አዲስ አበባ ዩኒቨርሲቲ

ADDIS ABABA UNIVERSITY

ENQUIRY. MASTERY. SERVICE.

DETERMINANTS OF ANEMIA AMONG UNDER–FIVE CHILDREN IN TANZANIA

BY:

WOLDESENBET MULU

A THESIS SUBMITTED TO THE DEPARTMENT OF STATISTICS IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF MASTER OF SCIENCE IN STATISTICS (BIOSTATISTICS)

ADVISOR: Prof. ESHETU WENCHEKO

June, 2024

ADDIS ABABA, ETHIOPIA

DECLARATION

I, the undersigned, declare that the thesis is my original work, has not been presented for

Consideration to this any other university, and all of the sources from which the thesis's components came duly acknowledged.

Name of candidate	**Signature**	**Date**
Woldesenbet Mulu	________________	______________________

This thesis proposal has been submitted for consideration with my approval as a university advisor.

Name of advisor	**Signature**	**Date**
Eshetu Wencheko (Professor)	_____________________	_______________________

ADDIS ABABA UNIVERSITY

Office of Graduate Program

College of Natural Science

Department of Statistics

This is to certify that the thesis prepared by Woldesenbet Mulu , entitled: “DETERMINANTS OF ANEMIA AMONG UNDER–FIVE CHILDREN IN TANZANIA” and submitted in partial fulfillment of the requirements for the Degree of Master of science in Statistics (Biostatistics) complies with the regulations of the University and meets the accepted standards with respect to originality and quality.

Signed by the Examining Committee:

	Name	Signature	Date
Examiner	______________________	______________	_____________
Examiner	______________________	______________	_____________
Advisor	______________________	______________	_____________

____

Chair of Department or Graduate program coordinator

ACKNOWLEDGEMENT

First of all I would like to thank the Almighty God for being on the side of me in the efforts towards my successful completion of the graduate study and helping me with everything.

Secondly, my deepest gratitude and heartfelt thanks go to my advisor, Eshetu Wencheko (Prof.), for giving his support, valuable comments, and advice by sharing his golden time throughout the completion of my thesis, especially by giving him time to balance my responsibilities as both a student and an employee. I also sincerely thank all other lecturers at Addis Ababa University, Statistics Department, who helped me with valuable advice for my work.

Finally, I would like to thank my classmates and friends who wish me academic success.

ACRONYMS AND ABBREVIATION

AIC = Akaki information criterion

BIC = Bayesian information criterion

EDHS=Ethiopia demographic and healthy survey

Hgb = Hemoglobin

MIS = Malaria indicator survey

PHS= Population and Housing

POM = Proportional odds model

SCD = Sickle cell disease

SSA = Sub Sahara Africa

TDHS = Tanzania demographic and health survey

U5 = Under five

WHO = World health organization

Abstract

Determinants of anemia among under–five children in Tanzania

Woldesenbet Mulu

Department of Statistics

Addis Ababa University, Ethiopia

Anemia is a health condition with low red blood cell count, which leaves the blood ability to carry oxygen insufficient to meet the body's metabolic needs at the cellular level. It is a global public health problem that affects people in both developing and developed countries. Worldwide the prevalence of anemia is 47% in children younger than 5 years 70% of children living in low-income countries. The primary purpose of this study is finding the major factors of anemia among children under- five years of age in Tanzania. The data source for analysis was the 2022 TDHS-MIS in under-five children anemia. Anemia level is divided into three categories non anemic, mild,,and moderate and above anemic. Included in the current study were 4149 children under the age of five. In order to determine the sociodemographic, clinical, and health-related variables and investigate the variation of anemia in clusters, descriptive statistics and multilevel model techniques were utilized. For statistical software STATA 16 was used for analysis. About 59.89% of under five children were anemic. Multilevel ordinal logistic regression model showed that child's age, mother's anemia status, geographical zone, stunting levels, mother's education level, mother's age, child's gender, and recent occurrence of diarrhea impact the anemia status of children.. It is imperative that the government prioritize the development and implementation of nutrition programs, emphasize healthy habits to prevent diarrhea, and educate women to enable reducing prevalence of anemia.

Contents

List of Tables

CHAPTER ONE

1. INTRODUCTION

1.1 Background of the Study

Anemia is a healthy condtion that is characterized by a decreased number of red blood cells or hemoglobin levels that result in the insufficient oxygen-carrying capacity of blood to meet the cellular metabolic demand of the body (WHO, 2011). It is a worldwide public health issue that is linked to a higher risk of morbidity and mortality, particularly in pregnant women and children under five (U5). It affects people in both industrialized and developing nations. Oftentimes, malaria parasitemia is the most prevalent illness linked to anemia (Mghanga et al., 2017).

Low blood hemoglobin levels, or anemia, are a significant worldwide public health issue. They are specifically common in many underdeveloped nations, where high rates of anemia are experienced by particular populations. This critical health concern impacts the well-being, daily functioning, and productivity of billions of individuals worldwide. Iron deficiency is a first cause of anemia, often intertwined with deficiencies in folate, vitamin B12, and infections (Milman, 2011).

Anemia is a public health issue that impacts communities in both affluent and impoverished nations. Although the main cause is iron deficiency, it is rarely found alone. More constantly it coexists with several other factors.similar as malaria, parasitic infection, nutritive scarcities, and hemoglobin apathy(McLean et al., 2009) .

Anemia is a serious public health issue that has a harmful impact on people's health and ability to support population expansion economically. Understanding the frequency and changes of anemia, developing public health initiatives, and directing clinical techniques for patient treatment and management all depend heavily on having a specific case definition. (Garcia-Casal et al., 2019) .

Anemia is a major global public health concern, with an estimated 47% of children under 5 having anemia globally. According to reports, the prevalence is around 30% in under-5 indigenous children who are underprivileged worldwide and 70% in children who live in low-income nations.(Mitchinson et al., 2019).

Anemia is a health condition in which the red blood cells count or hemoglobin concentration within them are lower than normal (WHO, 2011) . In the World Health Organization's 2015 report, an estimated 273 million children under the age of five were anemic; regional differences were found in 48.6% of Eastern Mediterranean countries, 53.8% of South-East Asian countries, 62.3% of African countries, 22.69% of

European countries, 22.3% of American and Western Pacific countries, and 21.9% in Western Pacific Region. (WHO, 2011).

In sub-Saharan Africa (SSA), anemia is among the leading causes of death for children admitted to hospitals. Anemia affects 75% of children under – five in East Africa; the frequency ranges from 44% to 76%. Anemia risk factors might change depending on the population and the environment.These factors include intestinal parasites, malaria, HIV infection, nutritional deficiencies, the habit of drinking tea during meals, hematological malignancies, and chronic conditions such as sickle cell disease (SCD)(Simbauranga et al., 2015) .

According to (EDHS, 2016), Over half of U5 children (57%) have some form of anemia, with mild anemia 25%, moderate anemia 29%, and severe anemia 3%. The prevalence of anemia declines with the age of the child; it peaks at 78% for 6 to 8 month olds and falls to 40% for 48 to 59 month olds. The prevalence of anemia generally decreases with increasing maternal educational status and improving household wealth.

In Ethiopia the occurrence of anemia exhibits geographical variation, with higher rates in the eastern region and lower rates in the southwestern area of the country. This spatial disparity in anemia prevalence may be attributed to differences in dietary habits, susceptibility to infectious diseases, access to healthcare facilities, or other factors (Ejigu et al., 2018) .

Anemia continues to be a significant public health issue in Sub-Saharan Africa among U5 children, with the prevalence of anemia ranging from 43% in Zaire to 74% in Tanzania. (Cornet et al., 1998).

Childhood anemia, while declining globally, remains a significant issue in countries with poor and moderate incomes like Tanzania. Despite efforts, regional disparities persist, even in agriculturally productive areas. The reasons for this decline and the remaining challenges related to child anemia in Tanzania require further investigation. The five-year multisectoral nutrition action plan of the Tanzanian government targeted to lower the prevalence of child anemia from the rate of 58% in 2015 by 30%. (Sunguya et al., 2020).However, according to the TDHS-MIS 2022, the prevalence has remained relatively stable.

1.2 Statement of the problem

Anemia is a worldwide public health problem affecting both developing and developed nations, with major consequences for human health as well as social and economic development. It occurs at all stages of human life cycle but is more prevalent in pregnant women and young children (WHO, 2008). Although the global anemia prevalence is estimated to be 47% in U5 children. Africa and Asia account for more than 85% of the absolute anemia burden in high risk groups (Balarajan et al., 2011) .

Anemia impacts a third of the world's population and contributes to increased morbidity and mortality, decreased work productivity, and impaired neurological development (Chaparro & Suchdev, 2019) . The prevalence of anemia also varies by geographic region. SSA, South Asia, Caribbean and the Oceania had the highest anemia prevalence across all age groups and both sexes in 2010 (Kassebaum et al., 2014) .

In the world a serious public health problem is anemia that regularly affects under five children menstruating adolescent girls and women, and pregnant and postpartum women. In 2019, anemia affected 40%, 30% and 30% of U5 children, pregnant women and women 15–49 years of age respectively. It is most prevalence in developing countries specially mid-range and lower economic level. Anemia increases the risk of infections and death, causes extreme fatigue, and impairs cognitive performance, loss of earnings, poor pregnancy outcomes, and poor growth and development. It is a strong indicator of overall health. Poor nutrition, infections, lack of iron in the blood, chronic diseases, heavy menstruation, and pregnancy issues. Anemia affected an estimated 103 million children and 106 million women in Africa. South-East Asia had 83 million children and 244 million women affected (WHO, 2023)

According to (TDHS-MIS, 2022), the prevalence of anemia among Tanzanian children decreased from 72% in 2004–05 to 59% in 2010, and has since remained relatively stable. The overall prevalence of anaemia among U5 children, based on capillary blood samples, is 59%**.** Among these children, 26 % exhibit mild anaemia**,** 31% have moderate anaemia, and 2% suffer from severe anaemia**.**

Therefore, this study attempts to identify determinate of U5 childhood in Tanzania. The statistical models selected for the purpose is ordinal multilevel regression based on its severity levels (no anemia, mild and moderate and above anemic).

1.3 Objectives of the study

1.3.1 The general objective

- ✓ The main objective of this study is to identify the main determinants of anemia among children under- five years of age in Tanzania.

1.3.2 Specific objectives

- ✓ To determine between cluster variations anemia among children U5 years of age in Tanzania.

- ✓ To examine the prevalence of anemia among U5 children.

1.4 Significance of the study

The use of this study can inform evidence-based interventions and public health programs tailored to the needs of the Tanzanian population. Additionally, the findings of this study could contribute to the global understanding of childhood anemia and serve as a basis for developing strategies to address this critical public health issue in similar socio-demographic contexts.

CHAPTER TWO

2. LITERATURE REVIEW

Several studies have been conducted to identify determinants of anemia among children. Most of these studies used survey or censuses data and revealed significant association between socioeconomic, demographic, nutritional and clinical factors and determinates of anemia in U5 children.

(Chungkham et al., 2021) undertook a study in India to identify determinants of anemia U5 children using a Bayesian geo-additive model. In the northern and central regions of India, the influence of unknown variables on childhood anemia is quite substantial at a 95% posterior credible interval. However, the majority of the states in India's northeast region displayed poor spatial impacts. The age of children increases, its effect on childhood anaemia decreases, which indicates, older children are less likely to have the risk of childhood anaemia. There was a non-linear correlation found between the age of the mother and childhood anemia. This suggests that mother's of all ages, but especially those between the ages of 15 and roughly 25, are more likely to produce children who are anemic.The risk of childhood anemia started declining among mothers of 25 years to around 37 years; then it increased.

(Dey & Raheem, 2016) employed a multilevel multinomial logistic regression model to identify covariates associated with anemia among U5 children in northeastern states of India. The results from the multivariate analysis showed that age of the child, mother's age at marriage, and number of children ever-born have significant effects on child anemia. Specifically Male children were at greater risk of having severe anaemia than female children (OR = 1.488; p = 0.010).

