

Factors affecting child malnutrition in Ethiopia

Ashenafi Argaw Yirga¹, Henry G Mwambi¹, Dawit Getnet Ayele², Sileshi Fanta Melesse¹

1. School of Mathematics, Statistics and Computer Science, University of KwaZulu-Natal, Pietermaritzburg, Private Bag X01, Scottsville, 3209, South Africa.
2. Bloomberg School of Public Health, The Johns Hopkins University.

Emails:

AshenafiArgawYirga: Email: ashu3argaw@gmail.com; Henry G Mwambi: Email: mwambih@ukzn.ac.za; Sileshi Fanta Melesse: Email: melesse@ukzn.ac.za

Abstract

Background: One of the public health problems in developing countries is child malnutrition. An important factor for children's well-being is good nutrition. Therefore, the malnutrition status of children under the age of five is an important outcome measure for children's health. This study uses the proportional odds model to identify risk factors associated with child malnutrition in Ethiopia using the 2016 Ethiopian Demographic and Health Survey data.

Methods: This study uses the 2016 Ethiopian Demographic and Health Survey results. Based on weight-for-height anthropometric index (Z-score) child nutrition status is categorized into four levels namely- underweight, normal, overweight and obese. Since this leads to an ordinal variable for nutrition status, an ordinal logistic regression (OLR)proportional odds model (POM) is an obvious choice for analysis.

Results: The findings and comparison of results using the cumulative logit model with and without complex survey design are presented. The study results revealed that to produce the appropriate estimates and standard errors for data that were obtained from complex survey design, model fitting based on taking the survey sampling design into account is better. It has also been found that for children under the age of five, weight of a child at birth, mother's age, mother's Body Mass Index (BMI), marital status of mother and region (Affar, Dire Dawa, Gambela, Harari and Somali) were influential variables significantly associated with underfive children's nutritional status in Ethiopia.

Conclusion: This child's age of a child, sex, weight of child at birth, mother's BMI and region of residence were significant determinants of malnutrition of children under five years in Ethiopia. The effect of these determinants can be used to develop strategies for reducing child malnutrition in Ethiopia. Moreover, these findings show that OLR proportional odds model is appropriate assessing the determinants of malnutrition for ordinal nutritional status of underfive children in Ethiopia.

Keywords: BMI, Ethiopian Demographic and Health Survey (EDHS), malnutrition, proportional odds model.

DOI: <https://dx.doi.org/10.4314/ahs.v19i2.13>

Cite as: Yirga AA, Mwambi HG, Ayele DG, Melesse SF. Factors affecting child malnutrition in Ethiopia. *Afri Health Sci.*2019;19(2): 1897-1909. <https://dx.doi.org/10.4314/ahs.v19i2.13>

Corresponding author:

Dawit Getnet Ayele,
Bloomberg School of Public Health,
The Johns Hopkins University
Cell: +1 (443) 554 6721
Email: ejgmul@yahoo.com, dayele1@jhu.edu

Introduction

Nutrition is an important factor for good health. In developing countries, child malnutrition is the leading public health problem and is a major cause of child morbidity and mortality. For malnutrition, under five children are the most vulnerable. The nutrition of infants and young children is a major concern to any society¹. The high-

est child malnutrition is found in the sub-Saharan Africa countries. Ethiopia is among those countries with the highest rate of stunting in sub-Saharan Africa. The proportion of underweight children is highest in the age range of 2 to 3 years (34%) and lowest among those under six months of age (10%). In general, 29% of children the under age of five are underweight, and 9% are severely underweight in Ethiopia. An estimated 159 million children under five years of age, or 23.8%, were stunting in 2016, 15.8% decrease from an estimated 255 million in 1990 worldwide⁵⁴. Even though the occurrence of stunting and underweight among children under five years of age worldwide has decreased since 1990, overall improvement is unsatisfactory and millions of children remain at risk^{2,3}. Malnutrition is the cause of substantial health problems in children that need due consideration. For that reason, reducing malnutrition of children is equal to improving the health status of the children. This is necessary in order to improve the health status of the future core segment of the society. This is crucial for economic growth and development of the society under consideration⁴.

To measure nutrition, Body Mass Index (BMI) is used and can be defined as the ratio of weight (kg) to squared height (m²). But BMI is not a direct measure of body fatness. BMI is dependent on age and gender for children and referred as BMI-for-age^{5,6}. Using this variable, the percentiles or quantiles of BMI for specified ages are of interest can be defined. Moreover, it gives a reference for individuals at that age with respect to the population. Determining BMI for children's weight status has been of interest to many researchers. Children's BMI under the age of five at or above the 95th percentile, between the 85th and 95th percentile and between the 5th and 85th percentile were classified as obesity, overweight and normal (healthy weight) respectively⁷. These intervals and cutoffs were as result of expert knowledge. The World Health Organization Expert Committee on Physical Status suggested the cutoff for underweight corresponding to BMIs less than the 5th percentile⁸.

