

# Determinants of Dehydration Status and Associated Risk Factors of Cholera Outbreak in Oromia, Ethiopia

Endale Alemayehu<sup>1,\*</sup>, Tsigereda Tilahun<sup>1</sup>, Eshetu Mebrate<sup>2</sup>

<sup>1</sup>Department of Statistics, College of Natural and computational Science, Ambo University, Ambo, Ethiopia

<sup>2</sup>Department of Statistics, College of Natural Sciences, Assosa University, Assosa, Ethiopia

## Email address:

endalestat@gmail.com (E. Alemayehu), tisgeti@gmail.com (T. Tilahun), eshexstat@gmail.com (E. Mebrate)

\*Corresponding author

## To cite this article:

Endale Alemayehu, Tsigereda Tilahun, Eshetu Mebrate. Determinants of Dehydration Status and Associated Risk Factors of Cholera Outbreak in Oromia, Ethiopia. *Biomedical Statistics and Informatics*. Vol. 5, No. 3, 2020, pp. 60-64. doi: 10.11648/j.bsi.20200503.11

**Received:** June 18, 2020; **Accepted:** July 24, 2020; **Published:** August 10, 2020

---

**Abstract:** *Introductions:* Cholera is a diarrheal disease caused by infection of the intestine with the gram-negative bacteria *Vibrio cholera*. According to updated global burden of cholera estimate 2019 in Ethiopia 68,805,272 populations are at risk of cholera with incidence rate of 4 per 1000 population and case fatality of 3.8% estimated annual number of cases 275,221. *Methods:* The main objective of this study is to identify the significant risk factors of dehydration status of cholera outbreak in Oromia regional state of Ethiopia. Ordinal logistic regression was used to model the data by incorporating the assumption behind this novel model. *Results:* The results of the study indicated that of the total 965 cholera patients, most of them 560 (58%) were severely dehydrated by cholera. The overall goodness of model ( $p\text{-value}=0.07$ ) shows that the model fits the data well. Besides, the proportional odds assumption also revealed that the slope coefficients in the model are the same across dehydration status ( $p\text{-value}=0.094$ ). For those have history of travel, the odds of severely dehydrated versus the combined some dehydrated and no dehydrated was  $\exp(1.133804)=3.11$  times higher than those have no history of travel ( $p\text{-value}<0.01$ ). All the other factors like history of contact with other patients, other sick patients in the family, Intravenous and Antibiotics drugs are statistically significant with 5% level of significance to determine the status of dehydration. *Conclusions:* The ordinal logistic regression was fitted the data well and most of the included factors were significant for the dehydration status of cholera outbreak.

**Keywords:** Cholera, Dehydration Status, Ordinal Logistic Regression

---

## 1. Backgrounds

Cholera is a diarrheal disease caused by infection of the intestine with the gram-negative bacteria *Vibrio cholerae*, either type O1 or O139. Both children and adults can be infected. It is one of the key indicators of social development and remains a challenge to countries where access to safe drinking water and adequate sanitation cannot be guarantee [1].

Reports revealed that approximately 1.3 billion people are at risk for cholera in the 69 cholera endemic countries. An estimated 2.86 million cholera cases occur annually in those countries [2]. Among these cases, there are an estimated 95,000 deaths has been occurred with majority burden was accounted by Sub-Saharan Africa. According to updated global burden of cholera estimate 2018 in Ethiopia

68,805,272 populations are at risk of cholera with incidence rate of 4 per 1000 population and case fatality of 3.8% estimated annual estimated number of cases 275,221 [2, 3].

According to WHO weekly Bulletin on outbreaks and other emergencies issued on that was 7 July 2019, from 25 April to 23 June 2019, 688 cholera cases, 23 confirmed and 15 associated deaths (Cholera case fatality rate: 2.30%) have been reported from five regions in Ethiopia. The most affected of all seven regions is Oromia (360 cases, 41%), followed by Amhara (202 cases, 23%), Afar (131 cases, 15%), Addis Ababa (125 cases, 14%), Somali (33 cases, 4%) and Tigray (19 cases, 2%). The main identified risk factor is said to be the contaminated water sources [4].