(Yusuf et al., 2019) used data extracted from 2231 U5 children based on the 2011 Bangladesh Demographic and Health Survey (BDHS2011). Anemia prevalence among children is 52.10%. Among these, 48.40% and 53.90% were from urban area and rural areas, respectively. The prevalence of mild, moderate and severe anemia among children was 57.10 %, 41.40% and 1.50%, respectively. the two-level logistic regression showed that children from mothers had anemia were more likely to have anemia than those from non-anemic mothers. Undernourished children were at higher risk to have anemia than normal weight or over nourished children. Children under two years of age had more risk to be anemic compared to children in higher age catagory groups with a decreasing pattern: 2 to 3 years; 3 to 4 years, and 4 to 5 years. Pre-school children from poor families were more likely to be anemic than those from rich or middle class.

A study on the sociocultural risk factors for anemia in U5 children living in Pakistan's Muzaffarabad, Azad Jammu and Kashmir was carried out by (Habib et al., 2020) .the overall prevalancy of anemia is 47.7 % and 52 .3% non anemia. Of the 201 children under five who were anemic, 40.4%, 43.2%, and 16.4% had mild, moderate, and severe anemia, respectively. A binary logistic regression analysis revealed that children who live in urban areas have a decreased risk of anemia (OR 0.21, 95% 0.14-0.33). Overall, research indicated that anemia was more common in children from lower-income families.The study also found that children of educated parents have a decreased risk of developing anemia. There were lower odds of anemia in children of educated mothers than in those of illiterate mothers (OR = 0.66 for 1 to 8 years of schooling and OR = 0.45 for 9 to 12 and 12+ years of schooling).

A systematic review and meta-analysis study by Tadesse et al., (2022) based on a total of 67,647 U5 in 20 African counters showed that the overall pooled prevalence of anemia among U5 children in Africa was 59% (95% CI: 55, 63). The analysis result showed that heterogeneity still exists: in terms of region, the sources of heterogeneity were Ethiopia, Tanzania, Lesotho, Ghana, and Uganda. The study showed that sex of a child, maternal educational status, place of residence, and family size were the pooled determinants of anemia among U5 children in Africa. Being female is a protective against anemia (AOR = 0.71; 95% CI: 0.57, 0.87). Mothers who were unable to read and write were 53% times more likely to have anemic child (AOR = 1.47; 95% CI: 1.31, 1.65). U5 children from rural areas were 20% less likely to be affected by anemia as compared to children from urban areas (AOR = 0.80; 95% CI: 0.67, 0.95). Under-five children from family size of less than five were 7% less likely to be affected by anemia (AOR = 0.93; 95% CI: 0.89, 0.98) those from five and more children.

A study was conducted by Adugna et al., (2023) to determine the variables linked to anemia in Liberian U5 children. The findings of the multivariable logistic regression showed that the age of the child, stunting, drinking water source, and the restroom facility were significant factors associated with anemia in U5 children. Children between the ages of 6 and 23 months and 24 and 42 months had, respectively, 2.4 and 1.8 times higher risks of anemia than children between the ages of 43 and 59 months. Stunted children were more anemic compared to children who were normal. The children from HHs with an inadequate water supply had 1.7 times more than those from HHs with an enhanced water supply. Children from households with unimproved toilet facilities have a have a higher risk of anemia than those from households with improved toilet facilities. The odds of anemia were 0.6 and 0.7 times lower in children than who were living in northwestern and north central regions, respectively, than U5 children from southeastern regions.

Using data from a DHS in Senegal, Malawi, Angola, Khulu & Ramroop,(2020)employed a generalized linear mixed model to find factors that are linked with anemia among U5 children in those three countries. Childs gender, mother's educational attainment, financial status, and nutritional status were important factor of anemia among U5 children; females are more likely to be affected by anemia than their male counterparts. Growing older raised the risk of being exposed to anemia. Compared to children living in urban areas, U5 children who live in rural areas are more likely to suffer from anemia .Those in low economy and middle-class households are more exposed to anemia that in rich HHs.

A study by (Asresie et al., 2020) used the 2016 Ethiopian DHS data. In all, 8,462 children in U5 were included in their research. The multivariable logistic regression analysis identified that wealth index, region, family size, mother's anemia status and age, stunting, underweight, and history of fever of the child were significantly associated with anemia. Children of mothers between the ages of 15 and 24 had a 1.4 times greater risk of getting anemia than children of mothers between the ages of 35 and 49. Children who were stunted had a 1.3-fold increased risk of having anemia compared to their non-stunted counterparts. Compared to children aged 6–23 months and 24–42 months, they had anemic odds that were about four times and two times higher than those aged 43–59 months, respectively. Children who had a history of fever two weeks before data collection were 1.3 times more likely to develop anemia than children who did not have a fever.

Jember et al., (2021) used the DHS data from 2016 to investigate the regional variance and covariate of childhood anemia among U5 children in ethiopia. An investigation of the causes of anemia was conducted using a mixed-effects logistic regression model. The findings were that religion, child's age, wealth index, mother's employment position at the time, maternal anemia, the number of U5 children living with HHs, and the child's fever during the previous two weeks, and stunting are statistically significant variables associated with anemia. The odds of developing anemia in children with family wealth index of poor, middle, richer and richest is lower by 26%, 40%, 35% and 43% in the poorest HHs.

Gebrehawerial & Lemma,(2020) used data from the 2016 Ethiopia DHS to identify covariates associated with anemia status of U5 children. The findings were: compared to male children, female children were less likely to have greater levels of anemia, underweight, feeding children less than 4 times a day, and anemia was positively correlated with insufficient dietary diversity. The odds of anemia were higher for children from anemic mothers than non-anemic mothers. Fever in the last two weeks significant determinant of anemia compared to the children had no fever. Children in U5 living in highly poor neighborhoods were more likely than other children to suffer from anemia.Compared to children from

Tigray (the reference region), children from the Somali, Dire-Dawa, and Harari regions showed greater levels of anemia; Children from Amhara, Benishangul, and SNNPS had 35%, 61%, and 38% lower chances of having higher levels of anemia than children from the Tigray region.

The Results revealed that more than 70% of the interviewed households were from rural areas in all three surveys. Said et al., (2021). The percentage of in rural areas home survey was 82.2%, 75.2%, and 81.0% in the years 2005, 2010, and 2015, respectively. The findings of hierarchical logistic regression analysis revealed that households that, In U5 children with limited dietary diversity, children in low age quartiles, and underweight children, the prevalence rate was highest in the low age quartile. Across all years, the prevalence was higher in the comparatively younger U5 children than in any other age group. Anemia was strongly correlated with child age, with younger children being more vulnerable than older ones.

CHAPTER THREE

3. DATA AND METHODOLOGY

3.1 Description of Study Area

The United Republic of Tanzania is situated in Eastern Africa, sharing borders with Kenya and Uganda to the north, Rwanda, Burundi, and the Democratic Republic of the Congo to the west, and Zambia, Malawi, and Mozambique to the south. Its eastern boundary is defined by the Indian Ocean, It has a coastline of 1424 kilometers along the Indian Ocean and covers a total area of 945,087 square kilometers, including 61,000 square kilometers of inland water. Zanzibar, consisting of Unguja and Pemba islands, has a total surface area of 2,654 square kilometers. The 2012 Population and Housing Census was conducted on August 26, 2012, after the union of Tanganyika and Zanzibar. Tanzania is divided into 31 administrative regions, with 44,928,923 people in total, including 43,625,354 on the mainland and 1,303,569 in Zanzibar (TDHS-MIS, 2022).

3.2 Source of data

The source of data for this study was The 2022 Tanzania Demographic and Health Survey and Malaria Indicator Survey (2022 TDHS-MIS) implemented by the Tanzania National Bureau of Statistics (NBS) and the Office of Chief Government Statistician (OCGS) in collaboration with the Ministries of Health of Tanzania Mainland and Zanzibar.

3.2.1 Study Population

The study population consisted of all children under the age of five residing in Tanzania, who had blood samples taken for analysis using the 2022 TDHS-MIS data. A total of 4149 children with complete hemoglobin measurements from blood samples were included in the study.

3.2.2 Study Design

A population-based cross-sectional survey was used to collect the 2022 TDHS-MIS data.Just the data directly collected from the respondents at a specific time is collected as a variable for many sample units at the same points in time (one time shot).

3.2.3 Sample Design

The sample design for the 2022 TDHS-MIS was completed in two phases with the goal of providing estimates for Zanzibar, Tanzania's mainland, and the entire nation. With 26 regions in Tanzania's mainland and 5 areas in Zanzibar, the sample design enables the estimation of indicators for each of the 31 regions. The 2022 TDHS-MIS followed a stratified two-stage sample design. The 2012 Tanzania Population and Housing Census (2012 PHC) enumeration areas (EAs) were the basis for the sampling points (clusters) that were selected in the first step. The EAs were selected with a probability proportional to their size within each sampling stratum. A total of 629 clusters were selected, there were 418 EAs from rural areas and 211 from urban areas out of the 629 EAs. 26 households from each cluster were systematically chosen for the second stage. The information was available on the Measure DHS website (http://www.dhs program.com), which could be accessed once a request for online authorization declaring our investigation's goal was made.

3.3 Variable in the study

3.3.1 The response variable

The response variable for this study was anemia level measured using the Hemoglobin that analysis was carried out on-site using a battery-operated portable analysis device. Based on the WHO hemoglobin level cut off points classified as: -

None Anemic: if Hgb value greater or equal to 11 mg/dl
Mild Anemic: if Hgb value ranges from 10.0 to 10.9 mg/dl
Moderate and above Anemic: if Hgb value less than or equal to 9.9 mg/dl

3.3.2 Independent covariates

The independent variables related to socio demography, nutrition, clinical conditions, and services were categorized as independent variables. Those variables were selected by earlier studies globally and national level.

No	**Covariates and representations**	**Categories**
	Socio demographic covariates	
1	Sex of child	Female ,male
2	Age of child	6-11 months ,12-23 months , 24-35 months, 36-47 months . 48-59 months
3	Place of residence	Rural and urban
4	Zone	Western, northern, central, southern highlands, southern, south west highlands lake, eastern, Zanzibar
5	Wealth index	Poorest, poorer, middle ,richer, richest

6	Educational level of the mother	No education, primary, secondary , higher
7	Age of mother	15-24 year , 25-35 year, above 35year
8	Toilet facility	Improved , unimproved
9	Source of drinking water	Improved, unimproved
	Clinical factors	
10	Maternal anemic status	No anemic , mild, moderate, severe
11	Diarrhea in a child in the last 2 weeks	No , yes
12	Fever in a child in the last 2 weeks	No, yes
	Nutrition and related factor	
13	Taking of vitamin A in the last 6 months before delivery	No, yes
14	Stunting status of child	Normal .moderate stunting ,sever stunting
15	Wasting status of child	Normal, moderate wasting, sever wasting
16	Underweight of children	Normal, moderate, severe underweight

3.4 Methodology

3.4.1 Descriptive statistics

Descriptive statistical analysis provides summary measures such as number and proportion to describe the sample information. It describes the frequency table created from the collected data.

3.4.2 Inferential statistics

Inference is about drawing or conclusions about population based on data obtained from a limited number of observations from a population. In this study we applied multilevel ordinal logistic regression model to analyze data. And for statistical analysis Stata 16 software ware used .

