Trends of Cholera Epidemics in West Africa: 2012-2017

Amadou Conde

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of Master of Science in Research Methodology

Prince of Songkla University

2022

Trends of Cholera Epidemics in West Africa: 2012-2017

Amadou Conde

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of Master of Science in Research Methodology

Prince of Songkla University

2022

Thesis Title	Trends of Cholera Epidemics in West Africa: 2012-2017
Author	Mr. Amadou Conde
Major Program	Research Methodology

Major Advisor	**Examining Committee**:
....	....Chairperson
(Dr. Nurin Dureh)	(Assoc. Prof. Dr. Apiradee Lim)
Co-advisor	....Committee
	(Dr. Nurin Dureh)
....	....Committee
(Asst. Prof. Dr. Attachai Ueranatasun)	(Asst. Prof. Dr. Attachai Ueranatasun)
	....Committee
	(Asst. Prof. Dr. Wandee Wanishsakpong)

The Graduate School, Prince of Songkla University, has approved this thesis as partial fulfillment of the requirements for the Master of Science Degree in Research Methodology

....
(Asst. Prof. Dr. Thakerng Wongsirichot)
Acting Dean of Graduate School

This is to certify that the work here submitted is the result of the candidate's own investigations. Due acknowledgement has been made of any assistance received.

....

(Dr. Nurin Dureh)

Major Advisor

....

(Asst. Prof. Dr. Attachai Ueranatasun)

Co-advisor

....

(Mr. Amadou Conde)

Candidate

I hereby certify that this work has not been accepted in substance for any degree and is not being currently submitted in candidature for any degree.

....

(Mr. Amadou Conde)

Candidate

Thesis Title Trend of Cholera Epidemics in West Africa: 2012-2017

Author Mr. Amadou Conde

Major Program Research Methodology

Academic Year 2022

ABSTRACT

The objectives of this research study were to find out the cases of cholera in West Africa regions from year 2012-2017 for the period of six years and to pinpoint the factors associated with the trends of cholera occurrence and intensity. The reported data of the study were gained from Humanitarian Data Exchange (HDX) in collaboration with World Health Organization (WHO). The model described replaces incidence by a bivariate outcome-comprising occurrence (taking values 0 if no case occurs or 1 otherwise), and intensity (incidence rate for cases that occur). Logistic regression was used to model occurrence, and log-linear regression was used to model intensity. The overall cases of cholera were totalized 326 778 cases in those 17 countries. The average proportion of cholera incidence rate per 100,000 population in many countries was below 1% and 0.1%. The highest average proportion was Sierra Leone with 22 729 cases for the population of 6.47 million, whereas Mali and Burkina Faso showed the lowest average proportion with 242 and 143 incidences for the population of 18.99 million and 20.29 million respectively. The DR. Congo was the highest cases rate about 178.129, second to Nigeria 52756 cases appeared and Burkina Faso only 143 cases. In addition, cholera rates were higher in September and October, but decreased in 2015-2017, according to the model. When incidence and intensity are combined, two northern countries (Mali and Niger) have both high occurrence and high intensity, but three coastal countries (Guinea-Bissau, Cote d'Ivoire, and Togo) have

both low occurrence and intensity. The incidence and intensity of four countries are considerably varied.

ACKNOWLEDGEMENTS

First, I would like to thanks my advisor and co-advisor Dr. Nurin Dureh and Asst. Prof. Dr. Attachai Ueranatasun respectively for their input, good advice and help throughout the accomplishment of this thesis. Also many thanks to Emeritus Prof. Dr. Don Mcneil for the guidance and help. I am grateful to TEH-AC scholarship for financing my master's degree.

Last, not the end, I would like to thank all my lecturers in department of Mathematics and computer science from the day one of this master's degree. My special thanks goes to Assoc. Prof. Dr. Apiradee Lim for her support.

Finally, I would like to thank my parents and siblings for their support and motivation.

Amadou Conde

CONTENTS

CONTENTS (Cont.)

Page

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

1.1 Background and rational

Cholera is a vital public health problem worldwide. Even though many of cholera contamination are difficult to detect, huge cholera outbreaks have been observed in Haiti, Vietnam, Congo Republic and Zimbabwe in recent years (WHO, 2016). Cholera cases have not been practical found in developed countries for over a century, because of their decent water and sewage treatment infrastructures. However, the causative agents (Vibrio cholerae O1 and O139) continue to boom wherever crowded housing conditions exist, and water and sanitation facilities are suboptimal (Gaffga, 2007).

Nowadays cholera is endemic in 47 countries, particularly in areas where the water, sanitation and hygiene (WASH) infrastructure is poor, and the most vulnerable groups in those particular areas are local. Epidemics occur both within and outside of endemic areas, often amid humanitarian crises, when the WASH infrastructure breaks down or is overwhelmed. In a situation where the population lacks immunity, a wider age range is affect, often with more severe clinical manifestations (Legros, 2018). Cholera is a bacterial disease of people caused by Vibrio cholerae (of classical or el tor biotypes) which characteristically causes serious loose bowels, and death (in those severely influenced) from water and electrolytes consumption.

Cholera has been calling a “Blue Death” due to a patient’s skin turning a bluish-grey color from extraordinary misfortune of fluids (Payton et al., 2010). Enteric disease-causing infections stay a significant problem throughout the world. An approximately 2-4 billion episodes of infectious diarrhea happen annually in developing nations, leading in 3-5 million fatalities, with the largest incidence and case-fatality rates among kids under the age of five (Rasti and Brown, 2019). Nearly half of all diarrhea instances are due to disease-causing bacteria generating one or more enterotoxins. Among these, the most serious disease is caused by V. cholera. Cholera is thought to be at least as prevalent now as it was 50 years ago, with approximately 100, 000 – 300,000 cases reported annually to WHO in 1995- 2004 (Yang, 2019). West Africa had been free of cholera during the entire twentieth century until August 1970 when a devastating epidemic of more than 150,000 cases and 20,000 deaths occurred. Cholera is still endemic in West Africa (WHO, 2017).

In a medical report from India, recorded proof of cholera epidemics dates back to 1563 (Okoh, 2018). In the nineteenth century, cholera spread from its apparent ancestral site in the orient to other parts of the world, producing pandemics in Europe (Rahman, 2018). The first pandemic was reported in 1817, showing a spread of the disease outside the Indian subcontinent along trade paths to south Russia's west. A second pandemic began in 1826, reaching the biggest European cities by the early 1830s (Mourya et al., 2019). The pandemic reached the United Kingdom in 1831 and the reaction was crucial in that it led to the establishment of local health boards and a "Cholera Gazette," which served as a clearinghouse to monitor the outbreak (Grove and Adamson, 2018). At that period cholera was believed to be spread by the “miasma” (like a fog) coming from the river, but the classic epidemiological study of John Snow in 1854 in London showed that the disease was linked with contaminated drinking water even before any bacteria were known to exist (Cottrel, 2019). Many past studies had focused on the incidence of cholera in Haiti, Africa countries, especially in DR. Congo.

However, there are limited studies that has been considered to study cholera outbreak in West Africa countries as a region. Hence, the present study is concerned about the trends, occurrence and intensity of cholera outbreak in West Africa from 2012 to 2017.

The results of the study would show the high and low cholera in terms of occurrence, intensity and seasonal pattern. The significant of the study will provide useful information in better understanding the trends, occurrence, and intensity on how to prevent the citizen from this deadly disease contamination. In addition, safe oral cholera vaccines should be used in conjunction with improvements in water and sanitation.

1.2 Objectives

The objectives of this study are as follows.

1. To investigate the trends of cholera occurrence and intensity in West Africa.
2. To compare the disparities of cholera occurrence and intensity in West Africa.