Different studies revealed that hand-washing, proper waste disposal, personal hygiene and making drinking water safe, health education, preparedness and surveillance at times and

after outbreaks are recognized as important measures in the control of the disease [2-4].

Now a day the problems of acute watery diarrhea are becoming the leading health cases in Ethiopia. As Ethiopia is one of WHO regions in having cholera epidemic and is among countries that have high risked peoples, this study was fully focused on modeling the dehydration of cholera with different factors. Thus, this study was aimed to identify all significant factors that are responsible for the cholera outbreak. At the same time the rapid difference in the dehydration status on the peoples with cholera also discussed so that the focus of its stage was identified. Therefore, the research questions on identifying the risk factors of cholera, addressing the difference in the dehydration status and application of appropriate statistical model has been covered in this study.

## 2. Methods

### 2.1. Study Area

This study was conducted in Oromia regional state of Ethiopia. The Oromia region is one of the nine regional states of Ethiopia, the homeland of the Oromo. It is bordered by the Somali Region to the east; the Amhara Region, the Afar Region and the Benishangul-Gumuz Region to the north; South Sudan, Gambela Region, and Southern Nations, Nationalities, and Peoples' Region to the west; and Kenya to the south. It is also the largest regional state covering 286,612 square kilometers.

### 2.2. Source Data

The data for this study was secondary data which is obtained from Ethiopian Public Health Institute (EPHI). All cholera patients reported in for a year in 2019 has been considered for this study. Thus, 965 patients in the study area and period were considered for the study.

### 2.3. Variables of the Study

#### 2.3.1. Response Variable

The response variable of the study was dehydration status which is coded as:

$$\begin{cases} \text{No Dehydration (coded as 0)} \\ \text{Some Dehydration (coded as 1)} \\ \text{Severe Dehydration (coded as 2)} \end{cases}$$

Dehydration is the loss of water and fluids from the body which leads to immediate death. The degree of dehydration is graded according to symptoms and signs that reflect the amount of fluid lost under different criteria<sup>3</sup>.

#### 2.3.2. Explanatory Variables

According to different literature and records of patient history, the following factors were identified. The independent variables are sex of patients, age categories of patients, travel history of patients, patients' history of

contacts with other peoples, whether other sick person in the family, whether the patients vomiting, Intravenous, Antibiotics and whether the patients had lab test. The factors are coded as follow.

Variable	Code
Sex	Female=0, Male=1
Age Group	<5=0, 5-14=1, 15-44=2, >44=3
History of travel	No=0, Yes=1
History of contact	No=0, Yes=1
Other sick person	No=0, Yes=1
Vomiting	No=0, Yes=1
Intravenous	No=0, Yes=1
Antibiotics	No=0, Yes=1
Lab test	No=0, Yes=1

### 2.4. Statistical Model

Ordinal logistic regression model is a type of logistic regression model that are used to analyze ordinal dependent variables. For this study since the outcome variable is in the form of ordinal scale (ordered dehydration status as no dehydration, some dehydration and severe dehydration).

Mostly considering the method is natural for multiple categorical response having more than two possible values. And also the most well-known of these ordinal logistic regression methods is the proportional odds model. The main idea under proportional odds model is that re-expressing the categorical variable regarding a number of tertiary variables based on internal cut-points in the ordinal scale.

Therefore, the ordinal logistic regression model can be formulated as:

$(\ln \gamma_i) = \beta_i - [\beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n]$  where  $I$  goes to the number categories minus one. As it is known, this is the part the researchers really want to investigate. So having seen the status of cholera diseases, it is clear to make ordinal response like no dehydrated, some dehydrated and severe dehydrated.