3.4.2.1 Ordinal logistic regression model

Logistic regression is a common modeling approach when the response variable is binary or polytomous. This model allows one to predict the log odds of outcomes of a response variable from a set of covariates that may be continuous, discrete, categorical, or contain both(Hosmer, Lemeshow, & Sturdivant, 2013). However, based on the formats of the categorical values of the outcome variable, the logistic regression model contains, such as binary logistic regression, multinomial logistic regression, and ordinal logistic regression when a response variable has only two values, more than two values, or more than two values with natural order or rank, respectively (Hosmer & lemeshow, 2000). In this study, the anemia level has three values (non-anemic, mild anemic, moderate and above anemic). Hence, we used the ordinal logistic regression model. Ordinal logistic regression is a statistical analysis method that can be used to model the relationship between an ordinal outcome variable and a set of continuous or categorical explanatory variables.(O'Connell, 2006).

There are different ordinal logit models, such as the adjacent category logit model, cumulative logit model, partial logit model, and generalized logit model (Agresti, 2013; Hosmer & lemeshow, 2000).

3.4.2.2 Testing parallel lines

In ordinal logistic regression models, there is a fundamental assumption that belongs to ordinal odds. This assumption states that parameters shouldn't change for various response variable groups. A test of parallel lines is used to determine whether it is reasonable to assume that the values of the location parameters are constant across categories of the response. Likelihood Ratio Test, Score Test, Wald Chi-Square Test, and other related tests are used to test parallel lines of assumption (Agresti, 2002).

The test of parallelism contains: minus 2 log-likelihood for the constrained model (proportional odds), the model that assumes the planes or surfaces are parallel across the category of the response variable, and minus 2 log-likelihood for the general model that assumes planes or surfaces are separated across the category.

In a way, this assumption states that the categories of the dependent variable are parallel to each other. When the assumption does not hold, it means that there is no parallelity between categories. An alternative may be used, such as the partial proportional odds model (Fullerton & Xu, 2012).

3.4.2.3 Proportional Odds Model (POM)

The proportion Odds Model (POM) belongs to the class of generalized linear models and is frequently used for the analysis of ordinal categorical data. It is a generalization of a binary logistic regression model. When there are more than two ordered categories of outcome variables (Agresti, 2002).Proportional Odds Model is used for modeling the outcome variable that has more than two categories with k collection of independent variables by defining the cumulative probabilities, cumulative odds and cumulative logit for the J-1 categories of the independent, this model simultaneously use all cumulative logits (McCullagh, 1980) and (Hosmer & lemeshow, 2000).

Consider the response variable Y with J categories coded in $\mathrm{j} = 1,2 \ldots J$ and $\mathrm{x} = (\mathrm{x1}, \mathrm{x2} \ldots \mathrm{xk})$ the vector of explanatory variables (co-variables). The J categories of Y conditionally to the values of co-variables occur with probabilities $p_1, p_2, \dots, p_J$, that is $\mathrm{pr}(Y = 1) = p_1, \mathrm{pr}(Y = 2) = p_2, \dots, (Y = j) = p_j$.For Y, the response with the J ordinal categories given that of K explanatory variables the cumulative probability at or below category j can be defined as the sum of the category probabilities

$$\mathrm{pr}(Y \le j \mid X) = \pi_j(X) = p_1 + p_2 + \cdots + p_j, \text{ for } j = 1,2 \dots, J - 1$$

Then the odds of the first $J - 1$ cumulative probabilities are,

$$\mathrm{odd}\, s(\mathrm{pr}(Y \le j)) = \frac{pr(Y \le j)}{1 - pr(Y \le j)} = \frac{\pi_j}{1 - \pi_j}\ , \mathrm{j} = 1,2, \dots, \mathrm{J} - 1. \quad (3.1)$$

Given that the dependent variable's categories seem to be arranged according to the degree of anemia, using the standard ordered logit model, also known as the proportional odds model. The cumulative probabilities reflect the ordering, with $\mathrm{pr}(Y \le 1) \le \mathrm{pr}(Y \le 2) \le \cdots \le \mathrm{pr}(Y \le J) = 1$. Models for cumulative probability do not use the final one.

$\mathrm{pr}(Y \le J)$ since it is necessarily equals to one

The proportional odds model is the log odds of the first $J - 1$ cumulative probabilities as:

$$\mathrm{logit}\big(pr(Y \le j)\big) = \log\left(\frac{pr(Y \le j)}{1 - pr(Y \le j)}\right) = \log\left(\frac{\pi_j}{1 - \pi_j}\right), \mathrm{j} = 1,2, \dots, \mathrm{J} - 1, \quad (3.2)$$

Each cumulative logit uses all the response categories. The relationship between the predictors and response variable is not a linear function in logistic regression instead; the logistic regression function is used, which is the logit transformation of π.

$$\text{where } \pi_j = \frac{\exp\big(\alpha_j - (\beta_1 X_1 + \cdots + \beta_K X_K)\big)}{1 + \exp\big(\alpha_j - (\beta_1 X_1 + \cdots + \beta_K X_K)\big)} \quad (3.3)$$

The log-odds/logit of the cumulative probability (π_j), which is the probability of response Y less than or equal to category j is modeled as a linear function of the predictor variables as:

$$\mathrm{logit}(pr(Y \le j)) = \log\left(\frac{\mathrm{pr}(Y \le j)}{1 - \mathrm{pr}(Y \le j)}\right) = \log\left(\frac{\pi_j}{1 - \pi_j}\right) = \alpha_j - (\beta_1 X_1 + \cdots + \beta_K X_K), 0 \le \pi_j \le 1, j = 1,2, \dots, J - 1 \quad (3.4)$$

where α_j = threshold value, $X_1, X_2 \dots, X_k =$ sets of factors or predictors

In proportional odds model, every single cumulative logit has its own threshold value, coefficients of the equality are independent from dependent variable categories, which are shown as " J " ($j = 1,2, \dots, J - 1$). Thus coefficients of the independent variable will be equal to each other in every cumulative logit model (McCullagh and Nelder, 1989).

This model estimates multiple equations of cumulative probability simultaneously and assumes a linear relationship for each logit and parallel regression line. The direction and intensity of the link between independent factors and the dependent variable's log odds are indicated by the logistic regression coefficients.. However, The interpretation of the regression coefficients differs from that of standard regression coefficients; for the categorical explanatory variable, it is the effect (more or less likely) of the estimated category of the independent variables in comparison to the reference category on the log odds of the dependent variable's categories being at higher levels. This model is also known as the proportionate odds model since the independent variable's influence remains constant across logit functions.

3.4.3 Multilevel ordinal regression model

When evaluating clustered data, the multilevel modeling approach is taken into consideration. Multilevel modeling is used in many fields, including biology, medicine, social science, education, environmental research, and more, and has grown to be essential for categorical responses (Perera et al., 2016).

In the ordinary logistic regression model, it is considered that the responses to each observation are independent. However, this assumption is violated in multilevel studies because, when the data are clustered, responses are correlated with each other within the cluster. When responses are correlated, the standard errors of the ordinary logistic regression model may be biased. In such cases, it is important to adjust the logistic regression model to address the clustered effect in order to obtain unbiased standard errors. . In order to adjust for the cluster effect, model-based goodness of fit testing methods described in(Lipsitz et al., 1996) and (Abeysekera & Sooriyarachchi, 2008) are incorporated within the multilevel model.

The multilevel ordinal logistic regression model involves examining a hierarchical structure in relation to an ordinal dependent variable. It is important to analyze the ordinal categorical dependent variable in relation to one or multiple independent variables. The ordinal categorical dependent variable follow the logistic distribution and nested with higher levels (Khiari & Rejeb, 2015; O'Connell, 2010).

Multilevel proportional odds model is the most common and applicable approach for hierarchical ordinal data and is an extension of the multilevel binary logistic regression model. It is also the generalization of a single-level proportional odds model for nested data structures, or hierarchical. For the K-level ordinal response, there are K-1 probabilities and a cumulative logit model for the response for the child and cluster. (Snijders, 2011).

The clustering of data points within the geographical enumeration area in this study provides a naturally occurring two-level hierarchical structure for the data, meaning that children are nested within the enumeration area.

3.4.3.1 Heterogeneous Proportion

The first logical step in applying multilevel analysis correctly is to test for proportional heterogeneity between geographic enumeration areas. The chi-square test is most frequently used to determine whether the proportions between EA are heterogeneous. Test result provided by,

$$x^2 = \frac{(O_i - E_i)^2}{E_i} \tag{3.5}$$

where, Oi is the observed and Ei is the expected counts in the cell of contingency table.

3.4.3.2 The empty model

The term "population of groups" in the Empty level-2 model for ordinal outcome variable refers to level two units, such as a cluster (enumeration area), and it describes the probability distribution for group-dependent probabilities without accounting for explanatory variables.

Let the $\gamma_{ij}^{(C)}$ be ordered categorical response of i^{th} individual in the j^{th} cluster with C ordered categories coded as $C = 1,2, \ldots, c$.

$$\text{logit}\left(\gamma_{ij}^{(C)}\right) = \text{logit}\left(\frac{\gamma_{ij}^{(C)}}{1-\gamma_{ij}^{(C)}}\right) = \alpha^{(C)} + u_{oj} \tag{3.6}$$

which measure the odds of $\gamma_{ij}^{(C)}$ being in the category less than or equal to C as compared to greater than category C. Where, $\alpha^{(C)}$ cutoff points [intercept] for ordered categories and u_{oj} is random effect of level-2 and assumed normal distribution$N(0, \sigma_{uo}^2)$.

3.4.3.3 Intra-class Correlation Coefficient

It is frequently useful to express the cluster variance for a multilevel model in terms of an intraclass correlation (ICC). ICC represents the proportion of unexplained variance that is at the cluster level, given by:

$$ICC = \frac{\sigma_{uo}^2}{\sigma_{uo}^2 + \sigma_e^2} \tag{3.7}$$

where, σ_{uo}^2 is variance of cluster and σ_e^2 is variance of children level units. For a logistic regression model, binary or ordinal, the level-1 (children) variance, which is not estimated, equals the variance of the standard logistic distribution $\frac{\pi^2}{3} = 3.29$ (Agresti, 2002).

3.4.3.4 Random intercepts model

In a random intercept model, the predictor variables' effect is taken to be constant at all levels, while the variation in the outcome variable is allowed to vary across the grouping variable's levels. The model assumes that each group (enumeration area) has a different baseline level of anemia, represented by a random intercept term. That is:

$$\text{logit}(p_{ij}) = \beta_o + u_{0j} + \beta_1 X_{ij1} \tag{3.8}$$

where, p_{ij} is the probability of anemia for individual I in group j, β_o is a fixed intercept , u_{0j} the random intercept for group j and X_{ij1},…., X_{ijp} are the predictor variables for individual i.

3.4.3.5 Random coefficient Model

In the random intercept model, the intercept is the only random effect, which implies that the groups differ with respect to the average value of the dependent variable [status of anemia]. In the random intercept model, it is assumed that the effects of explanatory variables are the same for each EA.

Now the slope is also allowed to vary across the EA. Also known as a random slope model, this type of model allows the effect of the predictors to vary across cluster. This implies that variations may occur not only in the baseline level but also in the predictors' effects across cluster. In the identical situation, the formula for a random coefficient model would be

$$\text{logit}(\gamma_{\text{cij}}) = \beta_o + \sum_{h=1}^{k} \beta_h X_{hij} + U_{oj} + U_{1j} X_{1ij} \tag{3.9}$$

There are two random group effects U_{oj} , which are the random intercepts and U_{1j}, which are the random coefficients. It is assumed that the mean of both are zero. Their variances are represented by$\delta^2{}_0, \delta^2{}_1$, respectively and their covariance is $\delta^2{}_{01}$. The model for a single explanatory variable discussed above could be extended by including more variables that have random effects. Assuming that there are$X_1, X_2, \dots, X_k$, level-one explanatory variables, we consider the model where all X variables have varying slopes and a random intercept.

$$\text{logit}(\gamma_{cij}) = \beta_o + \sum_{h=1}^{k} \beta_h X_{hij} + U_{oj} + \sum_{h=1}^{k} U_{hj} X_{hij} \quad (3.10)$$

The first part of equation $\beta_o + \sum_{h=1}^{k} \beta_h X_{hij}$ is called the fixed part of the model and the second part, $U_{oj} + \sum_{h=1}^{k} U_{hj} X_{hij}$ is called the random part.

