marginal. Open defecation increases the likelihood of contamination of water source or food and, incidentally, the risk of transmission of water-borne diseases such as cholera. Therefore, an increase in the rate of open defecation should logically be associated with an increase in the incidence of cholera.

Various studies have found such a positive relationship between the rate of open defecation and incidence of cholera at the population level. For instance, in Kenya, Cowman et al. (2017) found that districts with lower rates of open defecation had a lower incidence of cholera. These seemingly absurd findings may result from study design, using countries as units of analysis. Analyzing at the population level assumes that all communities or geographic entities in a country have a similar rate of open defecation or incidence of cholera. However, disparities within a country can be significant. High rates of open defecation or even high incidence of cholera may be confined in a portion of a country instead of being generalized to the whole country. For instance, WHO's report on cholera indicated that some African countries endemic to cholera have an open defecation rate 4 to 7 times lower than the African average. On the other hand, some countries that are not endemic to cholera have a rate of open defecation 2 to 3 times higher than the African average. Further investigation is needed to fully understand these discrepancies.

Hypotheses Testing

My hypotheses respectively stipulated that: (1) social vulnerability was associated with the incidence of cholera, (2) health system capacity was also associated with the incidence of cholera, and (3) the interaction of social vulnerability and health system capacity impacted the incidence of cholera. The three hypotheses were tested after extraction of the principal components, using PCA. Of the three main components or composite variables, two were loaded on social vulnerability labeled as *SocVuln_A* and *SocVuln_B*. The first composite of social vulnerability or SocVuln_A comprised of four predictors: the rate of access to safe drinking water, the rate of people living under the poverty line, the density of physicians, and the density of nurses and midwifery. SocVuln_B included five primary independent variables: the adult literacy rate, the rate of access to improved sanitation, the rate of open defecation, the gender inequality index, and the income inequality index or Gini Index. The last composite variable loaded on health system capacity and was labeled as "HealthSystCap." It comprised three indicators: the total expenditure of health as a percentage of the GDP, the proportion of out-of-pocket payment, the general government expenditure on health as a percentage of total expenditure on health.

The results showed that none of the two components of social vulnerability were associated with the incidence of cholera. However, there was a positive and statistically significant association between the interaction of the two components of social vulnerability and the incidence of cholera. The results also indicated that there was a statistically significant but negative relationship between the health system capacity and the incidence of cholera. The results of testing the association between the interaction of composite predictors and the outcome showed that the interaction of the two components of social vulnerability and the health system capacity had a positive and statistically significant association with the incidence of cholera.

Thus, the results suggest that the interaction of several factors, more than the effect of each parameter individually, impacts the incidence of cholera. This finding is consistent with the theory of the web causation of diseases from which derived the concept of social determinants of diseases. The conceptual framework of social determinants of diseases stipulates that the underlying cause of occurrence and dissemination of diseases is a web of complex interaction and feedback loops between the socioeconomic and environmental conditions in which people live (Catalyst, 2017). These results are also consistent with findings from other previous studies on the effect of social vulnerability on the incidence of infectious diseases. For instance, Stanturf, Goodrick, Warren, Charnley, and Stegall (2015) found a geographic association between components of social vulnerability at the district level and cases of Ebola in Liberia. However, Stanturf et al. (2015) operational definition of social vulnerability included elements such as food insecurity, population displacement, access to free medical care, and access to free land. Cordoba and Aiello (2016) also noticed that social factors influenced the transmission of influenza as well as the ability to control influenza outbreaks in the United States of America. Their social factors included access to health care and school and workplace policies. Bishwajit and Ghosh (2014) also described the nexus between social determinants such as poverty, illiteracy, food insecurity, and infectious diseases, including HIV and tuberculosis in South Asia.

Also, the negative association between the health system capacity and the incidence of cholera seems to suggest that the health system capacity had a negative impact on the high incidence of cholera, which is conceptually odd. First, these apparent illogical results may be due to the intrinsic values of the indicators that loaded on this composite variable in covering all aspects of the health system capacity. Indeed, health system capacity, like state capacity, is a broad concept that can be measured through numerous proxy indicators. The three indicators that loaded on the composite variable HealthSystCap (total expenditure of health as a percentage of the GDP, the proportion of out-of-pocket payment, and general government expenditure on health as a percentage of total expenditure on health) may not necessarily capture other aspects of the health system capacity. Other factors of the health system capacity which were not included in the analysis, such as the universal health care coverage, and availability of vaccine and medicines may as well contribute to the effect of the health system capacity on the incidence of cholera in a country. For instance, Filauri (2010) found a positive and statistically significant correlation between state capacity and the incidence of cholera. However, she measured state capacity by indicators such as control of corruption, external debt stocks, GDP, government effectiveness, foreign direct investment, political stability, regulatory quality, and secondary-school enrolment.

Second, the negative association between the health system capacity and the incidence of cholera may also be due to the opposing directions of the three predictors included in the composite variable on the incidence of cholera. On the one hand, a positive association between the proportion of out-of-pocket payment and the incidence of cholera could epidemiologically be justifiable. Studies have shown that the proportion of out-of-pocket was one of the barriers to care-seeking behaviors (Xu et al., 2003), which in turn increases the risk of transmission and amplification of diseases. But, on the other hand, such a positive association between the two other predictors (total expenditure of health as a percentage of the GDP and general government expenditure on health as a percentage of total expenditure on health) and the incidence of cholera would appear awkward. The two indicators measure the level of the government's participation in health expenditure in a country. It can be expected that as the portion of the government's participation in health expenditure grows, population health outcome also improves, including in disease prevention and control. In line with the above common sense, Kim and Lane (2013) found a negative relationship between government expenditure on health and infant mortality. Thus, the tree indicators loaded on the composite HealthSystCap had effect in opposite directions on the incidence of cholera. On one side, prevention and control of cholera require adequate funding. On the other, substantial out-of-pocket payment limits the capacity to control cholera. Thus, while the proportion of out-of-pocket negatively affects the outcome, the other two indicators are meant to positively impact the dependent variable. In fact, in the PCA, the indicator "proportion out-of-pocket payment" negatively loaded on the component HealthSystCap while the two others loaded positively on the composite variable considering together of the loading. Therefore, considering the factor loading, and the intrinsic values of the predictors, the negative association between HealthSystCap and the incidence of cholera makes sense. Finally, the results indicated that the interaction of the two components of social vulnerability and the health system capacity had a positive association with the incidence of cholera. These results suggest, again, that interaction, more than individual factors alone, has an impact on the incidence of cholera.