1.3 Expected advantages

1. The description of the occurrence and intensity of cholera is performed

2. Statistical model for analyzing the disparities of occurrence and intensity is obtain

3. The results of this study will assist in better understanding of the danger of cholera in most affected countries in West Africa, this useful information will allow the decision makers to create public awareness and be more proactive

1.4 Literature reviews

1.4.1 Occurrence of Cholera

Africa accounted for the majority of the global disease burden during the existing seventh cholera pandemic. More than 40 years after its resurgence in Africa in 1970, cholera remains a serious public health issue, characterized by a high disease burden, frequent outbreaks, persistent endemicity, and high CFRs, particularly in the region of the central African Great Lakes, which may serve as cholera reservoirs. Cases occur all year, with an increase in frequency during the rainy season. Elsewhere in Sub-Saharan Africa, cholera is mostly found in small outbreaks with the constant threat of widespread epidemics (Shimada, 2008). African countries reported 3,221,050 suspected cholera cases to the World Health Organization between 1970 and 2011, accounting for 46% of all cases reported globally. Sub-Saharan Africa accounted for 86% of reported cases and 99% of deaths worldwide in 2011, excluding the Haitian epidemic. Because of differences in modalities, completeness, and case definition in national cholera data, the number of cholera cases may be much higher than what is reported to the WHO. Adjusting for underreporting, one source estimates 1,341,080 cases and 160,930 deaths in Africa (52.6% of 2,548,227 estimated cases and 79.6% of 209,216 estimated deaths worldwide). Another estimate is 1,411,453 cases and 53,632 deaths per year (50 percent of the estimated 2,836,669 cases and 58.6% of the estimated 91,490 deaths worldwide) (Balakrish, 2006). Until 2017, African countries reported more than 4 million cholera cases to the World Health Organization. Explosive outbreaks occurred in the Democratic Republic of the Congo (DRC), Ethiopia, Nigeria, Somalia, South Sudan, Sudan, and Zambia in 2017. In 2017, the total number of cholera cases and deaths reported to the World Health Organization from Africa was 179,835 and 3,220, respectively, with reported case fatality rates ranging from zero in many countries to 3.2% in Zambia, 5.2% in Angola, and 6.8% in 5 Chad. These high case fatality rates reflect significant barriers to adequate case management in large parts of Africa (Deen et al., 2019).

1.4.2 Intensity of Cholera

Cholera is an acute illness with severe hypoperfusion complications that can include stroke, acute tubular necrosis with renal dysfunction, and aspiration pneumonia from the extensive vomiting (Mandell et al., 2016). Despite advances in disease knowledge and treatment, cholera remains a major public health problem in several countries. Unless the disease is diagnosed and treated by health professionals as soon as possible, the case fatality rate in its epidemic form can be as high as 30%. (Mahmud et al., 2014).

The World Health Organization (WHO) reported that cholera is an acute diarrheal disease that can kill within hours if left untreated. Researchers have estimated that each year there are 1.3 million to 4.0 million cases of cholera, and 21,000 to 143,000 deaths worldwide due to cholera (Ali et al., 2015).

1.4.3 Trends of Cholera

Several countries were still dealing with cholera outbreaks. In total, 54% of cases were reported in Africa, 13% in Asia, and 32% in Hispaniola. Cholera remains a major health issue, primarily affecting developing-world populations who lack access to adequate water and sanitation resources (WHO, 2017). Between 1970 and 2011, half of all cases in Africa were reported from just seven countries: Angola, the Democratic Republic of the Congo, Mozambique, Nigeria, Somalia, Tanzania, and South Africa. In conversely to the global trend of decreasing case fatality ratios (CFRs), CFRs in Africa have remained steady at about 2%. (Balakrish et al., 2006). Cholera has been recorded in many nations around the world, despite attempts being made by many nations to prevent the spread of the illness (WHO, 2017a). Officially reported cases, however, do not reflect the overall burden of the disease due to under-reporting of cases due to political and economic reasons, fear of loss of tourism and trade, and poorly functioning surveillance systems (WHO, 2017b). In 1992 estimated 235,000 clinical instances of cholera in Bangladesh in 1991, but none were recorded formally that year (WHO, 1992). The WHO has approximately 5-10 percent of the real number of instances recorded globally (WHO, 2017).

CHAPTER 2

METHODOLOGY

The chapter of research methodology expresses the methods considered in the study that includes the following components.

(1) Data source and study area

(2) Data collection and management

(3) Path diagram and variables

(4) Statistical methods

2.1 Data source and study area

Cholera epidemics data for West Africa were derived from Humanitarian Data Exchange (HDX) in collaboration with World Health Organization (WHO) from the period of 2012-2017. These data provide information on country code, country name, and affected cases date by week and year, the 17 countries from the data are geographically located in West Africa namely: Nigeria, Guinea, Guinea Bissau, Burkina Faso, Mali, Cote D'Ivoire, Chad, Niger, Togo, Ghana, Benin, Central Africa, Cameroon, Congo, Democratic Republic Congo, Sierra Leone and Liberia as shown in Figure 1. The data is available from the Humanitarian Data Exchange (HDX) website at https://centre.humdata.org, supported by the US Census Bureau .Populations and projections for all countries in the world together with other useful population data are freely available from the US Census Bureau website at https://www.census.gov.

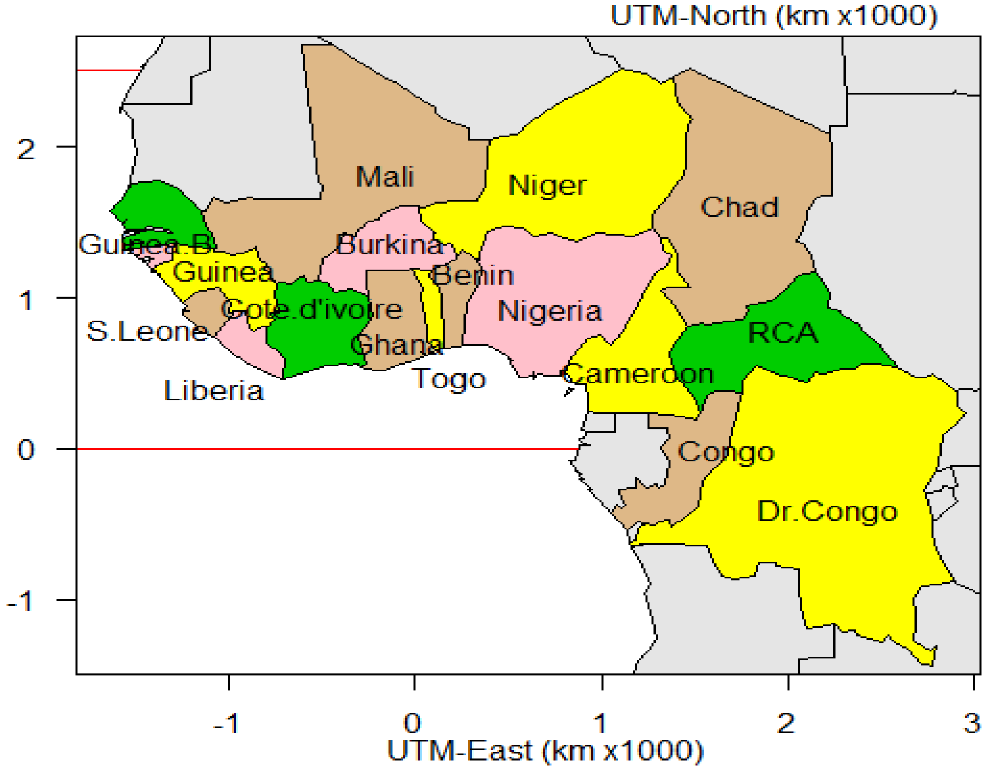


Figure 0.1 Geographical location of the study area

2.2 Data management

Cholera epidemics data for West Africa from Humanitarian Data Exchange (HDX) and US Census Bureau were recorded as a text file. Checking of errors was performed, to find out wrong codes, and missing values. Before analyzing the data, errors cleaning was performed, only few errors were found and omitted. For further analysis normalize data from the table into just four fields was done, note that 2015 has 53 weeks, so it was removed. Senegal data for only three years was presented, it was excluded. Thus, all of the missing values and errors were omitted in order to proceed further analysis. All of these data management was done by using R program.