Consequently, if the researchers know that a category is ordinal then there are special models that tell us how the covariate relates to someone being higher or lower on the scale. A proportional odds model that deals limiting some factors to have common parameters and leaves other factors free to have separate parameters is called a partial proportional odds model. Therefore, the basic form the generalized linear written as follow: -

$$\text{link} \gamma_i = \frac{\beta_i - [\beta_1 X_1 + \beta_2 X_2 + \dots + \beta_K X_K]}{\exp(\alpha_1 \gamma_1 \alpha_2 \gamma_2 + \dots + \alpha_m \gamma_m)}$$

Where  $\gamma_i$  is the cumulative probability for the  $i^{th}$  category and  $\beta_i$  is the threshold for  $i^{th}$  categories,  $\beta_1, \beta_2, \dots, \beta_k$  are the parameters to be estimated,  $X_1, X_2, \dots, X_K$  are the covariates and  $k$  is the number of covariates. From this equation the numerator to the right side determine the location of the model. The denominator equation could determine the scale. The  $\alpha_1, \alpha_2, \dots, \alpha_m$  are coefficients for the scale components and  $\gamma_1, \gamma_2, \dots, \gamma_m$  are  $m$  covariates for the scale components (selected from the same set of variables as  $X$ 's).

This method was selected as the right model to see the ordinal outcome variable that may have better advantage to examine the effects of explanatory variables at each category of response variable. Hence, it enhances the betterment of the study [10].

### 3. Results and Discussions

#### 3.1. Results

The results of the study indicated that of the total 965 cholera patients, most of them 560 (58%) were severely dehydrated by cholera; whereas only 123 (12.7%) was not dehydrated and the other 282 (29.2%) was registered as somewhat dehydrated.

The graphical presentation in the figure 1 below is classification of dehydration status by age category expressed for male and female patients separately. It clearly shows that both male and female patients aged 15-44 were the most

severely dehydrated compared to the other age groups. In all dehydration status, the 15-44 age groups were counted to be the largest group included as cholera infected in the study area and/or period. The graphical results supported the frequency results above that the severe dehydrated were exceeding the other stages.

The table 1 below is the summary of chi-square test of association between dehydration status and the identified factors. It indicated that there is no enough evidence to say that sex of patients their age group and whether they vomiting have significant association with dehydration status. Despite, history of travel, history of contact, other sick person in the family, IV, antibiotics and lab test has statistically significant association with dehydration status of peoples those have cholera infection. Thus, to now the magnitude and strength of those factors with the dehydration status, the researcher intended to fit appropriate statistical model.

**Table 1.** Chi-square test of association between dehydration status and the identified factors.

Independent variables	Dehydration status			
	Test	Chi-square value	df	P-value
Sex	Pearson Chi-square test	1.99	2	0.37
Age group		3.64	6	0.73
Hx travel		13.21	2	0.001**
Hx contact		57.16	2	0.000***
Other sick		7.82	2	0.02*
IV		382.19	2	0.000**
Antibiotics		350.98	2	0.000***
Lab-test		14.47	2	0.000***
Vomiting		4.99	2	0.08

The likelihood ratio chi-square test is -780.21446 with p-value of 0.000. This indicated that the model as the whole is statistically significant as compared to the null model. It is to mean that at least one of the considered factors was significant in determining the dehydration status.

The results presented under table 2 below were the coefficient model table which is used to know the significance of individual factors. It indicated the variables

like history of travel, history of contact with other patients, other sick patients in the family, Intravenous and Antibiotics drugs are significant with 5% level of significance. Thus, the researcher can conclude that those factors are potential in determining the status of dehydration. The goodness of model fit of table 3 (Appendix) shows that the model fit the data well. The test of parallel lines of table 4 (in the appendix) is also revealed that the assumption is fulfilled.

**Table 2.** Parameter estimation under ordinal logistic regression.

Dehydration	Coef.	Std. Err.	Z	P>z	[95% Conf.	Interval]
Hx_travel						
Yes	1.133804	.280116	4.05	0.000***	.5847867	1.682821
Hx_contact						
Yes	1.008163	.1505741	6.70	0.000***	.303283	1.7130428
Other_sick						
Yes	.3112465	.2089097	1.49	0.0136*	.207021	1.28209
IV						
Yes	1.000286	.3262449	3.07	0.002**	.3608578	1.639714
Antibiotics						
Yes	.8170224	.3198448	2.55	0.011*	.1901382	1.443907
Lab_test						
Yes	.4792511	.2505636	1.91	0.056	-.0118446	.9703468
/cut1	-1.738741	.1287864			-1.991157	-1.486324
/cut2	.1866333	.1085266			-.026075	.3993415

### 3.2. Discussions

The descriptive results of the study indicated that of the total 965 cholera patients, most of them 560 (58%) were severely dehydrated by cholera. Many other previous studies also revealed that the status of dehydration of patients was categorized under severity.