Limitations of the Study

The results of my study are subject to several limitations related to the design of the study, the sampling frame, and sample size, as well as the nature of secondary data, obtained and used in the analysis, especially the classification of the dependent variable into two categories. First, the interpretation of the results, generated from ecological research, is strictly limited and applied to the groups used as units of analysis, which were, in this case, the African countries. As such, the results cannot be extrapolated to apply to sub-national entities such as provinces, districts, or counties, even more so to individuals within the African countries. Also, the study results present a picture of homogenous countries painted in one single color. It failed to capture the differences between sub-national entities within each country. There are similarities but also differences in terms of socioeconomic indicators, which in some countries, especially in geographically large countries, can be significant between different sub-national geographic entities within a country. The rates of socioeconomic indicators such as access to safe water, open defecation, people living under the poverty line, the density of physicians and nurses may significantly differ from one sub-national entity to another. There may also be significant differences between urban and rural areas or densely and sparsely populated areas. Likewise, the same discrepancies may even exist in the incidence of cholera.

Another limitation is the complexity of the concept of social vulnerability and health system capacity. Although the selected indicators can be used as a proxy to measuring the two concepts, their effect and impact on health outcomes go beyond the scope of the 12 variables selected for this study. For instance, social norms and culture, as well as the political context, may play a significant role and modulate population-level vulnerability. Likewise, governance can be a significant modulator of the health system capacity. Therefore, the results of my study may have not necessarily captured different contours of what determines social vulnerability and health system capacity as related to the incidence of cholera. Moreover, the categorization of the incidence of cholera into *high* and *low* can mask differences in the magnitude of the incidence of cholera between the units of analysis.

Finally, the sample frame was limited, with only 47 units. Hence, no probabilistic sampling or a random selection of units was applied. Thus, my sample cannot be deemed representative of other developing countries, including other African countries that are not members of the African region of the WHO. Therefore, the results cannot be generalized to countries other than those included in the study. The limited sample size may also have affected the power to detect small effect sizes of the relationship between the predictors and the outcome.

Recommendations

Infectious diseases remain a significant public health problem in African countries. They account for 85% of the 120 to 150 public health events, or emergencies, recorded every year by the WHO in Africa. Of this, cholera outbreaks account for 25 to 30% of all the events (WHO, 2017b). There is an urgent need to control then eliminate cholera which depletes already impoverished African countries of millions of dollars. Social vulnerability to infectious diseases such as cholera amplifies the emergence and transmission of cholera within a country. Cholera outbreaks, in turn, exacerbate social vulnerability, thus creating a vicious cycle. Many studies have been conducted on cholera and its risk factors in Africa. However, most of the previous studies characterized risk factors at the individual level, which may provide a better understanding of the disease physiopathology or epidemiology and, thus can guide in the development of behavioral interventions. However, the characterization of individual-level risks alone may not be sufficient for the development of population-level public health policies and strategies. Instead, the determination of population-level risks or other social determinants of health such as the rate of health literacy, population trust in government, or the availability of social support, is often required. Such parameters can be included in future studies of population-level determinants of cholera outbreaks in Africa.

The results of my study have shown that the interaction between several social vulnerability and the health system capacity impacts the incidence of cholera. Therefore, there is a need to define what constitutes social vulnerability or health system capacity in relation to infectious diseases in general, and cholera in particular. The identification of core indicators of social vulnerability and health capacity as related to cholera can guide the development of indices, which can then be used to determine the level of social vulnerability and health system capacity in relation to the risk of cholera. Other future studies can include more or other indicators to better characterize social vulnerability or health system capacity. Also, the results have indicated that the interactions between social vulnerability and health system impact the incidence of cholera. However, they did not determine which factors are the most critical in those interactions. Further studies can quantitatively assess the level of contribution of each element of social vulnerability (such as the density of physicians and nurses, and gender inequality) and health system capacity such as health financing, in the interactions that affect the incidence of cholera in Africa. Finally, my study approach can be applied for sub-national levels to identify and map the geographic sub-national entities that are more at risk of cholera outbreaks based on the parameter of social vulnerability or health system capacity used in this study. Thus, an ecological approach can be used to map, within a country, districts, provinces, or counties that have a higher social vulnerability to cholera or lower health system capacity to respond to cholera. The same approach can also be used to determine countries that present higher vulnerability or lower capacity, within African geographic entities such as West Africa, East, and the horn of Africa, or Southern Africa,

Implications

Governments and multilateral organizations such as the World Bank, the WHO, and other agencies are engaged in the efforts to eliminate cholera in Africa. Effective policies and strategies are being developed and implemented. The study results have shown that it is a net of factors and their interactions rather than each factor individually that impacts the incidence of cholera. Therefore, policy and strategies to control cholera in Africa should target several factors of vulnerability and health system capacity at the same time rather than focusing on interventions targeting one or only a few indicators. The study results also suggest that the cholera elimination programs should indeed be one of the flag bearer programs of "health in all policies" approach, making the elimination of cholera a convergent target and focus of several sectors, not health alone. Thus, various development programs such as education, poverty alleviation, and fight against economic and gender inequalities should play a role in cholera prevention efforts along with the health sector. At household and individual levels, such a convergent program against cholera can contribute to addressing several social determinants of health. Also, such a program will ultimately help to alleviate underdevelopment and support the attainment of sustainable development goals while controlling and eliminating cholera at the same time.

Conclusions

The purpose of my study was to assess the relationship between social vulnerability, health system capacity, and the incidence of cholera in the countries of the African region of the WHO. Although the literature showed evidence of the correlation between some socioeconomic indicators or between social vulnerability and infectious diseases, including cholera, the impact of the health system capacity and its interaction with social vulnerability on the incidence of cholera was yet to be assessed. Anchored in the ecosocial theory of the distribution of diseases, I used the quantitative and ecological approach. All the 47 countries of the African region of WHO were included in the analysis. Data of twelve initial predictors were collected to characterize social vulnerability and health system capacity. Two series of inferential analyses were conducted. First, bivariate and multivariable logistic regression tests explored the association between the twelve independent variables and the outcome. Second, logistic regression was also conducted to test the hypotheses after the reduction of the twelve predictors into three main components, using the Principal Component Analysis or PCA. The results indicated that of the twelve initial predictors, only two had an association with the outcome.