2.3 Path diagram and variables

2.3.1 Path diagram

Path diagram consists the determinants and outcomes. The variables of this study are determined as determinants for countries, years and weeks while the outcomes variables are determined as occurrence and intensity as it is shown on the below path diagram.

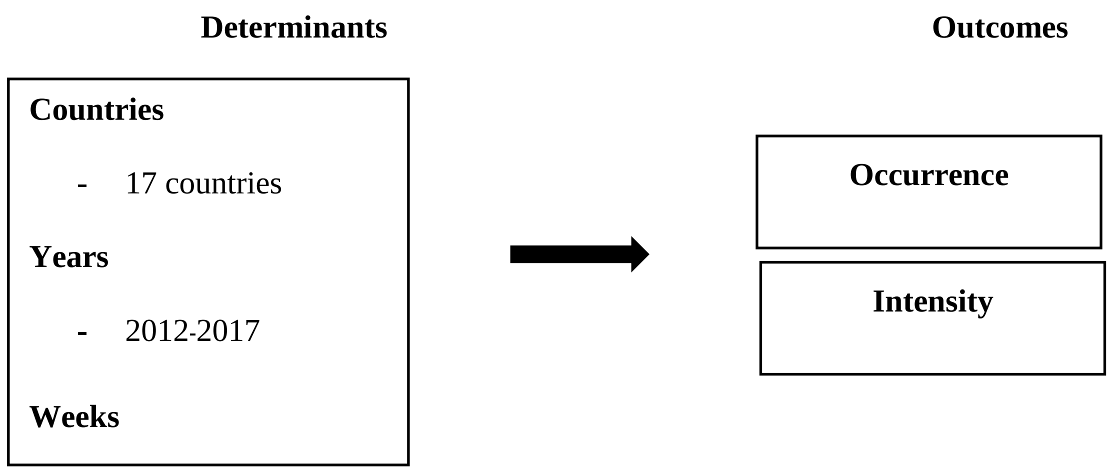


Figure 0.2 Path diagram

2.3.2 Variable Types

The data used for the study consist of year and week, (17 country), year (2012-2017) and occurrence and intensity rate. The variables roles and data types are shown in Table 2.1.

Table 0.1 Variable types and roles

Variables	Types (Number of categories)	Roles
Occurrence	Continuous	Outcome
Intensity	Continuous	Outcome
Country	Ordinal (17)	Determinant
Year	Ordinal (6)	Determinant
Week	Ordinal (52)	Determinant

2.4 Statistical methods

2.4.1 Cube root transformation of incidence rate

Incidence rates was computed as the number of cholera cases divided by the number of population and multiply by 100,000. Then it was transformed using the cube root to satisfy the distributional and variance homogeneity requirements. The cube root transformation was appropriate for this dataset since it avoided over-transforms and handled zero counts. We did not attempt to impute such outliers despite apparent anomalies in the database.

2.4.2 Poisson model

Since occurrence and intensity rates were considered as count data being the number of people who affected from cholera, Poisson regression model is suitable for fitting models with count data, which are non-negative and integer values. The likelihood function for the Poisson distribution with observed counts of y is given by,

$$Pr(Y = y) = \frac{e^{-\lambda}\lambda^{y}}{y!}$$

Where λ is the Poisson parameter, which equals to both the mean and the variance, e is the base of the natural logarithm and the number of occurrences of the event is y, which the probability giving by the Poisson function. Since λ is the adjust parameter in the Poisson model for the variation in observed count that tracks the Poisson distribution. The Poisson distribution mean must be greater than zero. It would be unfitting to assume that,

$$\lambda = \mathrm{a} + b_i x_i + \cdots + b_t x_t$$

The restriction of Poisson distribution is the mean must be equal to variance. Poisson regression model can be fitted to use generalized linear models (GLMs) with the log link function. Suppose that λ_{ijt} is a number of affected cholera cases in West Africa$i$, week j, and year t. Then the Poisson regression model is given by

$$\ln(\lambda_{ijt}) = \ln(P_{ijt}) + \mu + \alpha_i + \beta_j + \gamma_t$$

The parameter λ is the mean of λ_{ijt}, P_{ijt} is the population in West Africa i, countries by week j and year t, α is the effect of West Africa, β is the effect of week and γ is the effect of year. We assume that the effect of variables α_1, β_1 and γ_1 equal to zero. Poisson model assumption is frequently violated because of the problem of over dispersion. This means that the variance is greater than mean. Therefore, substitute model, which is negative binomial, was then considered.

2.4.3 Negative binomial regression

The negative binomial is another regression model for count data when Poisson regression does not fit to the data. The distribution of observed counts y takes the form:

$$Prob(Y = y) = \frac{\Gamma(y + k)}{\Gamma(y + 1)\Gamma(k)} \left(\frac{k}{k + \lambda}\right)^k \left(\frac{\lambda}{k + \lambda}\right)^y$$

In this equation Γ is the gamma function and k is known as the dispersion parameter, (is greater 0). Unlike the Poisson distribution when the mean must be equal to the variance, the negative binomial is $\lambda + \lambda^2 / k$. The negative binomial is equivalence to the Poisson distribution model if k the dispersion parameter is equal to zero. Thus if k is equal to zero, Poisson model is appropriate whereas the negative binomial is appropriate if k is significantly different from zero.

2.4.4 Logistic regression model

A statistical model used for analysis of epidemiological data is the logistic regression model that requires the response variables, which are dichotomous. It is suitable for binary outcome data and can manage general exposure variables not just only dichotomous. In this research study, the outcome is the occurrence, thus it is rational to use the logistic regression model. This model takes the form (McNeil, 1996)

$$ln\left(\frac{p}{1-p}\right) = a + b_1x_1 + b_2x_2 + \cdots + b_jx_j$$

Where p denotes the probability of occurrence of the outcome and x_j represent the j determinant. This equation may be inverted to give an expression for the probability p as

$$P = \frac{1}{1 + exp(-a - \sum_{j=1}^{p} b_jx_j)}$$

The functional form of the right-hand side ensures that its values are always between 0 and 1, which rational given that they are probabilities. Thus, the odds ratio for comparing two levels of the determinant x_j is given by

$$OR = exp\left(\sum_{j=1}^{p} b_j(x_j^{(1)} - x_j^{(0)})\right)$$

The parameter *b* in the model may be interpreted directly as natural logarithm of the odds ratio.

2.4.5 Log- linear regression models

Log-linear regression model have two important advantages, the first advantage it is flexible and the second advantage it is interpretable. It is also associated with analysis of variance and regression, which contains all the modeling flexibility.

$$\text{Log}(Y) = \alpha + \beta X$$

The interpretation of log-linear equation is one-unit change in X leads to 100*β percent change in Y. These models were used to describe occurrence (taking values 0 if no case occurs or 1 otherwise) and intensity (incidence rate for cases that occur). Where logistic regression was used to model occurrence, and simple linear regression after log-transforming ("log-linear" regression) was used to model intensity, the model's assessment was done with the normal quantile-quantile (Q-Q) plots.

2.5 Ethics Approval

Ethical approval for the study was obtained from Ethical Review Committee for Human Research, Prince of Songkla University, Pattani Campus, No. PSU.PN.1-006/63, March 31, 2020.

CHAPTER 3

RESULTS

This chapter presents the finding on the trends in weekly cholera incidence in West Africa from 2012-2017. The disparities of the cholera occurrence and intensity also illustrated through the models.