A wide range of nutrient-related deficiencies and disorders are included in malnutrition⁹. Different studies have been conducted in Ethiopia and regionally. These local and regional studies have shown an increase in malnutrition with increase in age of the child¹⁰⁻¹³. A study

conducted by Teller and Yimer (2000) in the Southern Nations, Nationalities and Peoples Region (SNNPR) of Ethiopia showed that women belonging to low economic status households were affected by malnutrition¹³. Another study conducted in India showed that 60% of deaths of children under the age of five are related to malnutrition. Malnutrition for children is strongly correlated with mother's poor nutritional status^{14,4,12,15}. Education is also one of the most significant factors that enables or empowers women to provide suitable care for their children, which is an important determinant of children's growth and development^{16,17,18}.

The main objective of this study is to use ordinal proportional odds model to identify the determinants of under-five children's nutritional status as a function of age and other relevant factors. This study will assist policy makers to know and understand the areas they need to focus on in order to improve the planning and assessment of health policies to avoid child mortality associated with malnutrition and to enhance children's health, diet and growth.

Methods and materials

Ethiopia is one of the sub-Saharan Africa countries located in the Eastern Africa region. In 2000, Ethiopia conducted the first Ethiopian Demographic and Health Survey (EDHS). As a continuous study, subsequent EDHSs were conducted in 2005, 2011 and 2016. These surveys are periodic cross-sectional surveys administered at the household level. The 2016 Ethiopian Demographic and Health Survey results were used for this study. Central Statistical Agency of Ethiopia was the responsible organization for the survey¹⁹. The 2016 EDHS sample was designed to provide estimates for the health and demographic variables of interest for Ethiopia as a whole; this comprised both urban and rural areas of Ethiopia and 11 geographical areas. For the survey, 17,817 households were included in data collection. The 2007 Population and Housing Census results were used as the sampling frame²⁰⁻²².

Study variable

The response variable for this study is under-five children's nutritional status in Ethiopia, which is an ordinal categorical variable. The explanatory variables used in this study are:- child's age, sex of a child, weight of child at birth, mother's current age, mother's BMI, educational

attainment of mother, mother's work status, religion, religion, wealth index, place of residence (rural or urban), and current marital status. The socio-economic and demographic factors used in this study were suggested by several researchers. These factors were referred to as intermediate variables for the determinants of children's nutritional status⁵³.

Statistical methods

An outcome with more than two categories is known as a polytomous outcome. Let J denote the number of categories for such an outcome. Out of N observations, Y_1, Y_2, \dots, Y_J are the frequencies in categories 1, 2, ..., J with corresponding probabilities, $\pi_1, \pi_2, \dots, \pi_J$, respectively. The distribution is the multinomial distribution and can be expressed as follows:

$$P(Y_1, Y_2, \dots, Y_J) = \frac{N!}{\prod_{j=1}^J Y_j!} \times \prod_{j=1}^J \pi_j^{Y_j}$$

The distribution leads to the multinomial (polytomous) logistic regression which is an extension of binary logistic regression. The link function is the multinomial logit model because the probability distribution for the outcome variable is assumed to be a multinomial rather than a binomial distribution. For a polytomous response, it is further important to note whether the response is nominal (consisting of unordered categories) or ordinal (consisting of ordered categories). An outcome variable that has two or more nominal categories can be modeled using multinomial logistic regression. It estimates the odds of being at any category compared to being at the baseline category (comparison or reference category). The model can be treated as a combination of a series of binary logistic regression models.

Suppose Y can take on values coded as 1, 2, ..., J . Next pick one of the outcome levels say J as the reference level. If we assume we have P covariates then the model is formulated as:

$$\log \left[\frac{\pi(Y=j|x_1, x_2, \dots, x_p)}{\pi(Y=J|x_1, x_2, \dots, x_p)} \right] = \beta_{j0} + \beta_{j1}x_1 + \beta_{j2}x_2 + \dots + \beta_{jp}x_p, \quad (1)$$

, where $j=1, 2, \dots, J-1$; J is the outcome from the base category, which can be any category but is generally the highest one; β_{j0} are the intercepts, and $\beta_{j1}, \beta_{j2}, \dots, \beta_{jp}$ are the regression coefficients. Since the model includes $J-1$

comparisons, it estimates $J-1$ logit function for each predictor²³.