The test of hypotheses showed that the two composite variables of social vulnerability (SocVuln_A and SocVuln_B) were not associated with the incidence of cholera. However, their interaction was positively associated with cholera. The results after reduction also showed that the health system was negatively associated with cholera. These inconclusive results made sense, given the elements that formed the composite variable HealthSystCap. These results also indicated that the interaction between social vulnerability and the health system was positively associated with the incidence of cholera. These results are consistent with several previous studies and in line with the concept of social determinants of health. However, the results are subject to some limitations. The sample may not be representative since the probabilistic sampling or a random selection of units was not conducted. Hence, the results of this study cannot be generalized to other countries. Due to the ecological design of the study, the results also cannot be applied to sub-national geographic entities or even to individuals within those countries. Future research needs to further explore the contribution of other indicators of vulnerability and health system capacity in relation to cholera. Nevertheless, these results suggest that policy and strategies to control or eliminate cholera should be multisectoral and target several indicators from various sectors and programs. Such programs can contribute to addressing several determinants of health while achieving the goal of cholera elimination in Africa.

References

Abdussalam, A. F. (2015). A spatial autocorrelation analysis of cholera patterns and social risk factors in Nigeria. *Academic Research International*, *6*(4), 136–147. Retrieved from http://www.journals.savap.org.pk/

Ackers, M. L. (1998). Are there national risk factors for epidemic cholera? The correlation between socioeconomic and demographic indices and cholera incidence in Latin America. *International Journal of Epidemiology*, *27*(2), 330–334. Retrieved from doi:10.1093/ije/27.2.330

Akanda, A. S., Jutla, A. S., Gute, D. M., Sack, R. B., Alam, M., Huq, A., … Islam, S. (2013). Population vulnerability to biannual cholera outbreaks and associated macro-scale drivers in the Bengal Delta. *The American Journal of Tropical Medicine and Hygiene*, *89*(5), 950–959. doi:10.4269/ajtmh.12-0492

Al-Arydah, M. T., Mwasa, A., Tchuenche, J. M., & Smith, R. J. (2013). Modelling cholera disease with education and chlorination. *Journal of Biological Systems*, *21*(04), 1340007. doi:10.1142/S021833901340007X

Ali, M., Gupta, S. S., Arora, N., Khasnobis, P., Venkatesh, S., Sur, D., ... Ganguly, N. K. (2017). Identification of burden hotspots and risk factors for cholera in India: An observational study. *PloS One*, *12*(8), e0183100. doi:10.1371/journal.pone.0183100

Ali, M., Lopez, A. L., You, Y., Kim, Y. E., Sah, B., Maskery, B., & Clemens, J. (2012). The global burden of cholera. *Bulletin of the World Health Organization*, *90*, 209–218. doi:10.2471/blt.11.093427

Ali, M., Nelson, A. R., Lopez, A. L., & Sack, D. A. (2015). Updated global burden of cholera in endemic countries. *PLOS Neglected Tropical Diseases, 9*(6), e0003832. doi:10.1371/journal.pntd.0003832

Anand, S., & Bärnighausen, T. (2004). Human resources and health outcomes: Cross-country econometric study. *The Lancet, 364*(9445), 1603–1609. doi:10.1016/S0140-6736(04)17313-3

Anbarci, N., Escaleras, M., & Register, C. A. (2012). From cholera outbreaks to pandemics: The role of poverty and inequality. *The American Economist, 57*(1), 21–31. doi:10.1177/056943451205700102

Asiedu, E., Gaekwad, N. B., Nanivazo, M., Nkusu, M., & Yi, J. (2015). On the impact of income per capital on health outcomes: Is Africa different? Retrieved at https://www.aeaweb.org/conference/2016/retrieve.php?pdfid=1019

Babones, S. J. (2008). Income inequality and population health: Correlation and causality. *Social Science & Medicine 66*, 1614–1626. doi:10.1016/j.socscimed.2007.12.012

Bain, R., Cronk, R., Wright, J., Yang, H., Slaymaker, T., & Bartram, J. (2014). Fecal contamination of drinking-water in low- and middle-income countries: A systematic review and meta-analysis. *PLoS medicine, 11*(5), e1001644. doi:10.1371/journal.pmed.1001644

Baker, D. W., Wolf, M. S., Feinglass, J., Thompson, J. A., Gazmararian, J. A., & Huang, J. (2007). Health literacy and mortality among elderly persons. *Archives of Internal Medicine, 167*(14), 1503–1509. doi:10.1001/archinte.167.14.1503

Beckfield, J. (2004). Does income inequality harm health? New cross-national evidence. *Journal of Health and Social Behavior, 45*(3), 231–248. doi:10.1177/002214650404500301

Berkman, L. F., & Kawachi, I. (2014). A historical framework for social epidemiology. In *Social epidemiology* (2nd ed.). New York, NY: Oxford University Press

Bhattacharya, S. K., Sur, D., Ali, M., Kanungo, S., You, Y. A., Manna, B, … Clemens, J. D. (2013). 5 year efficacy of a bivalent killed whole-cell oral cholera vaccine in Kolkata, India: A cluster-randomised, double-blind, placebo-controlled trial. *The Lancet Infectious Disease*, *13*(12), 1050-1056. doi: 10.1016/S1473-3099(13)70273-1

Bi, Q., Azman, A. S., Satter, S. M., Khan, A. I., Ahmed, D., Riaj, A. A., ... Lessler, J. (2016). Micro-scale spatial clustering of cholera risk factors in urban Bangladesh. *PLoS Neglected Tropical Diseases*, *10*(2), e0004400. doi:10.1371/journal.pntd.0004400

Biggs, B., King, L., Basu, S., & Stuckler, D. (2010). Is wealthier always healthier? The impact of national income level, inequality, and poverty on public health in Latin America. *Social Science & Medicine, 71*(2), 266–273. doi:10.1016/j.socscimed.2010.04.002

Birkmann, J. (2006). *Measuring vulnerability to natural hazards: Towards disaster resilient societies*. New York, NY: United Nations University Office

Bishwajit, G., Ide, S., & Ghosh, S. (2014). Social determinants of infectious diseases in South Asia. *International Scholarly Research Notices*, *2014*, 1–10.

doi:10.1155/2014/135243

Blaikie, P., Cannon, T., Davis, I., & Wisner, B. (2004). *At risk: Natural hazards, people's vulnerability and disasters*. New York, NY: Routledge.