3.1 Trends in weekly cholera incidence

Figure 1 shows the time series plot of weekly incidence rate for 17 West African countries. These plots show rates of cholera cases per 100,000 population in successive weeks for years 2012-2017. The trend is approximately constant for the Democratic Republic of Congo, but rates vary a lot in other countries.

3.2 Seasonal patterns of weekly cholera incidence

Seasonal patterns of cholera incidence rates are depicted in Figure 3.2. The overall incidence rate is represented by the red line. Each county's actual incidence rate is represented by the black line. It shows that cholera rates rise after July, when monsoon rains come, with larger peaks in Sierra Leone, Guinea Bissau, Guinea, and Ghana. There are three exceptions: Ivory Coast, Congo-West, and Liberia.

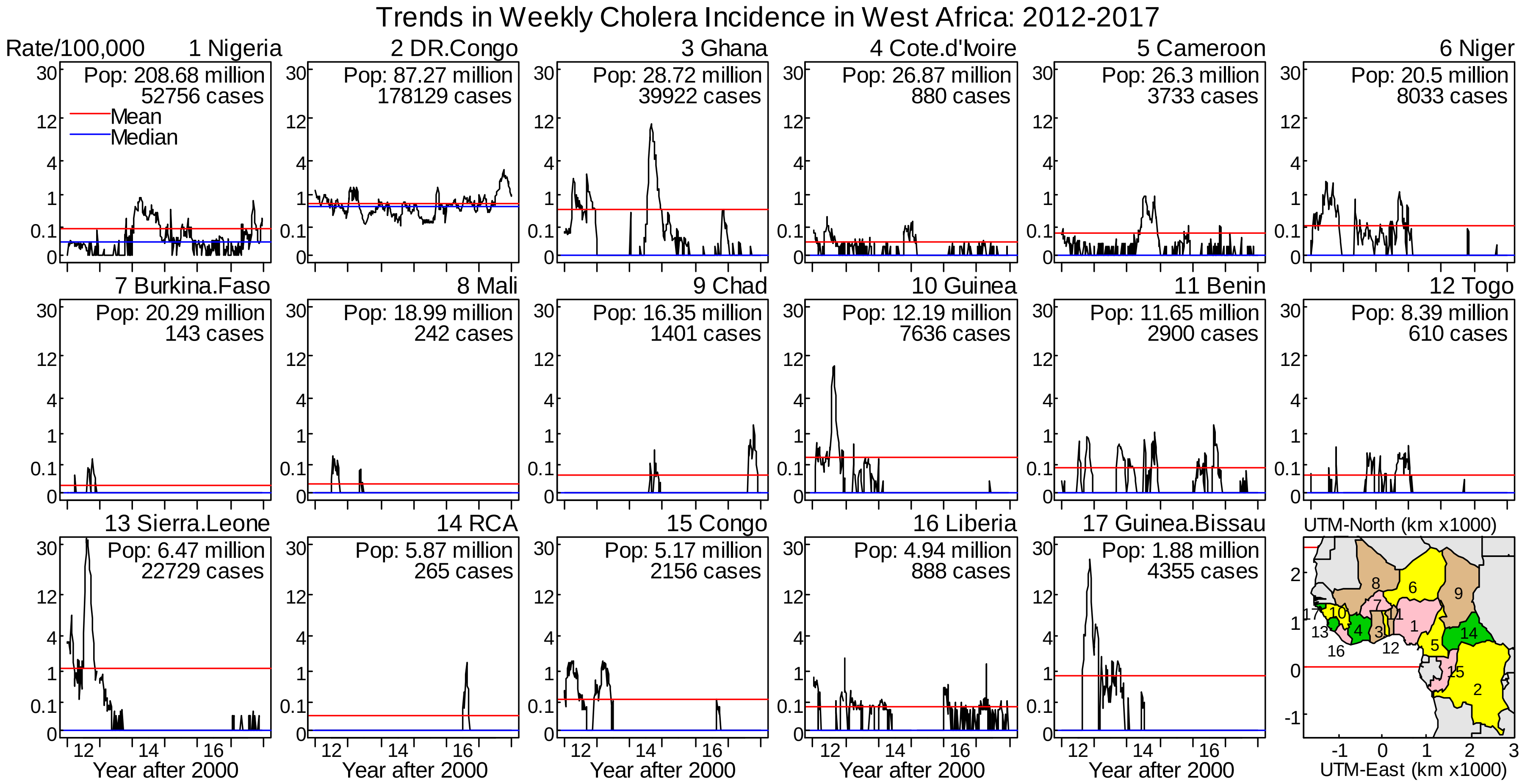


Figure 0.1 Time series plots of weekly incidence rate

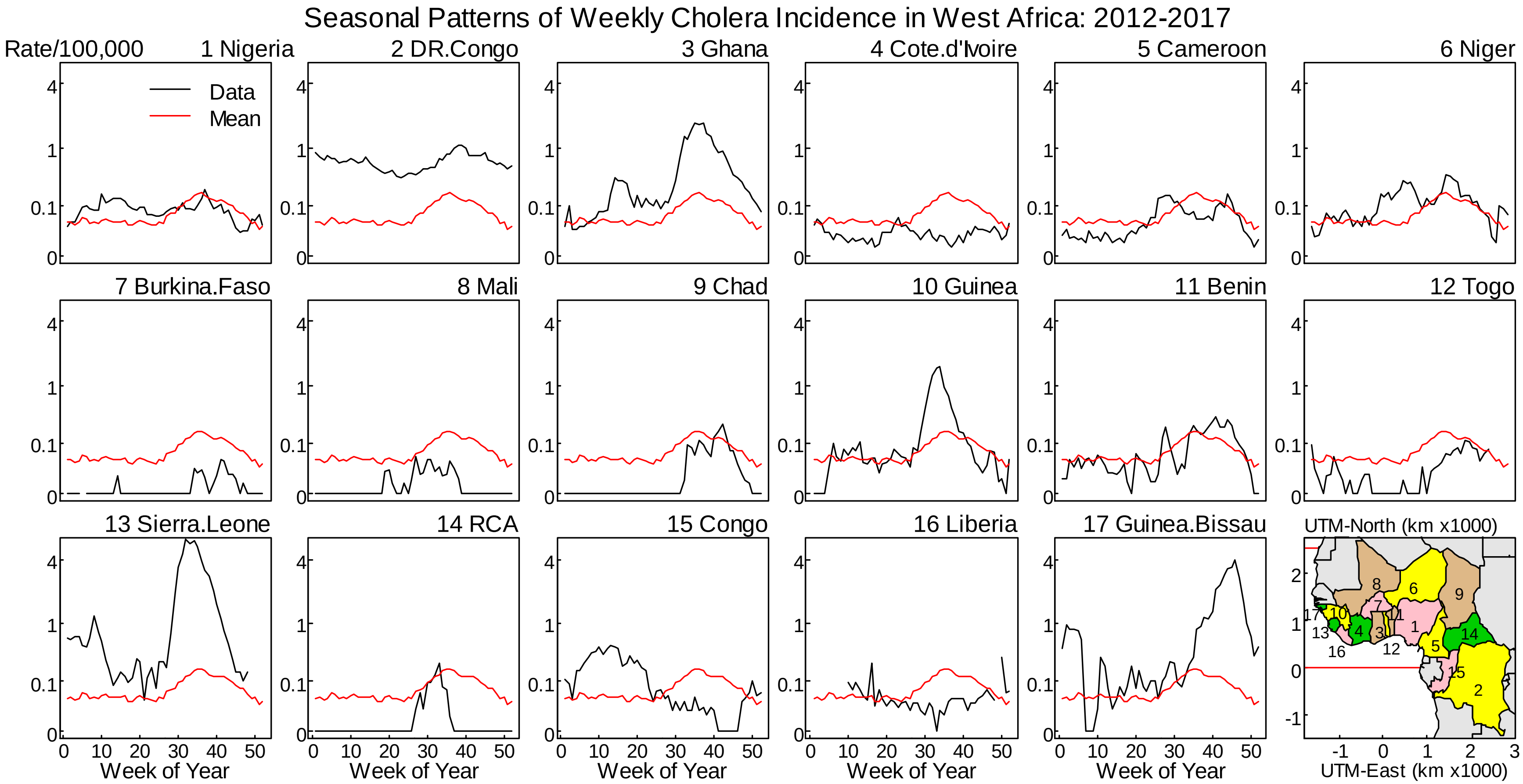
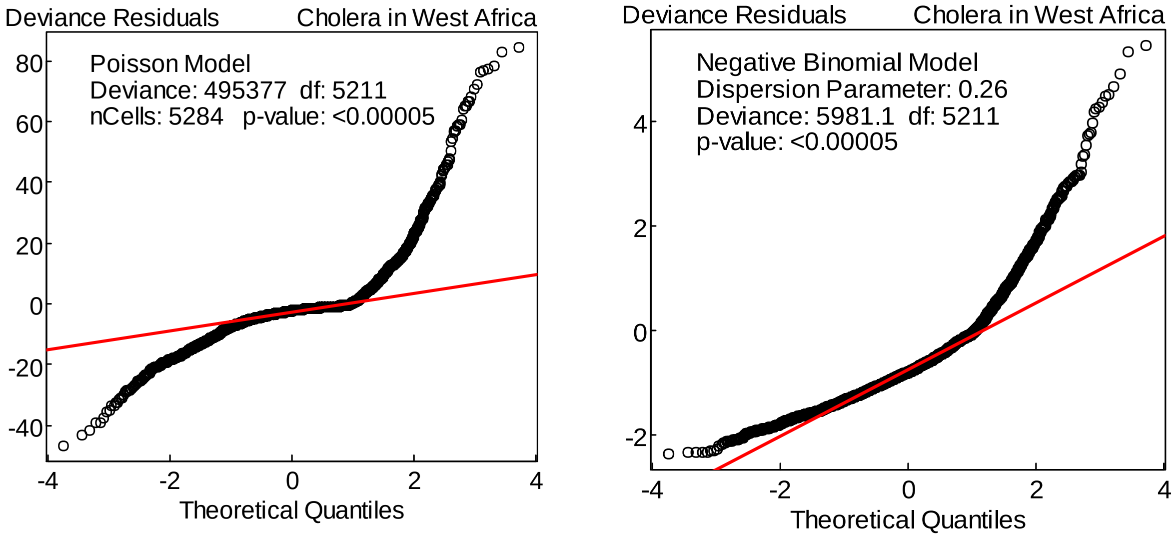