Commonly the maximum likelihood procedure is used to estimate parameters for the multinomial logistic regression model as it is the case with the binary logistic regression. For nominal categories, one of the categories is designated as a reference or base category and each of the remaining categories is compared with the reference category²⁴.

Ordinal logistic regression (OLR) considers any inherent ordering of the levels in the outcome variable and makes full use of the ordinal information^{25,26}. The incorporation of ordering can result in models that have simpler interpretations. Although ordinal outcomes can be simple and meaningful their optimal statistical treatment remains challenging to many applied researchers²⁷⁻²⁹. Moreover, these models have greater power than the multinomial logit models³⁰⁻³². However, a variable that can be ordered when considered for one purpose could be unordered differently when used for another purpose. Miller and Volker (1985) shows how different assumptions about the ordering of occupations result in different conclusions³³. Therefore, we need to think carefully before concluding that the outcome is ordinal³⁴. Although the categories for an ordinal variable can be ordered, the distances between the categories are unknown. Multinomial logistic for ordinal responses is normally called ordinal logistic regression. An ordinal logistic regression model is a generalization of a binary logistic regression model, when the outcome variable has more than two ordinal levels. In Stata, the ordinal logistic regression model assumes that the outcome variable is a latent variable, which is expressed in logit form as follows:

$$\log \left(\frac{P(Y \leq j|x)}{1 - P(Y \leq j|x)} \right) = \beta_{j0} + (-\beta_{j1}x_1 - \dots - \beta_{jp}x_p), \quad (2)$$

, where $Pr(Y \leq j|x_1, x_2, \dots, x_p)$ is the probability of being at or below category j , given a set of predictors $v=1, 2, \dots, p$. β_{j0} are the cutoff points (thresholds), and $\beta_{j1}, \beta_{j2}, \dots, \beta_{jp}$ are logit coefficients²³.

According to Agresti (2002), one way to use category ordering is to form logits of cumulative probabilities²⁴,

$$P(Y \leq j|x) = \frac{\exp(\beta_{j0} + X'\beta)}{1 + \exp(\beta_{j0} + X'\beta)} = \pi_1(x) + \dots + \pi_j(x), \quad j = 1, \dots, J \quad (3)$$

Equivalently the cumulative logits (logits of cumulative probabilities) can be defined as

$$\text{logit}[P(Y \leq j|x)] = \log \left[\frac{P(Y \leq j|x)}{P(Y > j|x)} \right] = \log \left[\frac{P_1(x)+P_2(x)+\dots+P_j(x)}{P_{j+1}(x)+P_{j+2}(x)+\dots+P_k(x)} \right], j = 1, 2, \dots, J - 1.$$

Each cumulative logit uses all J response categories. In Stata, the logit form of the ordinal logistic regression model that simultaneously uses all cumulative logits can be expressed as follows:

$$\text{logit}[P(Y \leq j|x)] = \beta_{j0} + (-X'\beta), \quad j = 1, 2, \dots, J, \quad (4)$$

where $P(Y \leq j|x)$ is the cumulative probability of the event $(Y \leq j|x)$, β_{j0} are the unknown intercept parameters increasing in j, and $\beta = (\beta_1, \beta_2, \dots, \beta_p)'$ is a vector of unknown regression coefficients corresponding to x. since $P(Y \leq j|x)$ increases in j for fixed x, the logit is an increasing function of this probability.

The cumulative logit model (4) satisfies

$$\begin{aligned} & \text{logit}[P(Y \leq j |x_1)] - \text{logit}[P(Y \leq j |x_2)] \\ &= \log \left(\frac{P(Y \leq j |x_1)/P(Y > j |x_1)}{P(Y \leq j |x_2)/P(Y > j |x_2)} \right) = \beta' (x_1 - x_2). \end{aligned}$$

An odds ratio of cumulative probabilities is called a cumulative odds ratio. The odds of the event $Y \leq j$ at $x=x_1$ is $\exp[\beta'(x_1 - x_2)]$ times the odds of the same event at $x=x_2$. The log cumulative odds ratio is proportional to the distance between x_1 and x_2 . The same proportionality constant applies to each logit. Because of this property, cumulative logit model, is called the proportional odds model^{24,35,36}.

Ordinal variables are often coded as consecutive integers from 1 to the number of categories. Because of this coding, it is tempting to analyze ordinal outcomes with the linear regression model. However, an ordinal response variable violates the assumptions of linear regression model, which can lead to incorrect conclusions^{37,38}. With an ordinal response, it is much better to use models that avoid the assumption that the distances between categories are equal. Although many models have been designed for ordinal outcomes, logit and probit models are commonly used as the link function in ordinal regression models³⁹. Most multinomial regression models for ordinal outcome variables are based on the logit function. The difference between both functions is typically only seen in small samples, because the probit link assumes the normal distribution of the probability of event, whereas the logit link assumes the logistic distribution. Details

about models for ordinal outcomes can be found in different literatures^{31,32,36,40-42}.