Bompangue, D., Giraudoux, P., Handschumacher, P., Piarroux, M., Sudre, B., Ekwanzala, M., ... Piarroux, R. (2008). Lakes as source of cholera outbreaks, Democratic Republic of Congo. *Emerging Infectious Diseases, 14*(5), 798–800. doi:10.3201/eid1405.071260

Branisa, B. (2013). Gender inequality in social institutions and gendered development outcomes. *World Development*, *45*, 252–268. doi:10.1016/j.worlddev.2012.12.003

Brinda, E. M., Rajkumar, A. P., & Enemark, U. (2015). Association between gender inequality index and child mortality rates: A cross-national study of 138 countries. *BMC Public Health, 15*(1), 97. doi:10.1186/s12889-015-1449-3

Bwire, G., Munier, A., Ouedraogo, I., Heyerdahl, L., Komakech, H., Kagirita, A., … Mengel, M. A. (2017). Epidemiology of cholera outbreaks and socio-economic characteristics of the communities in the fishing villages of Uganda: 2011–2015. *PLoS Neglected Tropical Diseases, 11*(3), e0005407. doi:10.1371/journal.pntd.0005407

Carpenter, A. (2014) Behavior in the time of cholera: Evidence from the 2008–2009 cholera outbreak in Zimbabwe. In W. G. Kennedy, N. Agarwal, & S. J. Yang. (Eds.) *Social computing, behavioral-cultural modeling and prediction* (pp. 237–244). doi:10.1007/978-3-319-05579-4_29

Carpenter, J. R. (2009). Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ*, *338*, b2393. doi:10.1136/bmj.b2393

Catalyst, N. E. J. M. (2017). Social Determinants of Health (SDOH). NEJM Catalyst, 1. doi:10.1056/CAT.17.0312

Castillo-Laborde, C. (2011). Human resources for health and burden of disease: An econometric approach. *Human Resources for Health, 9*(1), 4. doi:10.1186/1478-4491-9-4

Cerda, R., & Lee, P. T. (2013). Modern cholera in the Americas: An opportunistic societal infection. *American Journal of Public Health, 103*(11), 1934–1937. doi:10.2105/AJPH.2013.301567

Chen, L., Evans, T., Anand, S., Boufford, J. I., Brown, H., Chowdhury, M., ... Fee, E. (2004). Human resources for health: overcoming the crisis. *The Lancet, 364*(9449), 1984–1990. doi:10.1016/S0140-6736(04)17482-5

Cohen, D. A., Farley, T. A., & Mason, K. (2003). Why is poverty unhealthy? Social and physical mediators. *Social Science & Medicine*, *57*(9), 1631–1641. doi:10.1016/S0277-9536(03)00015-7

Confalonieri, U. E. C., Wilson, M. E., & Najar, A. L. (2006). Social and environmental vulnerability to emerging infectious diseases. *Scripta Varia, 106*, 195–212. Retrieved from http://www.pas.va/content/accademia/en/publications/scriptavaria.html

Cordoba, E., & Aiello, A. E. (2016). Social determinants of influenza illness and outbreaks in the United States. *North Carolina Medical Journal*, *77*(5), 341–345. doi:10.18043/ncm.77.5.341

Cowman, G. A. (2015). Cholera prevention and control in Kenya (Doctoral dissertation, The University of North Carolina at Chapel Hill). Retrieved at http//: uuid:ec570afe-a337-42ef-9b9d-7e1cdebd2c6a

Cowman, G., Otipo, S., Njeru, I., Achia, T., Thirumurthy, H., Bartram, J., & Kioko, J. (2017). Factors associated with cholera in Kenya, 2008-2013. Pan African Medical Journal, 28(101). doi:10.11604/pamj.2017.28.101.12806

De Camargo, K. R., Ortega, F., & Coeli, C. M. (2013). Modern epidemiology and its discontents. Revsta Saúde Pública, 47(5), 984-91. doi:10.1590/S0034-8910.2013047004777

Diez-Roux, A. V. (1998). Bringing context back into epidemiology: variables and fallacies in multilevel analysis. American Journal of Public Health, 88(2), 216-222. doi:10.2105/AJPH.88.2.216

Eisenstein, M. (2016). Poverty and pathogens. Nature, 531, S61-S63. Retrieved from https://ezp.waldenulibrary.org/login?url=https://search-proquest-com.ezp.waldenulibrary.org/docview/1774749064?accountid=14872

Ellis, P. D. (2010). The essential guide to effect sizes: Statistical power, meta-analysis, and the interpretation of research results. Cambridge University Press. ISBN 978-0-521-14246-5

Elston, J. W. T., Moosa, A. J., Moses, F., Walker, G., Dotta, N., Waldman, R. J., & Wright, J. (2016). Impact of the Ebola outbreak on health systems and population health in Sierra Leone. Journal of Public Health, 38(4), 673-678. doi:10.1093/pubmed/fdv158

Evans, D. B., Tandon, A., Murray, C. J., & Lauer, J. A. (2001). Comparative efficiency of national health systems: Cross national econometric analysis. *BMJ, 323*(7308), 307-310. doi: 10.1136/bmj.323.7308.307

Field, A. (2013). Discovering statistics using IBM SPSS Statistics (4th ed.). Sage Publications Ltd. 55 City Road, London EC1Y 1SP, UK: Sage.

Filauri, E. (2010). State capacity influences on the epidemiology of neglected tropical and vector-borne diseases in Africa (Doctoral dissertation, Walden University). doi: 10.1016/j.inhe.2011.01.001

Frankfort-Nachmias, C., & Nachmias, D. (2008). Research methods in the social sciences (7th ed.). New York: Worth.

Frohlich, K. L., Corin, E., & Potvin, L. (2001). A theoretical proposal for the relationship between context and disease. *Sociology of Health and Illness*, 23(6), 776-797. doi:10.1111/1467-9566.00275

Frost, J. (2019). Regression Analysis. An Intuitive Guide for Using and Interpreting Linear Models. ebook.

George, C. M., Monira, S., Sack, D. A., Rashid, M. U., Saif-Ur-Rahman, K. M., Mahmud, T., ... & Leontsini, E. (2016). Randomized controlled trial of hospital-based hygiene and water treatment intervention (CHoBI7) to reduce cholera. Emerging infectious diseases, 22(2), 233. doi: 10.3201/eid2202.151175

Global Task Force on Cholera Control. (2017). Ending Cholera – A Global Roadmap to 2030. Retrieved from http://www.who.int/cholera/publications/global-roadmap.pdf?ua=1

Gostin, L. O., & Friedman, E. A. (2015). A retrospective and prospective analysis of the west African Ebola virus disease epidemic: robust national health systems at the foundation and an empowered WHO at the apex. *The Lancet, 385*(9980), 1902-1909. doi: 10.1016/S0140-6736(15)60644-4

Graham, J. W. (2009). Missing data analysis: Making it work in the real world. Annual *Review of Psychology, 60*, 549-576. doi: 10.1146/annurev.psych.58.110405.085530

Guerra-Silveira, F., & Abad-Franch, F. (2013). Sex bias in infectious disease epidemiology: patterns and processes. *PloS One, 8*(4), e62390. doi: 10.1371/journal.pone.0062390

Hosmer Jr, D. W., Lemeshow, S., & Sturdivant, R. X. (2013). Logistic regression models for the analysis of correlated data. *Hosmer Jr. DW, et al. Applied logistic regression. New Jersey, USA: Wiley*, 313-376.