Figure 0.2 Seasonal Patterns of weekly cholera incidence by country

3.3 Model fitting and diagnostic

This subsection illustrates the results for a model of the cholera incidence. Poisson model and negative binomial were modeled since the cholera cases is a count data. Moreover, we suggest the alternative methods to model the incidence rate. The model described replaces incidence by a bivariate outcome-comprising occurrence (taking values 0 if no case occurs or 1 otherwise), and intensity (incidence rate for cases that occur). Logistic regression used to model occurrence, and simple linear regression after log-transforming ("log-linear" regression) used to model intensity.

The detail of each model was shown below. We use both Poisson and negative binomial, with week, year and country as predictive factors. There should be 5,304 (52 weeks x 6 years x 17 countries), but 20 are missing because Liberia has no data for weeks 1-9 in 2015, Sierra Leone has no data for weeks 49-52 in 2012 and Togo has no data for weeks 48-52 in 2012. Figure 3.3 (A) shows that the Poisson regression model fits very poorly according to the residual Q-Q plot. Therefore, we tried to another model which is negative binomial regression model with $\theta = 0.26$. This model shown reduces error but fit remains poor as figures 3.3 (B).



(A) (B)

Figure 0.3 Residuals Q-Q plot for Poisson regression model (A), and negative binomial model (B)

3.4 Confidence intervals for a logistic regression of cholera occurrence

Figure 3.4 shows confidence intervals of cholera occurrence for levels of each risk factor. However, confidence intervals for weeks are relatively wide compared to those for year and country. This may cause from the number of cholera occurrence in each week quite small. Thus, the four-weeks period graph is conducted as shown in Figure 3.5.

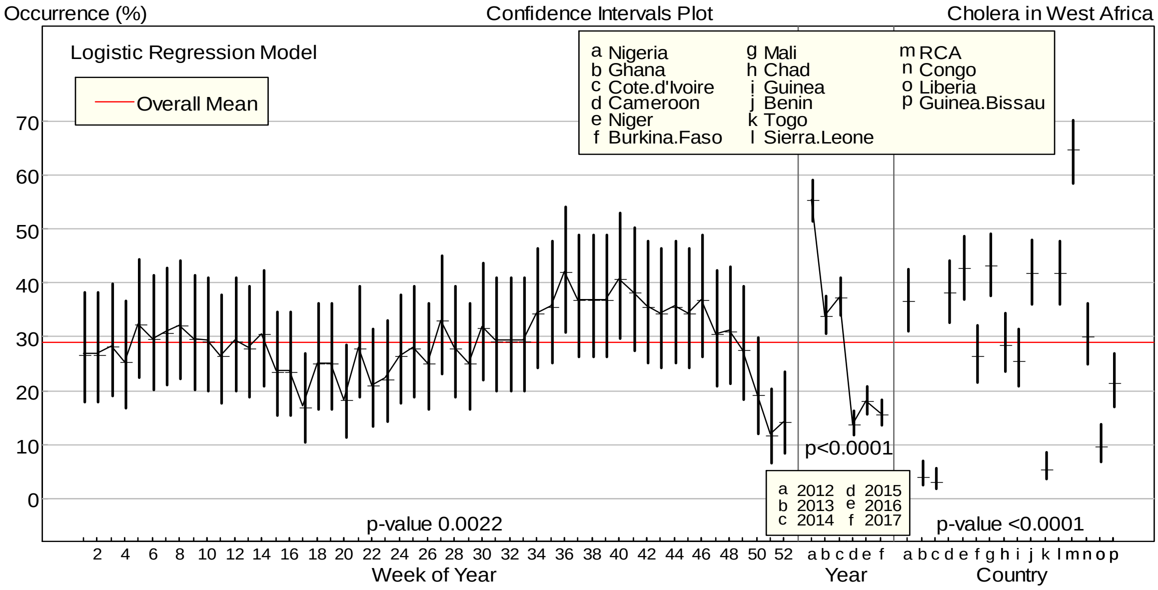


Figure 0.4 Confidence intervals plot for cholera occurrence

Figure 3.5 shows confidence intervals of occurrence for each risk factor using four-week periods, similar to month but with 13 periods each year, to show a seasonal pattern. Confidence intervals now have similar sizes. A confidence intervals plot shows that cholera rates were higher in September and October, lower in 2015-2017, higher in Guinea, Niger and Congo (west) and lower in Cote d'Ivoire, Cameroon, the Democratic Congo and Central African republics, Togo and Guinea Bissau.