Most the factors included in study have significant association with severity of cholera and its dehydration status. However, sex of patients and their age categories was no significant to determine the status of cholera dehydration. The results of Pearson chi-square test of association indicated that history of travel of patients ( $p$ -value<0.001), history of patients contact with other patients ( $p$ -value<0.001), whether there is other sick patients in the family ( $p$ -value<0.001), whether intravenous was given ( $p$ -value<0.05), whether antibiotics was given ( $p$ -value<0.01) and whether the patients had lab test ( $p$ -value<0.01) have significant association with the dehydration status of cholera outbreak at 1% and 5% level of significance. Hence, the result was consistent with the previous work by [2, 4].

The ordinal logistic regression, model presented in table 2 includes the significant factors of dehydration status of cholera outbreaks. For those have history of travel, the odds of severely dehydrated versus the combined some dehydrated and no dehydrated was  $\exp(1.133804)=3.11$  times higher than those have no history of travel ( $p$ -value<0.001). Thus, it is to mean that the severity of dehydration status of cholera outbreak has positive relationship with those have history of travel. It is agreed with previous work [2].

The odds of severely dehydrated versus some dehydrated and no dehydrated was  $\exp(1.008163)=2.74$  and  $\exp(0.3112465)=1.37$  for those have history contact and those have sick patients in their family respectively ( $p$ -value<0.01). Previous studies revealed that the variable was significant factors for cholera outbreak [5, 6].

Likewise the other factors like whether the patients had intravenous injection and whether the antibiotics treatment was given by the patients was included in the study and found significant with 1% and 5% level of significance. Other previous studies are also consistent with this result [7-9].

The results of cut1 (-1.738741) in table 2 above refers to the estimated cut-point on latent variable to differentiate no dehydrated from some dehydrated and severe dehydrated when the value of all predictors are evaluated at their reference categories. Whereas cut2 (0.186633) is the estimated cut-point used to differentiate some dehydrated from no dehydrated and severe dehydrated keeping the factors to be at their reference. Hence, the magnitude of the output is consistent with the study by [10].

Proportional odds assumption of ordinal logistic regression was met for this study ( $p$ -value=0.094) which is to mean that the effects of explanatory variables over all the categories of dehydration status was expected to be similar.

## 4. Conclusions

This study was aimed to identify the risk factors of cholera dehydration status in Oromia regional state with the secondary data obtained from EPHI. The ordinal logistic regression was used to determine the significant factors of the dehydration status of cholera outbreak. The basic research questions and objectives of the study were addressed. The assumption of test of parallel line was fulfilled for the study. All the factors except lab-test were significant under the ordinal logistic model.

## Declarations

During conducting the study, the investigators have included the following declaration.

### *Ethics Approval and Consent to Participate*

Since the data was secondary and has no any contact to the patients, the Ethical approval was not necessary to have. But, securing all the information was confirmed by EPHI.

### *Data Permission*

The data is not publically available; but it will provide by the researcher at any time of request.

### *Funding*

Not applicable.

### *Authors' Contributions*

This study was designed and compiled by EA as the principal investigator. The development of the basic research questions, identifying the problems and selecting appropriate statistical models have been done by him. The review of literatures and related works were addressed by TT and Edition of the overall progress of the work was supported by EM.

### *Abbreviations*

EPHI: Ethiopian Public health Institute, WHO: World Health Organization.

### *Consent for Publication*

Not Applicable.

## Competing Interests

The authors declare that they have no competing interest.

## Acknowledgements

We thank EPHI to have such compiled data.