IBM Corp. (2016). IBM SPSS Statistics for Windows, Version 24.0. Armonk, NY

IBM. (2017). SPSS statistics 25 guide reference: Missing values. Retrieved at https://www.ibm.com/support/pages/ibm-spss-statistics-25-documentation#en

Jaccard, J. (2001). Quantitative Applications in the Social Sciences: Interaction effects in logistic regression. NY SAGE Publications, 135. doi: 10.4135/9781412984515

Jeandron, A., Saidi, J. M., Kapama, A., Burhole, M., Birembano, F., Vandevelde, T., ... & Ensink, J. H. (2015). Water supply interruptions and suspected cholera incidence: a time-series regression in the Democratic Republic of the Congo. *PLoS*

Medicine, *12*(10), e1001893. doi: 10.1371/journal.pmed.1001893

Jones, B. A., Betson, M., & Pfeiffer, D. U. (2017). Eco-social processes influencing infectious disease emergence and spread. *Parasitology, 144*(1), 26-36. Retrieved at http://dx.doi.org.ezp.waldenulibrary.org/10.1017/S0031182016001414

Jutla, A., Whitcombe, E., Hasan, N., Haley, B., Akanda, A., Huq, A., Alam, M., Sack, R. B., and Colwell, R. (2013). Environmental Factors Influencing Epidemic Cholera. *American Journal of Tropical Medicine and Hygiene, 89*(3), 597–607. doi:10.4269/ajtmh.12-0721

Kabir, M. (2008). Determinants of life expectancy in developing countries. *Journal of Developing Areas*, *41*(2), 185 - 204. Retrieved at http://www.jstor.org/stable/40376184

Kawachi, I. (2014) in Berkman, L. F., & Kawachi, I. In M.M. Glymour, (Eds.). A historical framework for social epidemiology. Social Epidemiology, 1. Oxford University Press, 198 Madison Avenue, New York, NY, 10016, USA

Khatib, A. M, Ali, M., Seidlein, L., Kim, D. R., Hashim, R., Reyburn, R., Ley, B…, Deen, J. (2012). Effectiveness of an oral cholera vaccine in Zanzibar: findings from a mass vaccination campaign and observational cohort study. *Lancet Infectious Diseases, 12*, 837–844. doi:10.1016/S1473-3099(12)70196-2

Kim, T. K., & Lane, S. R. (2013). Government health expenditure and public health outcomes: a comparative study among 17 countries and implications for US health care reform. *American International Journal of Contemporary Research, 3*(9), 8-13. Retrieved at https://pdfs.semanticscholar.org/a585/271d7faa0ff66612a73f87a592e2ffb0713d.pdf

Kirigia, J. M., Sambo, L. G., Yokouide, A., Soumbey-Alley, E., Muthuri, L. K., and Kirigia, D. G. (2009). Economic burden of cholera in the WHO African region. BMC International Health and Human Rights, 9(8), doi:10.1186/1472-698X-9-8

Krieger, N. (1994). Epidemiology and the web of causation: Has anyone seen the spider?. *Social Science and Medicine, 39*(7), 887-903. doi:10.1016/0277-9536(94)90202-X

Krieger, N. (2001). Theories for social epidemiology in the 21st century: an ecosocial perspective. *International Journal of Epidemiology*, *30*(4), 668-677. doi: 10.1093/ije/30.4.668

Krieger, N. (2004). Ecosocial Theory. In Anderson, N.B. (Ed). Encyclopedia of Health and Behavior. ISBN: 9781412952576. doi: http://dx.doi.org/10.4135/9781412952576.n87

Krieger, N. (2011). Epidemiology and the people's health: theory and context. Oxford University Press.

Kruk, M. E., & Freedman, L. P. (2008). Assessing health system performance in developing countries: A review of the literature. *Health Policy, 85*(3), 263-276. doi: 10.1016/j.healthpol.2007.09.003

Kruk, M. E., Prescott, M. R., de Pinho, H., & Galea, S. (2009). Are doctors and nurses associated with coverage of essential health services in developing countries? A cross-sectional study. *Human Resources for Health*, *7*(1), 27. doi:10.1186/1478-

4491-7-27

Kwesiga, B., Pande, G., Ario, A. R., Tumwesigye, N. M., Matovu, J. K., & Zhu, B. P. (2018). A prolonged, community-wide cholera outbreak associated with drinking water contaminated by sewage in Kasese District, western Uganda. *BMC Public Health*, *18*(1), 30. doi:10.1186/s12889-017-4589-9

Leckebusch, G. C., & Abdussalam, A. F. (2015). Climate and socioeconomic influences on inter-annual variability of cholera in Nigeria. *Health and Place*, *34*, 107-117. doi: 10.1016/j.healthplace.2015.04.006

Lilje, J., Kessely, H., & Mosler, H. J. (2015). Factors Determining Water Treatment Behavior for the Prevention of Cholera in Chad. American Journal of Tropical *Medicine and Hygiene*, *93*(1), 57– 65. doi:10.4269/ajtmh.14-0613

Link, B. G., & Phelan, J. (1995). Social conditions as fundamental causes of disease. *Journal of Health and Social Behavior*, 80-94. doi: 10.2307/2626958

Little, R. J., & Rubin, D. B. (2014). Statistical analysis with missing data. John Wiley & Sons. doi:10.1002/9781119013563

Luby, S. P., Rahman, M., Arnold, B. F., Unicomb, L., Ashraf, S., Winch, P. J., ... & Leontsini, E. (2018). Effects of water quality, sanitation, handwashing, and nutritional interventions on diarrhoea and child growth in rural Bangladesh: a cluster randomised controlled trial. *The Lancet Global Health*, *6*(3), e302-e315. doi:10.1016/S2214-109X(17)30490-4

Luquero, F.J., Grout, L., Ciglenecki, I., Sakoba, K., Traore, B., Heile, M., Dialo, A.A., Itama, C., Serafini, M., Legros, D., Grais, R.F. (2013). First Outbreak Response

Using an Oral Cholera Vaccine in Africa: Vaccine Coverage, Acceptability and Surveillance of Adverse Events, Guinea, 2012. *Plos Neglected Tropical Diseases, 7*(10): e2465. doi:10.1371/journal.pntd.0002465

McCrickard, L. S., Massay, A. E., Narra, R., Mghamba, J., Mohamed, A. A., Kishimba, R. S., ... & Gibson, J. J. (2017). Cholera Mortality during Urban Epidemic, Dar es Salaam, Tanzania, August 16, 2015–January 16, 2016. *Emerging Infectious Diseases*, *23*(Suppl 1), S154. doi: 10.3201/eid2313.170529