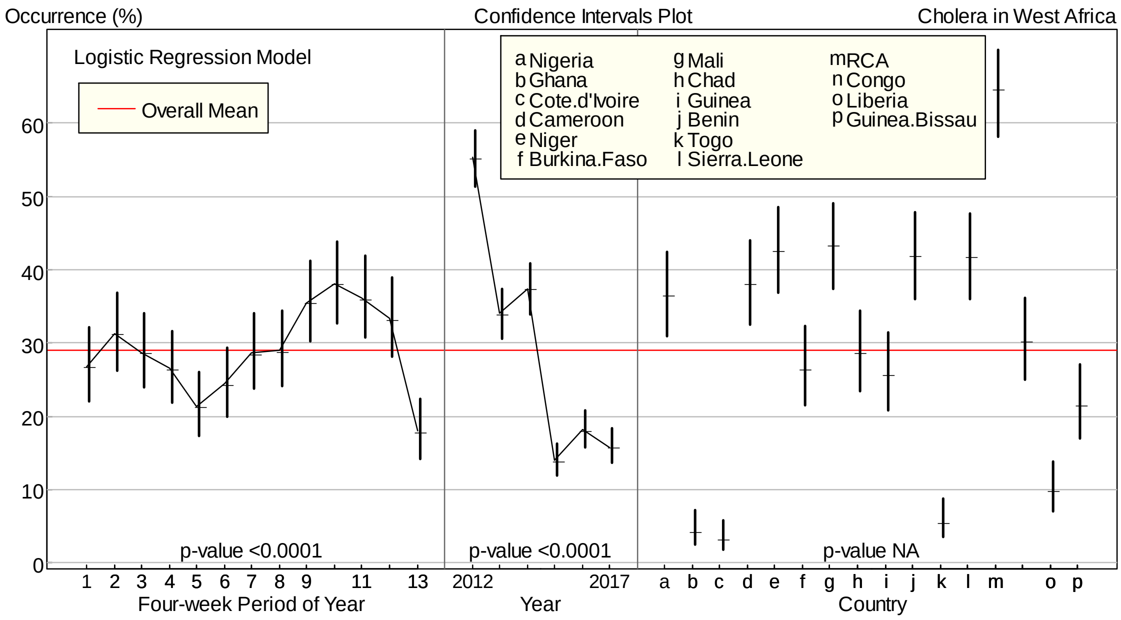


Figure 0.5 Seasonal Occurrence Pattern using Four-week Periods

3.5 Confidence intervals for a log-linear regression of cholera intensity

Figure 3.6 Shows the confidence interval for a log-linear regression of cholera intensity. The green dots denote crude intensity rates, computed directly from data for each factor. The black cross symbol denotes intensity rates computed from the model. They differ from those in the model, due to log transformation bias. The green line represents the overall mean of the intensity. The graph shows that cholera rates were higher in September and October, lower in 2015-2017. The countries with higher intensity rates were Guinea, Niger, and Congo (west) and lowered in Cote d'Ivoire, Cameroon, the Democratic Congo and Central African republics, Togo and Guinea Bissau. In addition, the Q-Q plot of residuals shows that a log-linear model also fits well to this data as shown in Figure 3.7.

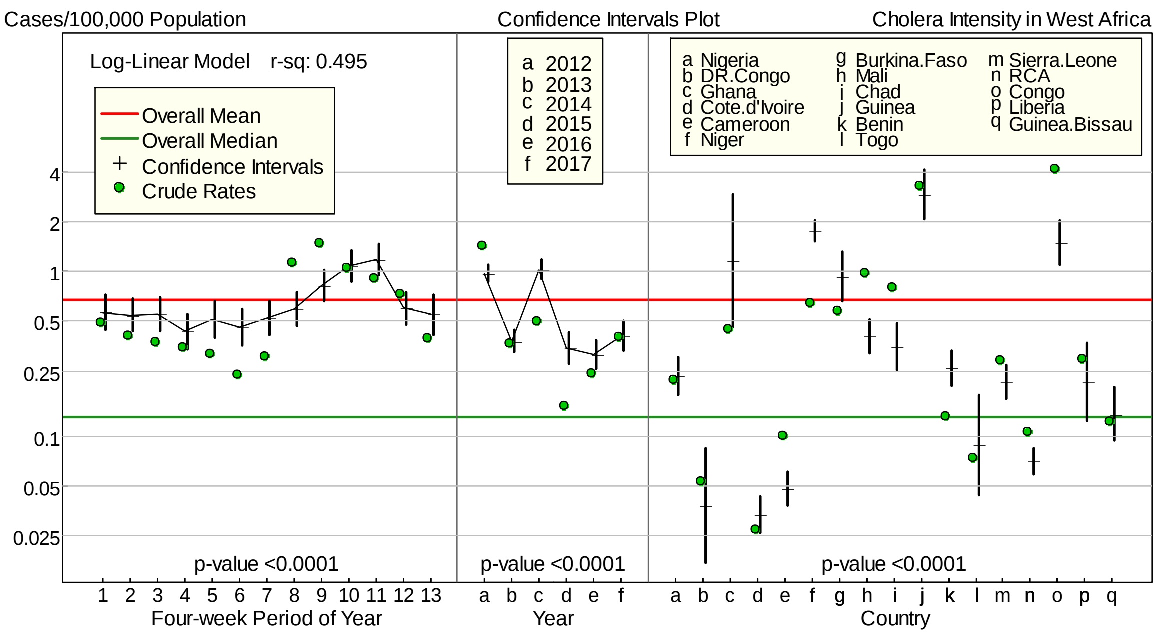


Figure 0.6 Confidence intervals plot for cholera intensity

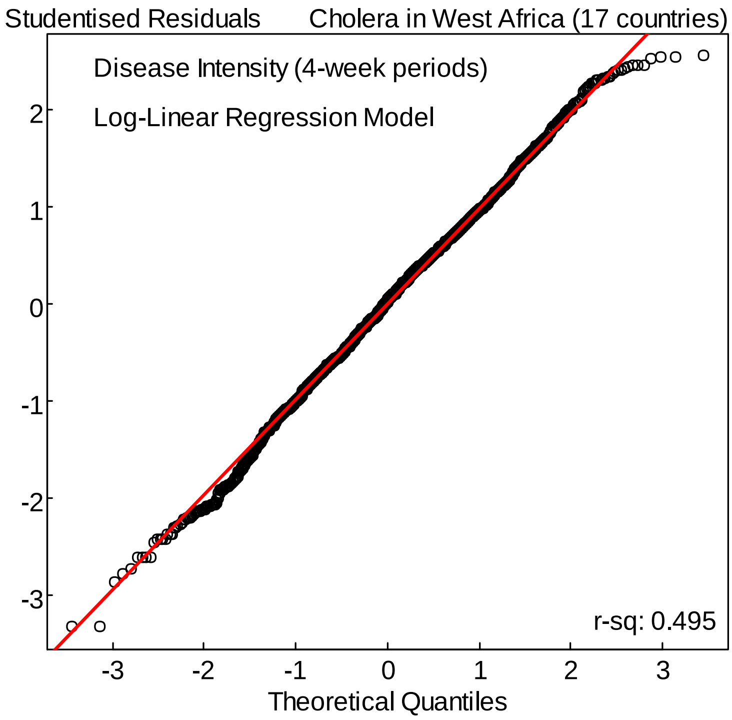


Figure 0.7 Residuals Q-Q plot for Log-linear regression model

3.6 Thematic Maps of Cholera Occurrence and Intensity

Confidence interval plots can be used to divide levels of a risk factor into three groups, depending on the placement of these intervals completely above, around, or below a specified level. Thematic maps show that these metrics have different patterns. For example, Congo (DRC) shows high occurrence and low intensity, whereas the reverse is true for Ghana, Cameroon and the RCA.