3.4.4 Methods of Parameter Estimation

The maximal likelihood estimation (MLE) approach is the most often used technique for estimating a logistic regression model's parameter. By using the maximum likelihood estimation method, values for the unknown parameters can be estimated that maximize the probability of obtaining the observed data set. In logistic regression, the model coefficients are calculated through the maximum likelihood technique, where the likelihood equations are nonlinear explicit functions of unknown parameters. The ordinal logistic regression model is applied to the observed responses by employing the maximum likelihood method.

In the model$P(Y \leq y_j \mid x)$, fitting a single parsimonious model that incorporates all j-1 cumulative logits differs from fitting individual logit models for each j. For estimating the parameters of the model define the binary indicator of the response variable for each observation or subject i. This shows that j categories of response collapsed in to binary outcome. Again let $(Y_{i1}, \ldots, Y_{ij})$ be binary indicators of the response for subject i. The likelihood function L is viewed as a function of β and α_i parameters. Parameters are estimated by maximizing the likelihood,or more usually, by maximizing the logarithm of the likelihood. Therefore, the likelihood function is defined as (Agresti, 2002).

$$l = \prod_{i=1}^{n} \left[\prod_{j=1}^{J} \pi_j(X_i)^{y_{ij}} \right]$$

$$\text{l} = \prod_{i=1}^{n} \left[\pi_1(X_i)^{y_{il}} * \pi_2(X_i)^{y_{i2}} * \pi_3(X_i)^{y_{i3}} * \cdots * \pi_J(X_i)^{y_{i,J}} \right]$$

$$l(\alpha, \beta) = \prod_{i=1}^{n} \left[\prod_{j=1}^{J} \left[P\left(\left(Y \leq y_j \mid x \right) \right) - P\left(\left(Y \leq y_{j-1} \mid x \right) \right) \right]^{y_{ij}} \right]$$

$$l(\alpha, \beta) = \prod_{i=1}^{n} \left[\prod_{j=1}^{J} \left[\frac{\exp^{(\alpha_j + z'\beta)}}{1 + \exp^{(\alpha_j + x'\beta)}} - \frac{\exp^{(\alpha_{j-1} + x'\beta)}}{1 + \exp^{(\alpha_{j-1} + z'\beta)}} \right]^{y_{ij}} \right]$$

$$l(\alpha, \beta) = \prod_{i=1}^{n} \left[\pi_1(X_i)^{y_{i1}} * \pi(X_i)^{y_{i2}} * \pi_3(X_i)^{y_{i3}} * \cdots * \pi_J(X_i)^{y_{ij}} \right]$$

$$L(\alpha, \beta) = \log(l(\alpha, \beta))$$

$$\log(l(\alpha, \beta)) = \prod_{i=1}^{n} \log\left(\left[\pi_1(X_i)^{y_{i1}} * \pi(X_i)^{y_{i2}} * \pi_3(X_i)^{y_{i3}} * \cdots * \pi_J(X_i)^{y_{ij}} \right] \right)$$

$$= \prod_{i=1}^{n} \left[y_{i1} \log \pi_1(X_i) + y_{i2} \log \pi_2(X_i) * \cdots * y_{ij} \log \pi_J(X_i) \right]$$

$$L(\alpha, \beta) = \prod_{i=1}^{n} \left[y_{i1} \log \pi_1(X_i) + y_{i2} \log \pi_2(X_i) + \cdots + y_{ij} \log \pi_J(X_i) \right] \qquad (3.11)$$

In general, Maximum Likelihood Estimation produces values for the unknown parameters that most closely align with the observed and expected probability values. A fisher scoring algorithm for MLE fitting of all cumulative link models was presented by McCullagh, (1980) Because of this, it is frequently employed as a effective technique to provide ML estimates for the parameters of ordinal logistic regression.

3.4.5 Model selection

Two widely utilized criteria for this purpose are the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC).(Akaike, 1974) and(Schwarz, 1978) ,respectively .The aim of AIC is to find a model that maximizes the likelihood of the data while taking into account the number of parameters used. The formula for AIC is.

$$.AIC = -2 * log(L) + 2 * p \quad 3.12$$

In the formula, L indicates the maximal likelihood of the model, which measures how well the model fits the data. The term p denotes the number of parameters in the model, including the intercept and any additional predictors.

The formula for BIC is : $BIC = -2 * log(L) + p * log(n)$ in the formula, the terms log(L) and p have the same meaning as in AIC. Additionally, the term log(n) represents the logarithm of the sample size (n).

3.4.6 Goodness of fit the model

After constructing the model for our data, it is important to investigate how effective the model is in representing the outcome variable. Pearson Chi-square, the Likelihood Ratio Test (Deviance), and the Wald test are the most commonly used measures of goodness of fit for categorical data. Among them, some are briefly discussed in the following section.

The likelihood ratio test is widely used to evaluate the overall goodness of fit of a logistic regression model. LRT is used to test the significance of a number of explanatory variables, and it is used to test the ratio of the maximized value of the likelihood function for the full model (L_{full}) over the maximized value of the likelihood function for the reduced model (L_{red}).

$$\text{LRT} = -2[L_{\text{red}} - L_{full}] \quad (3.13)$$

where, the log-likelihood function of the reduced model and L_{full} the log-likelihood function of the full model.

Wald Test: The Wald statistic is an alternative test that is commonly used to test the significance of individual logistic regression coefficients for each independent variable. The Wald test is one of a number of ways of testing whether the parameters associated with a group of explanatory variables are zero. If the Wald test shows significance for a specific explanatory variable, it indicates that the parameters linked to these variables are not zero, suggesting that these variables should be retained in the model. Conversely, if the Wald test is not significant, it implies that the explanatory variables can be excluded from the model.

Wald statistic is given by $W = \frac{\hat{\beta}^2}{se(\hat{(}2)}$. When the sample size is large, the statistic has an approximate chi-square distribution with one degree of freedom.

3.4.7 Variable selection

Variable selection means choosing from among many variables which to include in a particular model, that is, to select appropriate variables from a complete list of variables by removing those that are irrelevant or redundant. If we have p number of explanatory variables we do not believe that all of them at truly related to the response and we are interested in choosing which is the most important. The most automated variable selection such as forward step wise selection and back ward step wise selection.

Backward stepwise selection. It is a variable selection approach which starts from a model that comprises all variables under consideration, then starts deleting the least significance variables one after the other until a predetermined stopping rule is reached or until no variable is left in the model

CHAPTER FOUR

4. RESULTS

This chapter's goal is to describe and provide an analysis. Of the effect of major socio-demographic, health-related, and nutritional factor characteristics on determinants of anemia among U5 in Tanzania. In this study, descriptive analysis, POM, and multilevel logistic regression analysis were employed to identify the risk factors and examine the variation of anemia among clusters for U5 using Stata 16.

4.1 Descriptive statistics

Table 4. 1 Anaemia level among U5 years

Anemia level	Freq.	Percent	Cum.
Mild	1074	25.89	25.89
Moderate and above	1411	34	59.89
Not anemic	1664	40.11	100
Total	4149	100	

The data analysis in this study is based on a total of 4149 children under the age of 5. The overall prevalence of anemia among these children, based on capillary blood samples, revealed that 40.11 percent were not anemic, 25.89 percent had mild anemia, and 34 percent had moderate and above anemic.

Table 4. 2 anaemia level and Socio demographic Predictors

		Anemia level		
Socio demographic Predictors	Categories	Non anemic	Mild	Moderate and above
Sex of child	Male	827 (39.20%)	508(24.08%)	774(36.72%)
	Female	836(41.02%)	566(27.76%)	637(31.22%)
	Total	1664(40.11%)	1074(25.89%)	1411(34%)
Place of residence	Urban	443(39.36%)	291(25.86%)	392(34.78%)
	Rural	1221(40.37%)	783(25.90%)	1019(33.73%)
	Total	1664 (40.11%)	1074 (25.89%)	1411(34%)
Zone	Western	166(41.76%)	119(29.80%)	113(28.44%)
	Northern	177(44.29%)	87(21.62%)	136(34.09%)

	Central	236(52.98%)	111(24.87%)	99(22.15%)
	Southern highlands	117(5456%)	47(21.99%)	50(23.45%)
	Southern	51(32.33%)	35(22.12%)	72(45.55%)
	south west highlands	160(39.71%)	104(25.65%)	140(34.64%)
	Lake	563(38.07%)	372(25.16%)	544(36.77%)
	Eastern	150(28.59%)	162(30.89%)	213(40.52%)
	Zanzibar	40(33.45%)	36(30.84%)	42(35.72%)
	Total	1664(40.11%)	1074(25.89%)	1411(34%)
child's age in months	6-11 months	123(24.91%)	123(24.88%)	247(50.20%)
	12-23 months	295(28.22%)	276(26.42%)	474(45.35%)
	24-35 months	335(38.27%)	260(29.73%)	280(31.99%)
	36-47 months	420(48.61%)	211(24.42%)	233(26.96%)
	48-59 months	491(56.30%)	204(23.39%)	177(20.31%)
	Total	1664(40.11%)	1074(25.89%)	1411(34%)
mother's age	15-19	40(20.91%)	42(21.92%)	108(57.17%)
	20-24	332(35.86%)	245(26.46%)	349(37.68%)
	25-29	442(41.34%)	273(25.55%)	354(33.11%)
	30-34	370(43.58%)	212(24.95%)	267(31.47%)
	35-39	294(43.07%)	179(26.16%)	212(30.97%)
	40-44	148(44.44%)	96(28.89%)	89(26.67%)
	45-49	38 (39.32%)	27(27.77%)	32(32.79%)
	Total	1664(40.11%)	1074(25.89%)	1411(34%)
mother's education level	No education	326(36.21%)	223(24.79%)	351(39%)
	Primary	1007(41.55%)	623(25.70%)	793(32.76%)
	secondary	312(39.51%)	224(28.39%)	254(32.10%)
	Higher	19(52.48%)	4(11.01%)	13(36.51%)

	Total	1664(40.11%)	1074(25.89%)	1411(34%)
Wealth index	Poorest	348(37.17%)	230(24.59)	358(38.24%)
	Poorer	352(40.03%)	243 (27.69%)	283(32.27%)
	Middle	328(39.35%)	244(29.25%)	262(31.40%)
	Richer	318(40.69%)	188(24.11%)	275(35.20%)
	Richest	318(44.19%)	169(23.41%)	233(32.40%)
	Total	1664(40.11%)	1074(25.89%)	1411(34%)
Source drinking water	Improved	1144(40.39%)	727(25.67%)	962(33.94%)
	Unimproved	520(39.47%)	347(26.35%)	449(34.18%)
	Total	1664(40.11%)	1074(25.89%)	1411(34%)
Toilet	Improved	1187(41.34%)	736(25.65%)	947(33.01%)
	Unimproved	477(37.30%)	338(26.42%)	464(36.28%)
	Total	1664(40.11%)	1074(25.89%)	1411(34%)

The data from the TDHS-MIS 2022 presents the prevalence of anemia among children. Among the total number of male children, 827(39.20 %) are not anemic, 508(24.08%) have mild anemia, and 775(36.72%) have moderate and above anemia. In comparison, among female children, 836(41.02%) are not anemic, 566(27.76 %) have mild anemia, and 637(31.22 %) have moderate and above anemia. These percentages indicate that a higher proportion of female children are not anemic and have mild anemia compared to male children, while a higher percentage of male children have moderate anemia.

The children came from Urban show a higher rate of moderate anemia (34.78%) than rural children (33.73%). The tabulated data presents the prevalence of anemia among children across different zones, with anemia level .The Central and Southern Highlands zones exhibit the highest percentage of non-anemic individuals, with 52.18% and 54.56% respectively, conversely, the southern zone has the highest prevalence of moderate and above anemia at 45.55%**,** suggesting a significant public health concern in this area. and the children who came from Zanzibar zone is the highest prevalence of middle anemia compare with the other zone.