Mengel M.A., Delrieu I., Heyerdahl L., & Gessner B.D. (2014) Cholera outbreaks in Africa. In G. Nair, Y. Takeda (Eds.). Cholera outbreaks. *Current Topics in Microbiology and Immunology, 379*. doi: 10.1007/82_2014_36

Midi, H., Sarkar, S. K., & Rana, S. (2010). Collinearity diagnostics of binary logistic regression model. *Journal of Interdisciplinary Mathematics, 13*(3), 253-267. doi:10.1080/09720502.2010.10700699

Mintz, E.D. & Tauxe, R.V. (2013). Cholera in Africa: A closer look and a time for action. *The Journal of Infectious Diseases 208*(S1), S4–S7. doi: 10.1093/infdis/jit205

Mitchell, A. D., Bossert, T. J., Yip, W., & Mollahaliloglu, S. (2008). Health worker densities and immunization coverage in Turkey: a panel data analysis. *Human Resources for Health, 6*(1), 29. doi:10.1186/1478-4491-6-29

Morgenstern, H. (1982). Uses of ecologic analysis in epidemiologic research. *American Journal of Public Health, 72*(12), 1336-1344. doi: 10.2105/AJPH.72.12.1336

Murray, C. J., & Chen, L. C. (1993). In search of a contemporary theory for understanding mortality change. *Social Science and Medicine, 36*(2), 143-155.

doi: 10.1016/0277-9536(93)90206-J

Murray, C. J., & Frenk, J. (2000). A framework for assessing the performance of health systems. *Bulletin of the World Health Organization, 78*(6), 717-731. Retrieved from https://www.scielosp.org/scielo.php?pid=S0042-96862000000600004&script=sci_arttext&tlng=en

Murray, C.J.L. & Evans, D.B. (2003). Health systems performance assessment: debates, methods and empiricism. WHO Library Cataloguing-in-Publication, 1211 Geneva 27, Switzerland. ISBN: 92 4 156245 5

Nguyen, M. P., Mirzoev, T., & Le, T. M. (2016). Contribution of health workforce to health outcomes: empirical evidence from Vietnam. *Human Resources for Health, 14*(1), 68. doi:10.1186/s12960-016-0165-0

Nguyen, V. D., Sreenivasan, N., Lam, E., Ayers, T., Kargbo, D., Dafae, F., ... & Stroika, S. (2014). Cholera epidemic associated with consumption of unsafe drinking water and street-vended water—Eastern Freetown, Sierra Leone, 2012. *The American Journal of Tropical Medicine and Hygiene, 90*(3), 518-523. doi:10.4269/ajtmh.13-0567

Njuguna, J. (2016). Effect of eliminating open defecation on diarrhoeal morbidity: an ecological study of Nyando and Nambale sub-counties, Kenya. *BMC Public Health, 16*(1), 712. doi:10.1186/s12889-016-3421-2

Nkoko, D. B., Giraudoux, P., Plisnier, P. D., Tinda, A. M., Piarroux, M., Sudre, B., ... & Piarroux, R. (2011). Dynamics of cholera outbreaks in Great Lakes region of Africa, 1978–2008. *Emerging Infectious Diseases, 17*(11), 2026. doi:

10.3201/eid1711.110170

Nsagha, D. S., Atashili, J., Fon, P. N., Tanue, E. A., Ayima, C. W., & Kibu, O. D. (2015). Assessing the risk factors of cholera epidemic in the Buea Health District of Cameroon. *BMC Public Health*, *15*(1), 1128. doi:10.1186/s12889-015-2485-8

Norton, E. C., Wang, H., & Ai, C. (2004). Computing interaction effects and standard errors in logit and probit models. *The Stata Journal*, *4*(2), 154-167. doi:10.1177/1536867X0400400206

Null, C., Stewart, C. P., Pickering, A. J., Dentz, H. N., Arnold, B. F., Arnold, C. D., ... & Hubbard, A. E. (2018). Effects of water quality, sanitation, handwashing, and nutritional interventions on diarrhoea and child growth in rural Kenya: a cluster-randomised controlled trial. *The Lancet Global Health*, *6*(3), e316-e329. doi:10.1016/S2214-109X(18)30005-6

O'Connell and A. Ann (2005). Logistic regression models for ordinal response variables. Quantitative applications in the social sciences, 146, Thousand Oaks, Sage Publications, CA, USA

Olago, D., Marshall, M., Wandiga, S. O., Opondo, M., Yanda, P. Z., Kangalawe, R., ... & Kirumira, E. (2007). Climatic, socio-economic, and health factors affecting human vulnerability to cholera in the Lake Victoria basin, East Africa. *AMBIO: A Journal of the Human Environment*, *36*(4), 350-358. Retrieved from http://www.jstor.org/stable/4315838

Paynter, S. (2016). Incorporating Transmission into Causal Models of Infectious Diseases for Improved Understanding of the Effect and Impact of Risk Factors. *American*

Journal of Epidemiology, 183(6), 574-582. doi: 10.1093/aje/kwv234

Phelps, M., & Simonsen, L. (2016). Epidemiological description of unmitigated cholera epidemics in 19th century Denmark. In ASTHM annual meeting. doi: 10.13140/RG.2.2.20353.66407

Phelps, M., Perner, M. L., Davidsen, E., Andreasen, V., Jensen, P. K. M., & Simonsen, L. (2016). Epidemiological description of unmitigated cholera epidemics in 19th century Denmark. Poster session presented at American Society of Tropical Medicine and Hygiene Annual Meeting, Atlanta, United States.

Phelps, M. D., Simonsen, L., & Jensen, P. K. (2019). Individual and household exposures associated with cholera transmission in case–control studies: a systematic review. *Tropical Medicine and International Health, 24*(10), 1151-1168. doi:10.1111/tmi.13293

Pickett, K.E. and Wilkinson, R. G. (2015). Income inequality and health: A causal review. *Social Science and Medicine 128*, 316 – 326. doi: 10.1016/j.socscimed.2014.12.031

Pinzon-Rondon, A. M., Attaran, A., Botero, J. C., & Ruiz-Sternberg, A. M. (2015). Association of rule of law and health outcomes: an ecological study. *BMJ Open, 5*(10), e007004. doi:10.1136/bmjopen-2014-007004

Piot, P., Coltart CEM, Atkins KE. (2017). Preface: ‘The 2013– 2016 West African Ebola epidemic: data, decision-making and disease control’. Philosophical Transactions Royal Society B, 372. doi: 10.1098/rstb.2017.0020

Pleasant, A. (2014). Advancing health literacy measurement: a pathway to better health and health system performance. *Journal of Health Communication, 19*(12), 1481-1496. doi:10.1080/10810730.2014.954083

Pop, I. A., Van Ingen, E., & Van Oorschot, W. (2013). Inequality, wealth and health: is decreasing income inequality the key to create healthier societies?. *Social Indicators Research, 113*(3), 1025-1043. doi: 10.1007/s11205-012-0125-6. ISBN 978–0–19–997672–0

Porta, M. (2008). A Dictionary of Epidemiology. Oxford University Press.