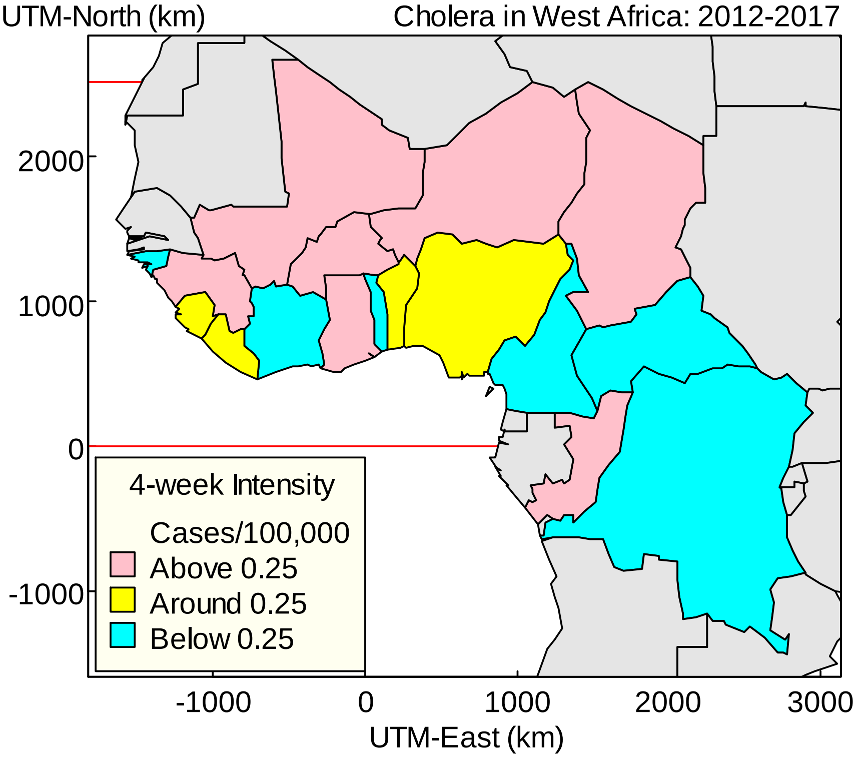


Figure 0.8 Thematic Maps of Cholera Intensity

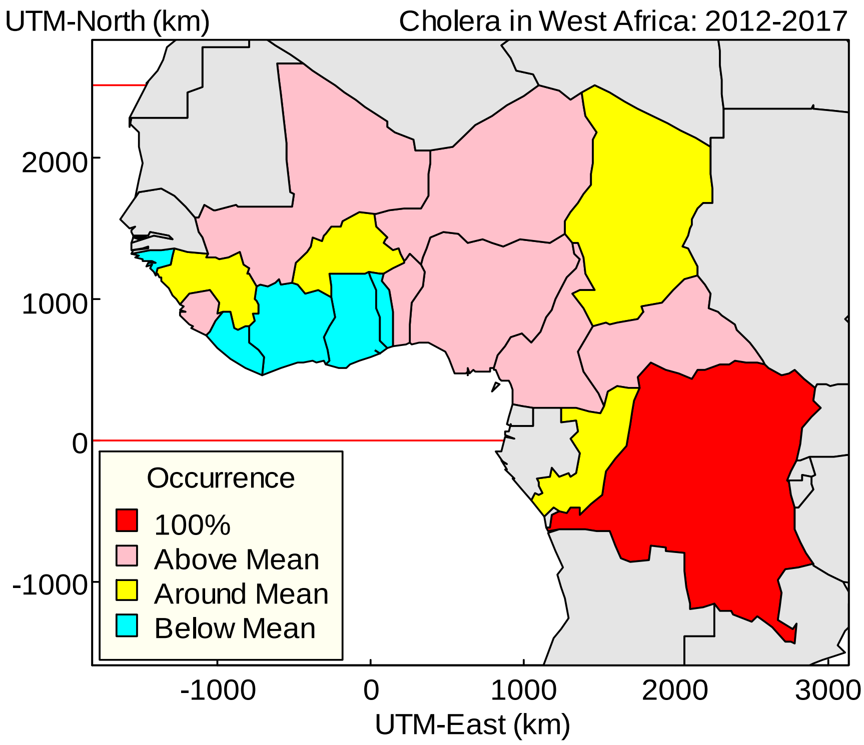


Figure 0.9 Thematic Maps of Cholera Occurrence

Figure 3.10 shows all combinations of occurrence and intensity levels. Spearman's correlation coefficient between occurrence and intensity is -0.04, indicating that these measures are unrelated. From this map, we can see that two northern countries (Mali & Niger) have both high occurrence and high intensity, whereas for three coastal countries (Guinea-Bissau, Cote d'Ivoire & Togo) occurrence and intensity are both low. In four countries (shaded pink or green) occurrence and intensity are quite different.

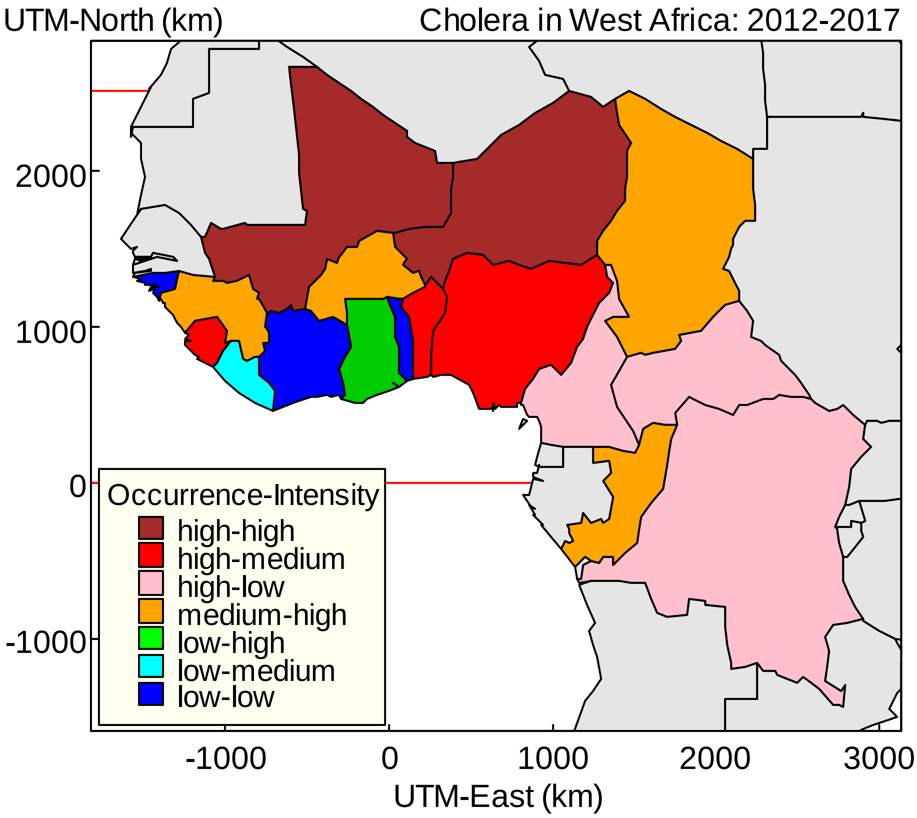


Figure 0.10 Occurrence-Intensity Map

CHAPTER 4

DISCUSSIONS AND CONCLUSION

This final chapter resume the discussion of the previous chapters in terms of results presented, the summary of findings; the limitations recommendations for the future studies also concluded.

4.1 Summary of findings

The overall cases of cholera in West Africa countries from 2012-2017 for the period of six years was totalized 326 778 cases in those 17 countries. The average proportion of cholera incidence rate per 100,000 population in many countries was below 1% and 0.1%.

The highest average proportion was Sierra Leone with 22 729 cases for the population of 6.47 million, whereas Mali and Burkina Faso showed the lowest average proportion with 242 and 143 incidences for the population of 18.99 million and 20.29 million respectively. The DR. Congo was the highest cases rate about 178.129, second to Nigeria 52756 cases appeared and Burkina Faso only 143 cases. In addition, for the trends Ghana showed high in 2014-2015, Sierra Leone started high peak in 2012 and dramatically decreased in 2017, also Guinea Bissau the trends increased in 2012 and no peak at in from 2015-2017.