The tabulation of anemia levels across different child age groups reveals a clear trend as children age, the prevalence of anemia decreases when the children age increase. The youngest group, aged 6-11 months**,** has the highest rate of moderate and above anemia at 50.20 %**,** which is a critical period for infant development and thus may require additional nutritional support. In contrast, children aged 48-59 months show the lowest rate of moderate and above anemia at 20.31 %. The younger mothers, aged 15-19 and 20-24, have the highest prevalence of moderate anemia at 57.17 %, 37.68 respectively from moderate and above. This suggests that younger mothers might be at a higher risk and may benefit from targeted nutritional interventions. As the age of mother's increases, the percentage of moderate anemia decreases, with mothers aged 40-44 and showing lower rates of moderate and above anemia at 26.67%. This indicates that anemia levels improve with increasing maternal age.

Among children with mothers who had no education, 36.21 % were not anemic, 24.79 % had mild anemia, and 39 % had moderate and above anemia. For children whose mothers had primary education, 41.55% were not anemic, 25.70 % had mild anemia, and 32.76 % had moderate and above anemia. Similarly, for children with mothers having secondary education 39.51 % were not anemic, 28.39 % had mild anemia, and 32.10 % had moderate and above anemia. and the children mother who higher education level 52.48 % were not anemic,11.01 % mild and 36.51 % were moderate and above anemic. Children of mothers with higher education levels tend to have lower rates of anemia, with a higher percentage being not anemic and lower percentages experiencing middle anemic. This highlights the importance of maternal education in influencing the health outcomes of U5 children, including the prevalence of anemia.

When examining the wealth index, there is a discernible pattern in the prevalence of anemia across different wealth categories. Among children from the poorest households is the highest prevalence of moderate and above anemia (38.24%) compared to children from the poorer (32.27%), middle (31.40%), richer (35.20%), and richest (32.40%) wealth categories.

Children using improved water sources have a slightly higher percentage (40.39%) with no anemia compared to those with unimproved sources (39.47%). children with moderate and above anemia levels showed a slightly higher percentage with unimproved water sources (34.18%) versus improved sources (33.94). Likewise, when examining the type of toilet sanitation facilities, there appears to be a notable relationship with the prevalence of anemia across different categories. Among children with access to improved toilet facilities, 33.01% had moderate and severe anemia, while this percentage was slightly higher for those using unimproved sanitation facilities (36.28%).

Table 4. 3 anaemia level and Clinical Predictors

		Anemia level		
Clinical Predictors	Categories	Non anemic	Mild	Moderate and above
mother's anemia level	Not anemic	1159(45.44%)	666(26.11%)	726(28.45%)
	Mild	259(35.51%)	184(25.32%)	285(39.17%)
	Moderate	208(27.25%)	195(25.53%)	360(47.22%)
	Severe	38(35.47%)	29(27.06%)	40(37.44%)
	Total	1664(4011%)	1074(25.89%)	1411(34%)
had diarrhea recently	No	1569(41.69%)	974(25.69%)	1248(32.92%)
	yes, last two weeks	95(26.38%)	100(28.02%)	163(45.61%)
	Total	1664(40.11%)	1074(25.89%)	1411(34%)
had fever in last two weeks	No	1481 (40.54%)	952(12.06%)	1220(33.30%)
	Yes	183(36.82%)	122(24.58%)	191(38.60%)
	Total	1664(40.11%)	1074(25.89%)	1411(34%)

Children of mothers who were non anemic status from 2554 is 45.44 % the children is not anemic, 26.11% having mild anemia, and 28.45 % experiencing moderate and above anemia, the children mother who had mild anemic from 728 is 35.51 %children is not anemic,25.32 % mild anemic and 59.17 % is moderate and above anemia ,the other side, children whose mother anemia status is moderate 27.25 % non-anemic,25.53 % middle and 47.22 % is moderate and above, the mother anemia status is severe 35.47 %children is non anemiac,27.06 % is mild anemic and 40% children is moderate and above anemic .

The data indicates that children who did not experience diarrhea in the past two weeks had a lower prevalence of moderate and above anemia (32.92%) compared to those who had a diarrhea recently (45.61%). Similarly, children without a recent fever had a lower prevalence of moderate anemia (33.30%) compared to those who had a fever (38,60%). These findings suggest a possible association between recent illness, such as diarrhea and fever, and a higher prevalence of moderate anemia among children.

Table 4. 4 anaemia level and Nutrition and related factors

		Anemia level		
Nutrition and related factors	Categories	Non anemic	Mild	Moderate and above
Vitamin A in last 6 months	No	729(39.78%)	488(26.64%)	614(33.58%)
	Yes	935(40.34%)	586(25.29%)	797(34.36%)
	Total	1664(40.11%)	1074(25.89%)	1411(34%)
Stunting l	Normal	1184(41.40%)	779(27.23%)	897(31.37%)
	Moderate	353 (23.52%)	222(23.73%)	370(39.18%)
	Severe	127(36.94%)	73(21.22%)	144(41.84%)
	Total	1664(40.11%)	1074(25.89%)	1411(34%)
Wasting l	Normal	1626(40.50%)	1033(25.75%)	1354(33.74%)
	Moderate	27(26.48%)	34(32.77%)	42(40.75%)
	Severe	11 (32.77%)	7(20.71%)	15(46.51%)
	Total	1664(40.11%)	1074(25.89%)	1411(34%)
Underweight	Normal	1488(40.70%)	959(26.42%)	1209(33.07%)
	Moderate	145(35.65%)	94(23.15%)	167(41.20%)
	Severe	31(35.56%)	21(24.01%)	35(40.43%)
	Total	1664(40.11%)	1074(25.89%)	1411(34%)

Children who did not receive vitamin A supplementation had a prevalence of moderate and above anemia at 33.58 %, while those who received vitamin A had a slightly higher prevalence at 34.36%.

The data reveals a notable association between child stunting levels and the prevalence of anemia. Among children with normal stunting levels, 31.37% had moderate and above anemia, while for those with moderate stunting, the prevalence of moderate and above anemia was higher at 39.18%. Additionally, children with severe stunting exhibited a prevalence of 41.84% for moderate and above anemia.

Among under-five children, 1626 (40.50%) were non-anemic, 1033 (25.75%) had mild anemia, and 1354 (33.74%) had moderate and above anemia for those with normal wasting levels. For children with moderate wasting levels, 27 (26.48%) were non-anemic, 34 (32.77%) had mild anemia, and 42 (40.75%) had moderate and above anemia. In the case of children with severe wasting levels, 11 (32.77%) were non-anemic, 7 (20.71%) had mild anemia, and 15 (46.51%) had moderate and above anemia. the other side children with normal weight have the highest prevalence of no anemia (40.70%).the children who is moderate underweight the highest percentage of mild (41.20%) compare to normal underweight (33.07%)

4.2 Results of multilevel logistic regression model

The data have a level two hierarchical structure with 4149 under-five children at level one, nested within cluster at level two. This is based on the idea that there might be differences in under-five children anemia between cluster that are not captured by the explanatory variables and hence might be regarded as unexplained variability within the set of all cluster.

4.2.1 Test of heterogeneity

Before beginning a multilevel analysis, it is necessary to test for heterogeneity in Tanzanian clusters related to U5 children's anemia. The test resulted in a Pearson chi-square value of 1600 with a p-value is equal to (0.000), indicating statistical significance at the 0.05 level. This suggests that there is significant heterogeneity among clusters concerning anemia in under-five children in Tanzania.

4.2.2 Multilevel model selection

We can select the appropriate model by using both Akaike information criterion (AIC) and Bayesian information criterion (BIC).

Table 4. 5 Model comparisons

Model	AIC(small is better)	BIC(small is better)
empty model	8693.137	8711.327
Random coefficient model	8448.206	8700.77
Random intercept model	8445.227	8678.849

Both the AIC and BIC values for empty model and random coefficient (slope) model are greater than random intercept and fixed effect model. In all the two criteria as shown Table 4.4,random intercept and the fixed effect model perform better than empty model and random coefficient(slope). Hence, the smaller value is better model (random intercept and fixed effect model).

This implies that the intercept varied across cluster, i.e., the variations of under-five children anemia among Tanzanian enumeration areas are non-zero. The random intercept and fixed slope model provides a better fit for predicting anemia variation among under-five children across Tanzanian enumeration areas.

Goodness of Fit the Model

Goodness of test	LRchi2	Df	Prob>chi2
Likelihood-ratio test	532.01	35	0.000

The likelihood ratio test is used to evaluate the model's overall goodness of fit. The difference between the models is indicated by the 35-degree-of-freedom LR chi-squared value (532.01). The statistical significance is confirmed by the p-value (0.0000). Based on the obtained results, we may reject the null hypothesis and interpret that at least one independent variable has significant effects on the outcome variable.

4.2.3 The empty model

Table 4. 6 Results of the empty model

Anemia_level	Coef.	Std. Err.	P>z	[95% Conf. Interval]	
Fixed effect					
/cut1	-.4362419	.0423722		-.5192899	-.3531939
/cut2	.7142814	.0434223		.6404139	.7993876
Random effect					
EA					
var(_cons)	0.369898	0.0608736		.2679174	.5106967

First, we fitted an intercept-only model, which is an empty model without any explanatory variables, to predict the likelihood of anemia in children under the age of five. A model in which no predictor (independent) variables are included and only the intercept differs across level EA is the most basic non-trivial specification of the hierarchical linear model. The empty model with random effect also assist to calculate the between cluster variations by the help of ICC

$$ICC = \frac{\sigma^2 u_0}{\sigma^2 u_0 + 3.29} = \frac{0.369898}{0.369898 + 3.29} = 0.1011$$

To assess the proportion of variation in anemia prevalence among under-five children attributed to cluster-level factors, we need to consider the random intercept in the multilevel logistic regression model. We can use the intra-correlation coefficient. The intra-correlation coefficient of empty model is ICC=0.1011, it implies 10.11% of the total variability in prevalence of anemia among U5 children between cluster level, The remaining variability (89.89%) arises from variations within the children themselves and other unknown factors.

4.4.4 Random Intercept and Fixed Effect Model

In the assessment of how explanatory variables impact the anemia status among children under five, there was a decrease of 0.1615 in the variance component (0.369898 - 0.2084025 = 0.1615) compared to the empty model. compare the variance of the random effect of the intercept and the fixed effect model decreased relative to the random effect of the intercept in the empty model. This decrease in the variance of the random effects of the intercept is attributed to the significant role played by the included fixed explanatory variables in explaining the variability in children's anemia status among clusters.