## Appendix

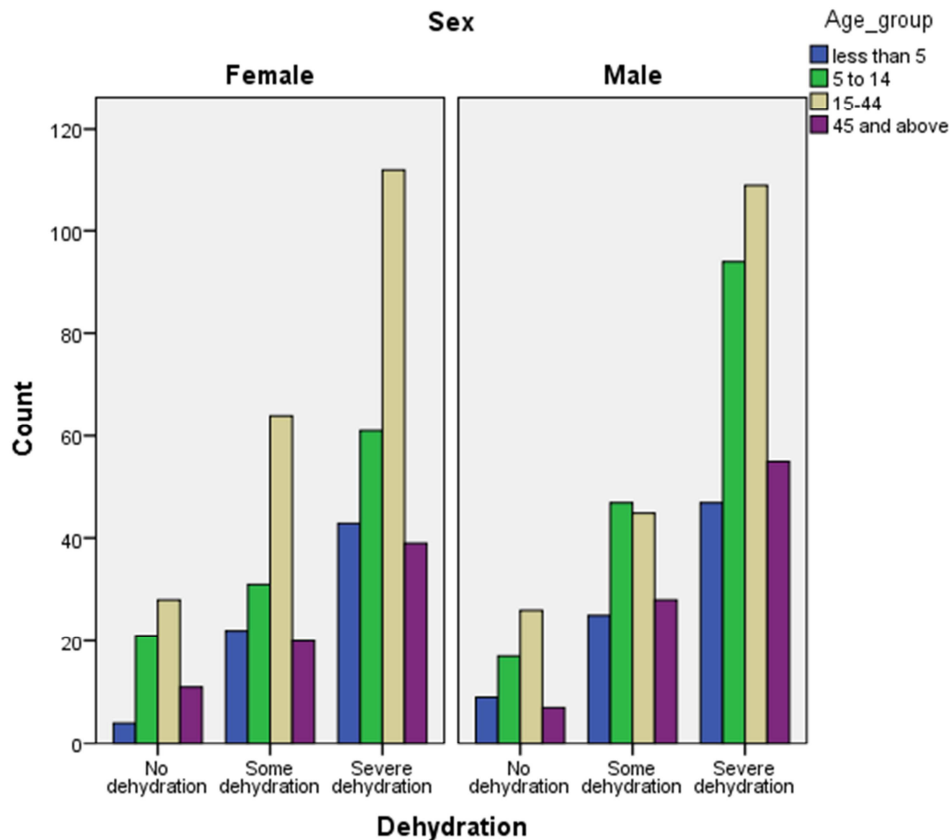


Figure 1. Classification of dehydration status by age category expressed for male and female patients.

Table 3. Goodness of model fit.

	Chi-Square	df	Sig.
Pearson	1395.069	428	.071
Deviance	1010.006	428	.083

Table 4. Test of parallel lines.

Model	-2 Log Likelihood	Chi-Square	df	Sig.
Null Hypothesis	1109.449			
General	1106.894	442.555	8	.094

## References

- [1] Dureab, F. *et al.* Risk factors associated with the recent cholera outbreak in Yemen: a case-control study. 1–6 (2018).
- [2] Health, E. Guideline on CHOLERA OUTBREAK MANAGEMENT Ethiopia.
- [3] Ohene, S., Klenyuie, W. & Sarpeh, M. Assessment of the response to cholera outbreaks in two districts in Ghana. *Infect. Dis. Poverty* 1–11 (2016) doi: 10.1186/s40249-016-0192-z.
- [4] Tull, K. Humanitarian interventions in Ethiopia responding to acute watery. 1–12 (2018).
- [5] Nsagha, D. S. *et al.* Assessing the risk factors of cholera epidemic in the Buea Health District of Cameroon. *BMC Public Health* (2015) doi: 10.1186/s12889-015-2485-8.
- [6] Phiri, P., Nzala, S. H. & Baboo, K. S. Factors Associated With the Recurring Cholera Outbreaks in Sinazongwe District of Southern Zambia. 42, 184–192 (2015).
- [7] Okugn, A. & Woldeyohannes, D. Food hygiene practices and its associated factors among model and non model households in Abobo district, southwestern Ethiopia: Comparative cross-sectional study. 1–9 (2018).
- [8] Open, C. & Journal, A. Treatment Outcome of Tuberculosis and Associated Factors at Gimbi Town Health Facilities Western Oromia,. 2, (2017).
- [9] César, H. *et al.* of Infectious Diseases East-Western Dynamic and Driving Factors of Cholera Epidemics in the Democratic Republic of the Congo. 3, 2009–2010 (2018).
- [10] Patience Ebekeyi *et al* Analysis of Factors Contributing to the Spread of Cholera in Developing Countries; 51 (2): 121-7 2019.