4.2 Discussions

Since the interesting outcome is the incidence rate of disease, most studies use Poisson and Negative binomial models to illustrate the association between the risk factors and incidence rate. However, in this study was found that those models were not fit to the data. Therefore, we suggest using the occurrence and intensity of cholera disease instead of incidence as the outcome. The logistic regression and log-linear regression to model the occurrence and intensity, respectively. The results showed that the logistic regression and log-linear regression could be fit to the data quite well compared to those models for Poisson and negative binomial based on the Q-Q residuals plot. These two models provide similar results for both seasonal patterns and time series trends. Thus, we may conclude that the way of analyses such data separately fit a logistic model for disease occurrence and a log-linear regression model for disease intensity after excluding non-cases to ensure that zero incidences are not included in the model can be an alternative method to the traditional Poisson regression and negative binomial regression.

The finding from this study was supported by many articles, which applied the logistic regression for determined the association between risk factors and cholera cases. The study by Nsenga (2020) was shown that logistic regression model fit well in predicting the cholera incidence. The model's overall percentage of accuracy in classification is 78.0%, and the area under the ROC is 0.903. Similar to the finding by Rajendran et al. (2007) which stated that the multinomial logistic regression model was an effective approach for predicting the acute diarrheal patient infected with Vibrio cholera compared with discriminant function analysis and log-linear models. Moreover, the study by Musa and Olayemi (2020) also confirmed that the logistic regression can depict the main risk factors statistically significant association with the responses of the patient to cholera treatment.

4.3 Conclusion

Another advantage obtained from this study is the alternative way of analysis the pandemic of cholera disease. Although many statistical and mathematical approaches have been applied to analyze the cholera pandemic in various aspects. However, most of the studies have been used methods complicated and difficult to generalize by researchers who are not familiar with the theory. Therefore, this study was introduced to analyze such kind of data with basic statistical modeling. Although the Poisson regression is one of the well-known methods for analyzing the incidence rate, and the negative binomial is one another method applied when the data is not satisfied with Poisson regression. But this analysis shown that those methods still irrelevant. Aside from the pandemic trend, which may be relevant to organizations or policymakers, the study found that this technique is advantageous to researchers in order for them to continue their work in future studies.

4.3 Limitation and future study

The present study analyzed the cases, incidence and intensity of cholera in West Africa countries; however, it is unable to analyze the death rate of cholera cases in those countries, or the ages, based on the dataset we used. Another factor is lack of information on other factors that could be associated with cholera such as different level of contaminations. Lastly, the patterns of occurrence and intensity differ among countries. Further study is needed to explain why this is so.

REFERENCES

Balakrish, G. (2006). Cholera due to altered El Tor strains of Vibrio cholerae O1 in Bangladesh. Journal of Clinical Microbiology, 44(11), 4211-4213.

Cottrell, V. (2019). Sanitationists, Sewers, and Shoe-Leather: Political and Social Subtexts in British Cholera Epidemiology, Student Library Research Awards. London, UK.pp.1848-1855.

Deen, J., Mengel, M. A., and Clemens, J. D. (2019). Epidemiology of cholera. Vaccine, 22(5), 128-133.

Gaffga, N. H., Tauxe, R. V., and Mintz, E. D. (2007). Cholera: a new homeland in Africa?. The American journal of tropical medicine and hygiene, 77(4), 705-713.

Grove, R., and Adamson, G. (2018). El Niño Events and the History of Epidemic Disease Incidence. In El Niño in World History. London: Palgrave Macmillan. pp. 159-178

Hosoda, Y., Yamaguchi, M., and Hiraga, Y. (1997). Global epidemiology of sarcoidosis: what story do prevalence and incidence tell us?. Clinics in Chest Medicine, 18(4), 681-694.

Legros, D. (2018). Global cholera epidemiology: opportunities to reduce the burden of cholera by 2030. The Journal of Infectious Diseases, 218(suppl_3), pp.137-140.

Mandal, S., Mandal, M. D., and Pal, N. K. (2011). Cholera: a great global concern. Asian Pacific journal of tropical medicine, 4(7), 573-580.

Mahmud, Z. H., Islam, S., Zaman, R. U., Akter, M., Talukder, K. A., Bardhan, P. K., and Alemu, W. (2014). Phenotypic and genotypic characteristics of Vibrio cholerae O1 isolated from the Sierra Leone cholera outbreak in 2012. Transactions of the Royal Society of Tropical Medicine and Hygiene, 108(11), 715-720.

Mourya, D. T., Yadav, P. D., Ullas, P. T., Bhardwaj, S. D., Sahay, R. R., Chadha, M. S., and Khasnobis, P. (2019). Emerging/re-emerging viral diseases and new viruses on the Indian horizon. The Indian Journal of Medical Research, 149(4), 447-450.

Musa, B.T., and Olayemi, O.S. (2020). Application of Logistic Regression Models for The Evaluation of Cholera Outbreak in Adamawa state Nigeria. International Journal of Mathematics and Statistics Studies, 8, 32-54.

Nsenga, N. (2020). Population-Level Determinants of Cholera Incidence in African countries [PhD.thesis]. Walden University.

Okoh, A.I., and Ekundayo, T. C. (2018). A global bibliometric analysis of Plesiomonas-related research (1990–2017). Plesiomonas One, 13(11).

Payton, M., Bush, T. L., Chung, G., Ziegler, B., Eden, P., McElroy, P., and Nguyen, H. N. (2010). Preclinical evaluation of AMG 900, a novel potent and highly selective pan-aurora kinase inhibitor with activity in taxane-resistant tumor cell lines. Cancer research, 70(23), 9846-9854.

Rahman, P. (2018). Global health crises and international cooperation: A comparative framing analysis of narratives told during cholera outbreaks in 1851 and 2017 (Doctoral dissertation). Arts and Social Sciences: School for International Studies. pp. 50-70.

Rajendran K, Ramamurthy T, Sur D. (2007). Multinomial Logistic Regression Model for Inferential Risk Age Groups for Infection Caused by Vibrio cholera in Kolkata, India. J Mod Appl Stat Methods. 6, 324-230.

Shimada, T., Arakawa, E., Itoh, K., Okitsu, T., Matsushima, A., Asai, Y., and Takeda, Y. (2008). Extended serotyping scheme forVibrio cholerae. Current Microbiology, 28(3), 175-178.

World Health Organization. (2017). Cholera and the Global Task Force on cholera control. Geneva: Switzerland.

World Health Organization (2017). Cholera vaccines. WHO position paper. Weekly Epidemiologic Rec, 76(16), pp. 117-240.

World Health Organization. (2000). W.H.O. report on global surveillance of epidemic prone infectious diseases, communicable diseases and surveillance response. (W.H.O./CDS/CSR/ISR/2000).Retrieved from https://www.who.int/csr/resources/ publications/surveillance/ WHO_CDS_CSR_ISR_2000_1/en/.

Yang, J. (2019). The K+ Activation Mechanism of V. cholerae ApbE (Doctoral dissertation). Illinois Institute of Technology, Retrieved from https://search.proquest.com.

VITAE

Name Mr. Amadou Conde

Student ID 6120320003

Education Attainment

Degree	Name of Institution	Year of Graduation
Bachelor of Mass (Communication)	International Islamic University Malaysia	2014

Scholarship Awards during Enrolment

1. Received the 2018 Scholarship awards for Master's and Ph.D. studies under the Thailand's Education Hub for ASEAN Countries (TEH-AC) granted by the Graduate School, Prince of Songkla University

2. Received thesis financial support granted by the Graduate School, Prince of Songkla University

List of Publication and Proceeding:

Conde, A., Dureh, N., and Ueranantasun, A. Trends and Patterns of Cholera Epidemic in West Africa: A Statistical Modeling Study. Journal of Water & Health. (Submitted for publication)