Table 4. 7 Results random intercept and fixed effect model

Variable	Odds Ratio	Std. Err.	P>z	[95% Conf. Interval]	
Children age (Ref <+11 month)					
12-23 months	.8115359	.0891023	0.057	.6544105	1.006387
24-35 months	.4965346	.0564441	0.000	.3973638	.6204556
36-47 months	.3516468	.0406857	0.000	.2802994	.441155
48-59 months	.2495936	.02939	0.000	.1981542	.3143863
Mother's anemia level(Ref=not anemic)					
mild	1.463414	.1241847	0.000	1.23918	1.728223
moderate	2.221171	.1886193	0.000	1.880611	2.623403
severe	1.376298	.270794	0.105	.9359099	2.023909
zone(Ref=western)					
northern	1.009364	.1713448	0.956	.7236892	1.407809
central	.6795332	.1133306	0.021	.4900605	.942262
southern highlands	.7419559	.1501487	0.140	.4990241	1.10315
southern	1.853974	.3961697	0.004	1.219588	2.818345
south west highlands	1.278092	.217569	0.149	.9155085	1.784275
lake	1.331085	.1840233	0.039	1.015141	1.745359
eastern	1.885643	.3023284	0.000	1.377161	2.581868
Zanzibar	1.470705	.3684297	0.124	.9000969	2.403046
Stunting(Ref=normal)					
Moderate Stunting	1.32899	.1026084	0.000	1.142358	1.546113
Severe Stunting	1.348106	.1585284	0.011	1.070603	1.697539
Mother education level(Ref=no education)					

primary	.8029366	.0672214	0.009	.6814266	.9461138
secondary	.822398	.0932325	0.085	.6585431	1.027022
higher	.8844229	.3288405	0.741	.4267482	1.83294
Mother age (REF=15-19)					
20-24	.6048324	.1006131	0.003	.4365547	.8379756
25-29	.5405178	.089999	0.000	.3900144	.7490994
30-34	.4856434	.082049	0.000	.3487449	.6762809
35-39	.508568	.0883425	0.000	.3618164	.7148417
40-44	.4968859	.094236	0.000	.3426286	.7205925
45-49	.6386726	.162253	0.078	.3881797	1.050809
Child sex (Ref=male)					
female	.8720565	.0543456	0.028	.7717891	.9853501
toilet facility (REF=improved)					
Unimproved	1.097935	.088224	0.245	.937948	1.285211
Diarrhea recently (Ref=no)					
yes, last two weeks	1.301406	.1478445	0.020	1.041629	1.62597
Wealth index(REF=poorest)					
poorer	.8635748	.0854071	0.138	.7114045	1.048295
middle	.8907326	.0916318	0.261	.7280853	1.089714
richer	.9616094	.1043284	0.718	.7774077	1.189456
richest	.8056307	.0967034	0.072	.6367406	1.019317
Wasting (Ref=normal)					
Moderate Wasting	1.262178	.2482314	0.236	.8584515	1.855777
Severe Wasting	1.454978	.511813	0.286	.7301832	2.899218
/cut1	-1.56371	.2351161		-2.024529	-1.10289

/cut2	-.3155608	.2338763		-.7739499	.1428283
Random part					
EA					
var(_cons)	.2084025	.0510322		.1289634	.3367745

In the Random Intercept and Fixed Effect Model results, the calculated intra-class correlation coefficient is, $ICC = \frac{\sigma^2 u_0}{\sigma^2 u_0 + 3.29} = \frac{0.2084025}{0.2084025+3.29}$ =0.0596.This value represents 5.96% of the variance in the prevalence of anemia among children under five years old that can be attributed to differences between clusters.

The constant term's between-cluster [level-2] variance for the prevalence of anemia in children U5 was estimated to be 0.2084025, which was 0.1615 less than the empty model. This suggests that those significant factors contributed to the variations in the prevalence of anemia in under-five children across clusters.

From Table 4.7 The results of multilevel ordinal logistic regression analysis showed that there were differences in the prevalence of anemia among U5 children by cluster. In addition, child's age, mother anemia, zone, stunting level, mother education level, mother age, sex of children, had diarrhea recently also found to be significant ($p < 0.05$) determinants of anemia among U5 children aged, whereas toilet facility, wealth index, wasting were insignificant ($p > 0.05$) predictors of variation in prevalence of anemia among children aged U5 years.

In this study age of child had also significant effect to anemia level among U5 children; children in the age group of 24-35 months had odds 0.497 times lower than children aged less than or equal to 11 months of being in a higher anemia category. This means that children aged 24-35 months had a 50.3% lower odds of being in a higher anemia category compared to children aged less than or equal to 11 months. Similarly, children in the age group of 36-47 months had odds 0.352 times lower, indicating a 64.8% lower odds, and children in the age group of 48-59 months had odds 0.25 times lower, indicating a 75% lower odds of being in a higher anemia category compared to children aged less than or equal to 11 months.

This study also showed that a mother's anemia level was significantly factor with a child's anemia level. The estimated odds ratio (OR=1.463) , that the odds of being (mild or moderate and above anemia) for a child from a mother who had milled anemia were 1.463 times the odds of a child from a mother who had not anemia, controlling for other variables in the model and random effect at level two (OR=1.463, CI: 1.239,1.728). Similarly, the estimated odds ratio (OR = 2.221) implied that the odds of being in the in the higher category for a child from a mother who had moderate anemia were 2.221 times the odds of a child from a mother who had no anemia, holding other variables constant.

Geographical area such as zone has a significant effect on anemia among U children .The odds ratio (OR) for children who live in the central zone were 0.68 this suggest that holding other variables constant, the odds being in higher category of anemia are about 32 % lower for children who live in the central zone compared children who live in the western. On the other hand, the odds ratios for children living in the southern, eastern, and lake zones were 1.854, 1.885, and 1.33, respectively. This means that children living in the southern zone have 85.4% higher odds of being in a higher anemia category compared to those in the western zone. Similarly, children in the eastern zone have about 88.5% higher odds, and children in the lake zone have about 33% higher odds of being in a higher anemia category compared to children in the western zone.

The odds of being in a higher category for a child who is moderately stunted are 32.9% higher than the odds for a child who is not stunted, controlling for other variables (OR = 1.329, CI: 1.142, 1.546).Similarly, the odds of being in a higher category for a child with severe stunting are 34.8% higher than the odds for a child with normal stunting, when other variables are held constant.

In this study mother education is determinate factor of anemia among of U5 children. The odds ratio(OR) for a child from a mother who had a primary education level : 0.803 this suggest that holding other variables constant, the odds being in higher category of anemia are about 19.7 %lower for child from a mother who had a primary education level compared to a child from a non-educated mother.

In this study, mother’s age is a significant determinant of anemia among under-five children. Children from mother age groups 20–24, 25–29, 30–34, 35–39, and 40–44 had odds ratios of 0.605, 0.541, 0.486, 0.506, 0.497, and 0.639 times lower, respectively, compared to children from the mother age 15–19 group, when other variables were held constant. This means that children of mothers aged 20–24, 25–29, 30–34, 35–39, and 40–44 have 39.5%, 45.9%, 51.4%, 49.4%, and

36.1% lower odds of being in a higher anemia category, respectively, than children of mothers aged 15–19.

Sex of child is also a significant determinant of anemia among under-five year's children. The odds ratio(OR) for female :0.872 this suggest that holding other variables constant ,the odds being in higher category of anemia are about 12.8 % lower for females compared males.

This study revealed that in the last two weeks before the date of survey date diarrhea had a significant effect to anemia among under-five children. ; OR for having diarrhea in the last two weeks: 1.301 ,Children who had diarrhea in the last two weeks have 30.1% higher odds of being in a higher anemia category, compared to those who did not have diarrhea, given that other variables held constant.

4.3 Proportional Odds Assumption

Score test for the proportional odds assumption for STATA procedure produced insignificant chi-square value of 48.83 with 35 degree of freedom [p-value=0.060] indicating that a parallel lines assumption has appropriate for the evidence the data. Since, we cannot reject the null hypothesis at 5% significant level. Thus, the model satisfies the proportional odds assumption. Hence, it is not necessary to go for another model.

4.3 Test of multi-collinearity

Table 4. 8 Result of Multicollinerty

Variable	VIF	Tolerance
Child age	1.07	0.9319
Mother anemia level	1.02	0.9790
Zone	1.07	0.9311
Stunting	1.03	0.9694
Mother education level	1.27	0.7870
Mother age	1.07	0.9386
Child sex	1.01	0.9939
Toilet	1.24	0.8088

Diarrhea recently	1.04	0.9574
Wealth index	1.35	0.7381
Wasting	1.01	0.9938
Mean of VIF	1.11	

One of the assumptions in logistic regression is among explanatory variables should not be severe multicollinearity . Multicollinearity occurs when two or more explanatory variables have a high degree of correlation with one another, such that they cannot obtain unique or independent information in the regression model. If there is a high enough degree of correlation between variables, it may be difficult to fit and interpret the model. The researcher checked the existence or absence of multi-colinearity among the independent variables by using a variance inflation factor (VIF). Based on the results displayed in the VIF values of less than 10. As a result, there is no multi-collinearity problem among these variables.

.

CHAPTER FIVE

5. DISCUSSION

5.1 Discussion

The main purpose of the research was to identify determinant factors of anemia among U5 children in Tanzania by applying multilevel ordinal logistic regression based on 2022 TDHS-MIS. Accordingly, descriptive analysis and multilevel model techniques were used to analyses the data. The results showed that, the prevalence of anemia among U5 children in Tanzania. From the total sample 4149, 59.89 % are anemic among this children 25.89% exhibit mild anemia and 34 % moderate and above anemia .This study is similarly presented in TDHS-MIS 2022 Final Report and in line with the pooled prevalence of anemia among U5 children in Africa by Tadesse et al., (2022).

Covariates which were included in this study were sex of child, age of child, place of residence, zone, wealth index, maternal education, age of mother, toilet facility, source of drinking water, maternal anemia status, diarrhea in a child in last 2 weeks, fever in a child in the last 2 weeks, taking of vitamin A in the last 6 months before delivery, stunting status of child, wasting status of child and underweight of children. Variables such as residence place, source of drinking water, recent fever in a child, and maternal intake of vitamin A in the six months before delivery were excluded from the final multilevel model using stepwise forward variable selection with a significance threshold of p >= 0.25.

Results based on the empty model suggest that children status of anemia differed across cluster. In addition to the empty model, random intercept and fixed effect model as well as a model with random coefficients (slopes), were considered. Based on, AIC and BIC), the two-level random intercept and fixed effect model provided a good fit for the data. The overall variance constant term random intercept and fixed effect models was found to be statistically significant implying that status of anemia differ across EA/cluster.

This study identify child age significantly factor with prevalence of anemia among under five children. As age of child increases the risk of being anemia decreased. This finding seemed to be consistent with other studies According to the findings of study done (Yusuf et al., 2019) children under 2 years of age had higher risk to be anemic compared to children in older age groups with a decreasing pattern. Agreeing with the finding study by (Adugna et al., 2023) The odds of having anemia in children aged 6–23 and 24–42 months were 2.4 times, and 1.8 times higher than those in the age group 43–59 months, respectively.in this study similarly to the finding by (Chungkham et al., 2021) finding shows The age of children increases, its effect on childhood anemia decreases, which indicates, older children are less likely to have the risk of childhood anemia in line finding (Said et al., 2021). although according to (Khulu & Ramroop, 2020)The likelihood of being exposed to anemia increased with increasing age.

This study showed significant influence of maternal anemia on the anemia status of children. This aligns with previous research conducted by Jember et al. (2021) and Asresie et al. (2020).Specifically, children with mild and moderate levels of anemia have 1.477 and 2.308 times higher odds of anemia compared to children whose mothers do not have anemia. These results similar observations in Ethiopia, as reported by (Gebrehawerial & Lemma, 2020), The odds of anemia were higher for children from anemic mothers than non-anemic mothers.

The study highlights the significance of the zone (geographical region) as a factor influencing anemia prevalence among children under five. This finding aligns with previous research conducted by (Adugna et al., 2023) and (Asresie et al., 2020). Specifically, the odds of moderate and severe anemia are 0.644 times lower for children living in the central zone compared to those in the western zone. Conversely, children residing in the southern and eastern zones have higher odds of moderate and severe anemia (with odds ratios of 1.861 and 1.844, respectively) compared to the western zone. These findings similarly reported by (Gebrehawerial & Lemma, 2020) reported that children from Somali, Dire-Dawa, and Harari regions exhibit higher anemia levels than those from Tigray (the reference region). Additionally, children from Amhara, Benishangul, and SNNPS regions have 35%, 61%, and 38% lower odds of higher anemia levels, respectively, compared to Tigray.

The findings of this study also show that a stunted children are significantly more vulnerable to anemia compared to those with normal growth. This finding aligns with other studies (Adugna et al., 2023) , (Asresie et al., 2020) , (Jember et al., 2021).