Sperandei, S. (2014). Understanding logistic regression analysis. *Biochemia Medica, 24*(1), 12–18. doi:10.11613/BM.2014.003

Poulos, C., Riewpaiboon, A., Stewart, J.F., Clemens, J., Guh, S., Agtini, M., Sur, D., Islam, Z., Lucas, M., Whittington, D. (2012). Costs of Illness Due to Endemic Cholera. *Epidemiology and Infection,140*(3), 1 – 16, doi:10.1017/S0950268811000513.

Pullan, R. L., Freeman, M. C., Gething, P. W., & Brooker, S. J. (2014). Geographical inequalities in use of improved drinking water supply and sanitation across sub-Saharan Africa: mapping and spatial analysis of cross-sectional survey data. *PLoS Medicine, 11*(4), e1001626. doi: 10.1371/journal.pmed.1001626

Qadri, F., Ali, M., Chowdhury, F., Khan, A. I., Saha, A., Khan, I. A., ... & Khan, J. A. (2015). Feasibility and effectiveness of oral cholera vaccine in an urban endemic setting in Bangladesh: a cluster randomised open-label trial. *The Lancet*, *386*(10001), 1362-1371. doi: 10.1016/S0140-6736(15)61140-0

Rajan, K., Kennedy, J., and King, L. (2013). Is wealthier always healthier in poor countries? The health implications of income, inequality, poverty, and literacy in India. *Social Science and Medicine 88*, 98 -107. doi: 10.1016/j.socscimed.2013.04.004

Ramesh, A., Blanchet, K., Ensink, J. H., & Roberts, B. (2015). Evidence on the effectiveness of water, sanitation, and hygiene (WASH) interventions on health outcomes in humanitarian crises: a systematic review. *PLoS One, 10*(9), e0124688. doi: 10.1371/journal.pone.0124688

Rancourt, N. (2013). Gender and Vulnerability to Cholera in Sierra Leone: Gender analysis of the 2012 cholera outbreak and an assessment of Oxfam's response. Oxfam GB. Oxfam House, Oxford, OX4 2JY, UK. ISBN: 978-1-78077-349-0

Rebaudet, S., Sudre, B., Faucher, B., & Piarroux, R. (2013). Environmental determinants of cholera outbreaks in inland Africa: a systematic review of main transmission foci and propagation routes. *The Journal of Infectious Diseases, 208*(suppl_1), S46-S54. doi: 10.1093/infdis/jit195

Reidl, J., & Klose, K. E. (2002). Vibrio cholerae and cholera: out of the water and into the host. *FEMS Microbiology Reviews, 26*(2), 125-139. doi: 10.1111/j.1574-6976.2002.tb00605.x

Root, E. D., Rodd, J., Yunus, M., & Emch, M. (2013). The role of socioeconomic status in longitudinal trends of cholera in Matlab, Bangladesh, 1993–2007. *PLoS Neglected Tropical Diseases, 7*(1), e1997. doi: 10.1371/journal.pntd.0001997

Rosewell, A., Bieb, S., Clark, G., Miller, G., MacIntyre, R., & Zwi, A. (2013). Human resources for health: lessons from the cholera outbreak in Papua New Guinea. *Western Pacific Surveillance and Response Journal, 4*(3), 9. doi: 10.5365/wpsar.2013.4.2.006

Rothman, K.J., Adami, H.O., & Trichopoulos, D. (1998). Should the mission of epidemiology include the eradication of poverty? *The Lancet, 352*(9130), 810-813. doi:10.1016/S0140-6736(98)01327-0

Saha, A., Rosewell, A., Hayen, A., MacIntyre, C. R., & Qadri, F. (2017). Improving immunization approaches to cholera. *Expert Review of Vaccines, 16*(3), 235-248. doi:10.1080/14760584.2017.1249470

Sambo, L. G., Kirigia, J. M., & Orem, J. N. (2013). Health financing in the African Region: 2000–2009 data analysis. *International Archives of Medicine, 6*(1), 10. doi: 10.1186/1755-7682-6-10

Sauvageot, D., Njanpop-Lafourcade, B. M., Akilimali, L., Anne, J. C., Bidjada, P., Bompangue, D., ... & Orach, C. G. (2016). Cholera incidence and mortality in Sub-Saharan African sites during multi-country surveillance. *PLoS Neglected Tropical Diseases, 10*(5), e0004679. doi: 10.1371/journal.pntd.0004679

Schaetti, C., Sundaram, N., Merten, S., Ali, S.M., Nyambedha, E.O., Lapika, B., Chaignat, C., Hutubessy, R. and Weiss, M.G. (2013). Comparing sociocultural features of cholera in three endemic African settings. *BMC Medicine, 11*(1), 206. doi: 10.1186/1741-7015-11-206

Sen, G., & Östlin, P. (2008). Gender inequity in health: why it exists and how we can change it: Final Report to the WHO Commission on Social Determinants of Health. Retrieved from who.int/social_determinants/resources/csdh_media/wgekn_final_report_07.pdf

Snow, J. (1856). On the mode of communication of cholera.” Edinburgh Medical Journal, 1(7), 668–670. Retrieved from https://johnsnow.matrix.msu.edu/work.php?id=15-78-4B

Sperandei, S. (2014). Understanding logistic regression analysis. Biochemia medica: *Biochemia Medica, 24*(1), 12-18. doi:10.11613/BM.2014.003

Song, M. K., Lin, F. C., Ward, S. E., & Fine, J. P. (2013). Composite variables: when and how. *Nursing Research, 62*(1), 45–49. doi:10.1097/NNR.0b013e3182741948

Stanturf, J. A., Goodrick, S. L., Warren Jr, M. L., Charnley, S., & Stegall, C. M. (2015). Social vulnerability and Ebola virus disease in rural Liberia. *PLoS One, 10*(9), e0137208. doi: 10.1371/journal.pone.0137208

Sterne, J. A., White, I. R., Carlin, J. B., Spratt, M., Royston, P., Kenward, M. G., ... & Carpenter, J. R. (2009). Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ, 338*, b2393. doi: 10.1136/bmj.b2393

Stoltzfus, J.D., Carter, J.Y., Akpinar-Elci, M., Matu, M., Kimotho, V.,…Langat, D., & Elci, O. C. (2014). Interaction between climatic, environmental, and demographic factors on cholera outbreaks in Kenya. *Infectious Diseases of Poverty, 3*(37), doi:10.1186/2049-9957-3-37

Sugimoto, J. D., Koepke, A. A., Kenah, E. E., Halloran, M. E., Chowdhury, F., Khan, A.