The finding of this study show mother's education primary level is less likely be affected the children by anemia compared to the mother is no education level, which similar with this finding the study by (Habib et al., 2020) the odds of the occurrence of anemia are lower among children of educated parents Compared to illiterate mothers and the study by (Tadesse et al., 2022) Mothers who were unable to read and write were 53% times more likely to have anemic child (AOR = 1.47; 95% CI: 1.31, 1.65) Similarly the mother's age is a significant determinant of anemia among under-five children consistent with other study (Chungkham et al., 2021) .

The study indicate that female children are less likely to be affected by anemia compared to males, which aligns with similar studies by(Tadesse et al., 2022) , (Gebrehawerial & Lemma, 2020) and Specifically the finding by (Dey & Raheem, 2016)Male children were at greater risk of having severe anemia than female children (OR = 1.488; p = 0.010). However, (Khulu & Ramroop, 2020) present contrasting results, suggesting that females are more likely to experience anemia than their male counterparts.

CHAPTER SIX

6. CONCLUSION AND RECOMMENDATION

6.1 Conclusions

The main purpose of this study was to identify socio demographic, health related factors, nutritional related factor and to assess the cluster variation of anemia levels among under five years based on TDHS-MIS 2022 data using multilevel ordinal logistic regression model.

The study findings indicated that approximately 59.89% of children under the age of five in Tanzania were identified as having anemia.

The multilevel logistic regression model identified factors contributing to anemia among under five children. From the methodological aspect, it was found that multilevel random intercept model is better compared to the variance component model and random coefficient model in fitting the data.

A key finding of this study is the significant cluster variation of anemia among children under five in Tanzania. The fixed part of the random intercept model highlighted the significant influence of variables including the child's age in months, mother's anemia status, geographical zone, stunting levels, mother's education level, mother's age, child's gender, and recent occurrence of diarrhea on anemia prevalence among young children, with statistical significance at the 5% level.

6.2 Recommendations

Based on this study finding we recommended the following:

- As the anemia levels of children vary among geographical zone, It is advisable that the implementation of maternal health related programs, polices and strategies be established by the government giving special attention for zones like southern, eastern and lake zone, which are identified as having higher risks, to effectively combat anemia in children under five years of age.
- To mitigate the risk of anemia among children under five in Tanzania, the government should prioritize expanding health extensions and awareness programs for mothers, emphasizing healthy practices to prevent diarrhea.

- To combat anemia among children under five in Tanzania, the government should prioritize developing and implementing nutrition programs that specifically target stunting in children reduce anemia risk
- Enhancing maternal education should attain at least a primary education level to improve health outcomes for their children.
- Finally, we recommend that further research should be conducted to identify others factors that affect and contribute to anemia among under five years children by considering variations.

References

Abeysekera, W. W. M., & Sooriyarachchi, R. (2008). A novel method for testing goodness of fit of a proportional odds model: An application to an AIDS study. *Journal of the National Science Foundation of Sri Lanka*, *36*(2), 125–135. https://doi.org/10.4038/jnsfsr.v36i2.144

Adugna, D. G., Kibret, A. A., Aragie, H., Enyew, E. F., Dessie, G., Melese, M., Simegn, W., Abebe, E. C., Admasu, F. T., & Dejenie, T. A. (2023). Prevalence and determinants of anemia among children aged from 6 to 59 months in Liberia: a multilevel analysis of the 2019/20 Liberia demographic and health survey data. *Frontiers in Pediatrics*, *11*(April), 1–9. https://doi.org/10.3389/fped.2023.1152083

Agresti, A. (2002). An Introduction to Categorical data Analysis. In *John Wiley Sons, Inc., New York.*

Agresti, A. (2013). 3Rd-Ed-Alan_Agresti_Categorical_Data_Analysis.Pdf. In *International encyclopedia of statistical science* (Vol. 47, Issue 4, pp. 755–758).

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, *19*(6), 716–723.

Asresie, M. B., Fekadu, G. A., & Dagnew, G. W. (2020). Determinants of Anemia among Children Aged 6-59 Months in Ethiopia: Further Analysis of the 2016 Ethiopian Demographic Health Survey. *Advances in Public Health*, *2020*. https://doi.org/10.1155/2020/3634591

Balarajan, Y., Ramakrishnan, U., Özaltin, E., Shankar, A. H., & Subramanian, S. V. (2011). Anaemia in low-income and middle-income countries. *The Lancet*, *378*(9809), 2123–2135. https://doi.org/10.1016/S0140-6736(10)62304-5

Chaparro, C. M., & Suchdev, P. S. (2019). Anemia epidemiology, pathophysiology, and etiology in low- and middle-income countries. *Annals of the New York Academy of Sciences*, *1450*(1), 15–31. https://doi.org/10.1111/nyas.14092

Chungkham, H. S., Marbaniang, S. P., & Narzary, P. K. (2021). "Childhood Anemia in India: an application of a Bayesian geo-additive model." *BMC Pediatrics*, *21*(1), 1–12.

https://doi.org/10.1186/s12887-021-03008-0

Cornet, M., Le Hesran, J. Y., Fievet, N., Cot, M., Personne, P., Gounoue, R., Beyeme, M., & Deloron, P. (1998). Prevalence of and risk factors for anemia in young children in southern Cameroon. *American Journal of Tropical Medicine and Hygiene*, *58*(5), 606–611. https://doi.org/10.4269/ajtmh.1998.58.606

Dey, S., & Raheem, E. (2016). Multilevel multinomial logistic regression model for identifying factors associated with anemia in children 6–59 months in northeastern states of India. *Cogent Mathematics*, *3*(1), 1159798. https://doi.org/10.1080/23311835.2016.1159798

EDHS. (2016). Central Statistical Agency (CSA) [Ethiopia] and ICF. 2016. Ethiopia Demographic and Health Survey 2016. Addis Ababa, Ethiopia, and Rockville, Maryland, USA: CSA and ICF. In *Central Statistical Agency (CSA) [Ethiopia] and ICF.* (Vol. 1, Issue 1).

Ejigu, B. A., Wencheko, E., & Berhane, K. (2018). Spatial pattern and determinants of anaemia in Ethiopia. *PLoS ONE*, *13*(5), 1–16. https://doi.org/10.1371/journal.pone.0197171

Fullerton, A. S., & Xu, J. (2012). The proportional odds with partial proportionality constraints model for ordinal response variables. *Social Science Research*, *41*(1), 182–198. https://doi.org/10.1016/j.ssresearch.2011.09.003

Garcia-Casal, M. N., Pasricha, S. R., Sharma, A. J., & Peña-Rosas, J. P. (2019). Use and interpretation of hemoglobin concentrations for assessing anemia status in individuals and populations: results from a WHO technical meeting. *Annals of the New York Academy of Sciences*, *1450*(1), 5–14. https://doi.org/10.1111/nyas.14090

Gebrehawerial, & Lemma, L. (2020). Factors Associated with Anemia Among Children 6–23 Months of Age in Ethiopia: A Multilevel Analysis of Data from the 2016 Ethiopia Demographic and Health Survey. *Pediatric Health, Medicine and Therapeutics*, *Volume 11*, 347–357. https://doi.org/10.2147/phmt.s258114

Habib, N., Abbasi, S.-U.-R. S., Afzal, A., Arooj, S., & Farid, S. (2020). Socio-Cultural Risk Factors of Anemia among Children under Five Years of Age in District Muzaffarabad, Azad Jammu & Kashmir, Pakistan. *Journal of Pharmaceutical Research International, December*, 100–112. https://doi.org/10.9734/jpri/2020/v32i3030911

Hosmer, Lemeshow, & Sturdivant, 2013. (2013). Applied Logistic Regression. In *John Wiley \& Sons* (Vol. 398). https://doi.org/10.2307/2532419

Hosmer & lemeshow. (2000). Applied logistic regression, 2. *John Wiley \& Sons*.

Jember, T. A., Teshome, D. F., Gezie, L. D., & Agegnehu, C. D. (2021). Spatial variation and determinants of childhood anemia among children aged 6 to 59 months in Ethiopia: further analysis of Ethiopian demographic and health survey 2016. *BMC Pediatrics*, *21*(1), 1–14. https://doi.org/10.1186/s12887-021-02901-y

Kassebaum, N. J., Jasrasaria, R., Naghavi, M., Wulf, S. K., Johns, N., Lozano, R., Regan, M., Weatherall, D., Chou, D. P., Eisele, T. P., Flaxman, S. R., Pullan, R. L., Brooker, S. J., & Murray, C. J. L. (2014). Plenary Paper RED CELLS , IRON , AND ERYTHROPOIESIS A systematic analysis of global anemia burden from 1990 to 2010. *Blood Journal*, *123*(5), 615–625. https://doi.org/10.1182/blood-2013-06-508325.The

Khiari, M., & Rejeb, J. ben. (2015). Determination of the Regional Impact on Innovation with an Ordinal Logit and a Multilevel Analysis. *Procedia - Social and Behavioral Sciences*, *195*, 592–602. https://doi.org/10.1016/j.sbspro.2015.06.139

Khulu, C., & Ramroop, S. (2020). Key determinants of anemia among youngsters under five years in Senegal, Malawi, and Angola. *International Journal of Environmental Research and Public Health*, *17*(22), 1–12. https://doi.org/10.3390/ijerph17228538

Lipsitz, S. R., Fitzmaurice, G. M., & Molenberghs, G. (1996). Goodness-of-Fit Tests for Ordinal Response Regression Models. *Applied Statistics*, *45*(2), 175. https://doi.org/10.2307/2986153

McCullagh, P. (1980). Regression models for ordinal data. *Journal of the Royal Statistical Society: Series B (Methodological)*. https://doi.org/10.1542/peds.2020-004879

McLean, E., Cogswell, M., Egli, I., Wojdyla, D., & De Benoist, B. (2009). Worldwide prevalence of anaemia, WHO Vitamin and Mineral Nutrition Information System, 1993-

2005. *Public Health Nutrition*, *12*(4), 444–454. https://doi.org/10.1017/S1368980008002401

Mghanga, F. P., Genge, C. M., Yeyeye, L., Twalib, Z., Kibopile, W., Rutalemba, F. J., & Shengena, T. M. (2017). Magnitude, Severity, and Morphological Types of Anemia in Hospitalized Children Under the Age of Five in Southern Tanzania. *Cureus*, *9*(7). https://doi.org/10.7759/cureus.1499

Milman, N. (2011). Anemia - Still a major health problem in many parts of the world! *Annals of Hematology*, *90*(4), 369–377. https://doi.org/10.1007/s00277-010-1144-5

Mitchinson, C., Strobel, N., McAullay, D., McAuley, K., Bailie, R., & Edmond, K. M. (2019). Anemia in disadvantaged children aged under five years; Quality of care in primary practice. *BMC Pediatrics*, *19*(1), 1–11. https://doi.org/10.1186/s12887-019-1543-2

O'Connell. (2006). Logistic Regression Models for Ordinal Response Variables. In *Logistic Regression Models for Ordinal Response Variables*. https://doi.org/10.4135/9781412984812

O'Connell, A. A. (2010). An illustration of multilevel models for ordinal response data. *Data and Context in Statistics Education: Towards an Evidence-Based Society. Proceedings of the Eighth International Conference on Teaching Statistics (ICOTS8)}*, *8*.

Perera, A. A. P. N. M., Sooriyarachchi, M. R., & Wickramasuriya, S. L. (2016). A Goodness of Fit Test for the Multilevel Logistic Model. *Communications in Statistics: Simulation and Computation*, *45*(2), 643–659. https://doi.org/10.1080/03610918.2013.868906

Said, F. A., Khamis, A. G., Habib, A., Yang, H., He, Z., & Luo, X. (2021). Prevalence and determinants of anemia among children in zanzibar, tanzania: Analysis of cross-sectional population representative surveys. *Children*, *8*(12). https://doi.org/10.3390/children8121091

Schwarz, G. (1978). Estimating the Dimension of a Model. *The Annals of Statistics*, *6*(2), 461–464. https://doi.org/10.1214/aos/1176344136

Simbauranga, R. H., Kamugisha, E., Hokororo, A., Kidenya, B. R., & Makani, J. (2015). Prevalence and factors associated with severe anaemia amongst under-five children hospitalized at Bugando Medical Centre, Mwanza, Tanzania. *BMC Hematology*, *15*(1), 1–9. https://doi.org/10.1186/s12878-015-0033-5

Sunguya, B. F., Zhu, S., Paulo, L. S., Ntoga, B., Abdallah, F., Assey, V., Mpembeni, R., & Huang, J. (2020). Regional disparities in the decline of anemia and remaining challenges among children in tanzania: Analyses of the tanzania demographic and health survey 2004–2015. *International Journal of Environmental Research and Public Health*, *17*(10), 1–15. https://doi.org/10.3390/ijerph17103492

Tadesse, S. E., Zerga, A. A., Mekonnen, T. C., Tadesse, A. W., Hussien, F. M., Feleke, Y. W., Anagaw, M. Y., & Ayele, F. Y. (2022). Burden and Determinants of Anemia among Under-Five Children in Africa: Systematic Review and Meta-Analysis. *Anemia*, *2022*. https://doi.org/10.1155/2022/1382940

TDHS-MIS. (2022). *Demographic and Health Survey and Malaria Indicator Survey 2022 Final Report. Dodoma, Tanzania, and Rockville, Maryland, USA: MoH, NBS, OCGS, and ICF*.

WHO. (2011a). The global prevalence of anaemia in 2011. *Geneva: World Health Organization; 2015*, 1–48.

WHO. (2011b). The global prevalence of anemia in 2011. Geneva: World Health Organization. *Cochrane Database Syst Rev*, *110*(6), 1–7. http://www.who.int//iris/bitstream/10665/148556/1/WHO_NMH_NHD_14.4_eng.pdf

WHO. (2023). *anemia*. May 1. https://www.who.int/news-room/fact-sheets/detail/anemia

WHO, C. (2008). Worldwide prevalence of anaemia 1993--2005. *WHO Global Database on Anaemia*.

Yusuf, A., Mamun, A. S. M. A., Kamruzzaman, M., Saw, A., Abo El-Fetoh, N. M., Lestrel, P. E., & Hussain, G. (2019). Factors influencing childhood anaemia in Bangladesh: A two level logistic regression analysis. *BMC Pediatrics*, *19*(1), 1–9. https://doi.org/10.1186/s12887-019-1581-9

Appendix

Table A: 1Parallel lines test

Brant test of parallel regression assumption

	chi2	p>chi2	Df
All	48.830	0.060	35
2.child_age	0.010	0.929	1
3.child_age	0.390	0.533	1
4.child_age	0.740	0.389	1
5.child_age	0.030	0.853	1
1.mother_anemial	0.030	0.857	1
2.mother_anemial	0.020	0.877	1
3.mother_anemial	0.080	0.775	1
2.zonee	5.240	0.052	1
3.zonee	0.000	0.986	1
4.zonee	2.030	0.154	1
5.zonee	1.950	0.163	1
6.zonee	1.080	0.298	1
7.zonee	1.420	0.234	1
8.zonee	0.540	0.463	1
9.zonee	0.210	0.650	1
2.Stunting_L	2.750	0.098	1
3.Stunting_L	2.450	0.117	1
1.mother_education	0.170	0.677	1
2.mother_education	0.010	0.939	1
3.mother_education	4.650	0.031	1
2.mother_age	0.040	0.833	1
3.mother_age	0.060	0.805	1

4.mother_age	0.230	0.628	1
5.mother_age	0.000	0.971	1
6.mother_age	1.080	0.299	1
7.mother_age	0.030	0.864	1
2.child_sex	3.120	0.077	1
1.toilet	1.000	0.318	1
2.diarrahea_recently	1.400	0.237	1
2.wealth_index	1.200	0.274	1
3.wealth_index	1.100	0.294	1
4.wealth_index	0.280	0.596	1
5.wealth_index	0.910	0.340	1
2.WHZ_Ordinal	1.890	0.169	1
3.WHZ_Ordinal	0.210	0.650	1

A significant test statistic provides evidence that the parallel regression assumption has been violated.

Table A: 2 all parameter estimates

Anemia_level	Odds Ratio	Std. Err.	z	P>|z|	[95% Conf.	Interval]
child_sex						
2. female	.8739523	.0545524	-2.16	0.031	.7733133	.9876885
residence						
2. rural	.9992087	.0982858	-0.01	0.994	.8240032	1.211668
zonee						
2. northern	1.015173	.1730258	0.09	0.930	.7268779	1.417812
3. central	.6878402	.1152113	-2.23	0.025	.4953517	.9551277
4. southern highlands	.7413305	.1507111	-1.47	0.141	.497696	1.10423
5. southern	1.866815	.399976	2.91	0.004	1.226666	2.841032
6. south west highlands	1.288934	.2205098	1.48	0.138	.9217386	1.80241
7. lake	1.342	.18706	2.11	0.035	1.021185	1.763602
8. eastern	1.878468	.3101525	3.82	0.000	1.359134	2.596241
9. zanzibar	1.455683	.3656245	1.49	0.135	.8897541	2.38157
child_age						
2. 12-23 months	.814918	.0901217	-1.85	0.064	.6561145	1.012158
3. 24-35 months	.4959376	.0566643	-6.14	0.000	.3964344	.6204156
4. 36-47 months	.3503053	.0407104	-9.03	0.000	.2789492	.4399144
5. 48-59 months	.2484931	.029296	-11.81	0.000	.197225	.313088
mother_education						
1. primary	.7935428	.067257	-2.73	0.006	.6720885	.9369454
2. secondary	.8087248	.0947517	-1.81	0.070	.6427949	1.017487
3. higher	.8713948	.3269953	-0.37	0.714	.4176354	1.818162
wealth_index						
2. poorer	.8570619	.0852183	-1.55	0.121	.7053043	1.041473
3. middle	.8804038	.0916532	-1.22	0.221	.717908	1.07968
4. richer	.9511056	.1057597	-0.45	0.652	.7648531	1.182713
5. richest	.7933553	.1000364	-1.84	0.066	.6196369	1.015776
mother_age						
2. 20-24	.5983019	.0997897	-3.08	0.002	.4314695	.8296419
3. 25-29	.5346231	.089205	-3.75	0.000	.3854959	.7414393
4. 30-34	.4792881	.0811861	-4.34	0.000	.3438844	.6680067
5. 35-39	.5034813	.0877008	-3.94	0.000	.3578603	.7083585
6. 40-44	.4895164	.0930646	-3.76	0.000	.3372413	.7105487
7. 45-49	.6338013	.1615748	-1.79	0.074	.3845532	1.044599
mother_anemial						
1. mild	1.458766	.1240733	4.44	0.000	1.234775	1.723389
2. moderate	2.221194	.1888081	9.39	0.000	1.88032	2.623863
3. severe	1.39278	.2749089	1.68	0.093	.9459565	2.050661
vitamin_lastsixmonth						
1. yes	1.006548	.0660028	0.10	0.921	.8851525	1.144592
toilet						
1. Unimproved	1.120756	.094215	1.36	0.175	.9505085	1.321498
source_drwater						
unimproved	.9028515	.0744884	-1.24	0.215	.7680495	1.061313
diarrahea_recently						
2. yes, last two weeks	1.321724	.1537058	2.40	0.016	1.052333	1.660079
faver_recently						
1. yes	.9326007	.0942595	-0.69	0.490	.7650033	1.136915
WHZ_Ordinal						
Moderate Wasting	1.23042	.2657051	0.96	0.337	.8058218	1.878743
Severe Wasting	1.481387	.5540803	1.05	0.293	.7116975	3.083485
underW						
Moderate underw	1.120326	.1368985	0.93	0.352	.8817217	1.423499
Severe w	.8883759	.2280472	-0.46	0.645	.5371458	1.469269
Stunting_L						
Moderate Stunting	1.310495	.1054699	3.36	0.001	1.119257	1.534408
Severe Stunting	1.327514	.1817551	2.07	0.039	1.015074	1.736122
/cut1	-1.614738	.2506622			-2.106027	-1.123449
/cut2	-.3654309	.2493009			-.8540517	.1231898
EA						
var(_cons)	.2101618	.0513043			.1302447	.3391153

Table A: 3 random coefficient model

Anemia_level	Odds Ratio	Std. Err.	z	P>z	[95% Conf.	Interval]
Children age						
12-23 months	.8143188	.0898134	-1.86	0.063	.6560141	1.010825
24-35 months	.4945961	.0564423	-6.17	0.000	.3954697	.618569
36-47 months	.3505036	.0406469	-9.04	0.000	.2792422	.4399505
48-59 months	.2479464	.0291746	-11.85	0.000	.1968798	.3122586
Mother's anemia status						
mild	1.464859	.1244274	4.49	0.000	1.240205	1.730208
moderate	2.233576	.1896036	9.47	0.000	1.891228	2.637894
severe	1.370197	.270425	1.60	0.111	.9306539	2.017335
Zone						
northern	1.008968	.171067	0.05	0.958	.7237011	1.406681
central	.6779789	.1128253	-2.34	0.020	.4892876	.9394382
southern highlands	.734175	.1478802	-1.53	0.125	.4947063	1.089561
southern	1.854072	.392643	2.92	0.004	1.224235	2.807943
south west highlands	1.302564	.2201653	1.56	0.118	.9352447	1.81415
lake	1.318787	.1813463	2.01	0.044	1.007224	1.726726
eastern	1.936839	.309705	4.13	0.000	1.415743	2.649736

zanzibar	1.463966	.3655776	1.53	0.127	.8973691	2.38831
Stunting						
Moderate Stunting	1.333873	.1028317	3.74	0.000	1.146814	1.551443
Severe Stunting	1.352191	.1592648	2.56	0.010	1.073448	1.703314
Mother Education						
primary	.7977045	.0667773	-2.70	0.007	.6769964	.9399346
secondary	.8178471	.0927538	-1.77	0.076	.6548404	1.02143
higher	.8699443	.3249729	-0.37	0.709	.418331	1.809101
Mothers age						
20-24	.5996467	.0999659	-3.07	0.002	.4325072	.8313761
25-29	.5377384	.0897692	-3.72	0.000	.3876795	.7458806
30-34	.4832111	.0817962	-4.30	0.000	.3467757	.6733256
35-39	.5068931	.088259	-3.90	0.000	.3603356	.7130591
40-44	. 4957633	.0942058	-3.69	0.000	.3416077	.7194839
45-49	.6392192	.1623609	-1.76	0.078	.3885488	1.051608
Child sex						
female	.8706351	.0542119	-2.22	0.026	.7706095	.9836441
toilet sanitation						
Unimproved	1.112002	.0891421	1.32	0.185	.9503208	1.301191

Diarrhea recently						
yes, last two weeks	1.286456	.1537385	2.11	0.035	1.017821	1.625991
Wealth index						
poorer	.8632244	.0852272	-1.49	0.136	.7113504	1.047524
middle	.8905344	.0916305	-1.13	0.260	.7278926	1.089517
richer	.9626631	.1043874	-0.35	0.726	.7783472	1.190626
richest	.8167125	.098051	-1.69	0.092	.6454722	1.033382
wasting						
Moderate Wasting	1.264941	.2496397	1.19	0.234	.8591778	1.862333
Severe Wasting	1.4173	.500765	0.99	0.324	.7091084	2.832768
/cut1	-1.568494	.2355143			- 2.030094	- 1.106895
/cut2	-.317639	.2342951			.7768488	.1415709
EA						
var(diarrahea_recently)	.0289978	2.30e-06			.0289933	.0290023
var(_cons)	.1863885	.0000178			.1863535	.1864235
EA						
cov(diarrahea_recently,cons)	.0735178	2.63e-06	2.8e+04	0.000	.0735126	.0735229