I., ... & Calderwood, S. B. (2014). Household transmission of Vibrio cholerae in Bangladesh. *PLoS Neglected Tropical Diseases, 8*(11), e3314. doi:10.1371/journal.pntd.0003314

Sun, G. Q., Xie, J. H., Huang, S. H., Jin, Z., Li, M. T., & Liu, L. (2017). Transmission dynamics of cholera: Mathematical modeling and control strategies. Communications in Nonlinear Science and Numerical Simulation, 45, 235-244. doi: 10.1016/j.cnsns.2016.10.007

Talavera, A. & Pérez, E.M. (2009). Is cholera disease associated with poverty? *The Journal of Infection in Developing Countries, 3*(6), 408-411. doi: 10.3855/jidc.410

Tambo, E., Ugwu, E. C., & Ngogang, J. Y. (2014). Need of surveillance response systems to combat Ebola outbreaks and other emerging infectious diseases in African countries. *Infectious Diseases of Poverty, 3*(1), 29. doi: 10.1186/2049-9957-3-29

Taylor, D. L., Kahawita, T. M., Cairncross, S., & Ensink, J. H. (2015). The impact of water, sanitation and hygiene interventions to control cholera: a systematic review. *PLoS One, 10*(8), e0135676. doi: 10.1371/journal.pone.0135676

United Nations Development Program – UNDP (2017). Gender Inequality Index. Human Development Reports. Retrieved from. http://hdr.undp.org/en/content/gender-inequality-index-gii

United Nations International Strategy for Disaster Risk Reduction. (2009). 2009 UNISDR Terminology on Disaster Risk Reduction. UNISDR-20-2009-Geneva,

Geneva, Switzerland

United Nations Entity for Gender Equality and the Empowerment of Women. (2017). Gender equality glossary. Retrieved at https://trainingcentre.unwomen.org/mod/glossary/view.php

Van Deurzen, I., Van Oorschot, W., & Van Ingen, E. (2014). The Link between Inequality and population health in low and middle income countries: Policy Myth or Social Reality? *PLoS One 9*(12): e115109. doi: 10.1371/journal.pone.0115109

Vandenbroucke, J.P. (1998). Is ‘the causes of cancer’ a miasma theory for the end of the twentieth century? International Journal of Epidemiology. 17(4), 708-709. doi:10.1093/ije/17.4.708

Waldman, R. J., Mintz, E. D., & Papowitz, H. E. (2013). The cure for cholera—improving access to safe water and sanitation. *New England Journal of Medicine, 368*(7), 592-594. doi: 10.1056/NEJMp1214179

Wang, R., Sedransk, J., & Jinn, J. H. (1992). Secondary data analysis when there are missing observations. *Journal of the American Statistical Association, 87*(420), 952-961. Retrieved from www.jstor.org/stable/2290631

Wolfe, M., Kaur, M., Yates, T., Woodin, M., & Lantagne, D. (2018). A Systematic Review and Meta-Analysis of the Association between Water, Sanitation, and Hygiene Exposures and Cholera in Case–Control Studies. *The American journal of Tropical Medicine and Hygiene, 99*(2), 534-545. doi:10.4269/ajtmh.17-0897

World Health Organization. (1986). Ottawa Charter for Health Promotion, 1986.

Retrieved at http://www.euro.who.int/__data/assets/pdf_file/0004/129532/Ottawa_Charter.pdf?ua=1,

World Health Organization. (2000). World health report 2000. Health systems: Improving performance. Geneva, Switzerland; 2000. World Health Organization, 1211 Geneva 27, Switzerland. ISBN 92 4 156198 X

World Health Organization. (2007). Everybody's business: Strengthening health systems to improve health outcomes. WHO's framework for action. Geneva, World Health Organization Document Production Services. ISBN 978 92 4 159607 7.

World Health Organization. (2010a). Cholera vaccines: WHO position paper. Weekly epidemiological record, 13(85), 17–128

World Health Organization. (2010b). Monitoring the building blocks of health systems: a handbook of indicators and their measurement strategies. WHO Document Production Services, Geneva, Switzerland. ISBN 978 92 4 156405 2.

World Health Organization. (2015). World Health Statistics 2015: Indicator compendium, Indicator Code Book World Health Statistics - World Health Statistics indicators retrieved from http://www.who.int/gho/publications/world_health_statistics/WHS2015_IndicatorCompendium.pdf

World Health Organization. (2016a). Cholera, 2015. Weekly epidemiological record. 91(38), 433-440. Retrieved from http://www.who.int/wer/2015/wer9040/en/

World Health Organization. (2016b). Global strategy on human resources for health:

Workforce 2030. World Health Organization. Geneva, Switzerland. ISBN 978 92 4 151113 1

World Health Organization. (2017a). World health statistics 2017: monitoring health for the SDGs, Sustainable Development Goals. ISBN 978-92-4-156548-6

World Health Organization. (2017b). 2016 Report on acute public health events assessed by WHO Regional Offices for Africa, the Americas, and Europe under the IHR (2005). Accessed at https://www.paho.org/hq/dmdocuments/2016/2016-Joint-Report-Afro-Amro-Euro.pdf

World Health Organization. (2018a). Cholera vaccine: WHO position paper, August 2017–Recommendations. Vaccine, 36(24), 3418-3420. doi: 10.1016/j.vaccine.2017.09.034.

World Health Organization. (2018b). Managing epidemics: Key facts about major deadly diseases. Geneva, Switzerland. ISBN: 978-92-4-156553-0

World Health Organization and United Nations Children's Fund. (2017). Progress on drinking water, sanitation and hygiene: 2017 update and SDG baselines. Geneva, Licence: CC BY-NC-SA 3.0 IGO.ISBN 978-92-4-151289-3

World Bank. (2012). Income Inequality by the Numbers: in History and Now. Policy Research Working Paper 6259. Accessed Retrieved from https://openknowledge.worldbank.org/bitstream/handle/10986/12117/wps6259.pdf?sequence=2, on June 28, 2017

World Bank. (2014). World development indicators 2014. World Bank Publications.

World Bank (2017). Country Income Classification. Accessed from https://datahelpdesk.worldbank.org/knowledgebase/articles/906519-world-bank-country-and-lending-groups

World Bank (2018) Metadata Glossary (2018). Retrieved from https://databank.worldbank.org/metadataglossary/gender-statistics/series/SI.POV.GINI