We label the four levels of under five children's nutritional status as 1, 2, 3, and 4 where we compare underweight, normal weight, overweight, and obese at the same time. Since this leads to an ordinal variable for nutritional status, an ordinal logistic regression (OLR) is an obvious choice for analysis.

There are many ways of generalizing the logit model to handle ordered categories, such as the partial proportional odds, continuation-ratio, adjacent-category logits, cumulative logits, and stereotype logistic models. Despite this diversity and the vast variety of studies on the subject their use in the public health area is still rare^{34,43,44,45}. This may be attributed not only to their complexity, but especially to the difficulty encountered when it comes to validating their assumptions⁴⁶. When the dependent variable has only two categories, the usual binary logistic model is appropriate.

The usual proportional odds model assumes that data are collected using simple random sampling by which each sampling unit has an equal probability of being selected from a population. When the data comes from a complex survey design with the use of different strata, clustered sampling techniques, and unequal selection probabilities, it is inappropriate to conduct the proportional odds model analysis for the ordinal response variable without taking the survey sampling design into account. Ignoring these features in data analysis may lead to biased estimates of parameters, incorrect variance estimates and misleading results. The parameters and their variance may be either overestimated or underestimated⁴⁷. In such cases, a specialized technique to produce the appropriate estimates and standard errors for ordinal outcome variable should be used. This method takes into account the weight in the survey sampling design.

Features of complex surveys such as sampling weights, strata, and clusters, have been illustrated in literature^{50,47}. In Stata, svy prefix command for survey data is used to fit the proportional odds model when taking all the elements of survey design features into account. It is necessary to specify strata, cluster and weights before fitting the model. For more details on how to use this command one can use the help svyset command in stata software.

Results

The proposed model namely the proportional odds model was applied to the 2016 Ethiopian DHS data and the

results of the application are herein discussed. In addition to the response and explanatory variables, we also assessed two-way interaction effects: unfortunately we did not find any significant interaction effect. Stata `ologit` command was used for model fitting.

Table 1 shows the results for the proportional odds model under the simple random sampling assumption. The log likelihood at each iteration shows that ordinal logistic regression, like binary and multinomial logistic regression, uses maximum likelihood estimation, which is an iterative procedure. Iteration 0 is the log likelihood of the “null” or “empty” model; that is, a model with no predictors. At the next iteration, the predictors are included in the model. At each iteration the log likelihood increases because the goal is to maximize the log likelihood. When the difference between successive iterations is very small, the model is said to have “converged”, and the iteration stops⁴⁹.

The value for log likelihood of the fitted model is -6009.1723, which is used in the likelihood ratio chi-square test of whether all predictors’ regression coeffi-

cients in the model are occurring at the same time zero and in tests of nested models. The likelihood ratio chi-square (LR χ^2) tests that at least one of the predictors’ regression coefficient is not equal to zero. The number in the parenthesis indicates the degrees of freedom of the Chi square distribution used to test the LR χ^2 statistic and is defined by the number of predictors in the model. The LR χ^2 statistic can be calculated by $-2(L(\text{null model}) - L(\text{fitted model})) = -2((-6362.3952) - (-6009.1723)) = 708.45$, where $L(\text{null model})$ is from the log likelihood of the model with no predictor variable (Iteration 0) and $L(\text{fitted model})$ is the log likelihood from the final iteration (assuming the model converged) with all the parameters. The P-value of the log likelihood ratio Chi-square test with 28 degree of freedom, $LR \chi^2(28) = 708.45$, $Prob > \chi^2 = 0.0000$, which indicates that at least one of the logit regression coefficient of the predictors were statistically different from 0. So the full model with all predictors provided a better fit than the null model with no independent variables in predicting cumulative probabilities of under-five children nutritional status (Table 1).

Table 1: Parameter estimates using proportional odds model with complex survey design

Parameters	Coeff.	St.error	OR	P-value	[95% C. I for β]
cut1	-3.2418	0.3566			(-3.9423,-2.5414)
cut2	1.7371	0.3504			(1.0489,2.4252)
cut3	2.9871	0.3628			(2.2745, 3.6998)
Current age of child	-0.3186	0.0274	0.7271	0.0040	(-0.3725,-0.2648)
Sex of child (ref. Male)					
Female	-0.2417	0.0834	0.7852	0.0000	(-0.4056,-0.0779)
Weight of child at birth (ref. Small)					
Large	0.5481	0.1221	1.7301	0.0000	(0.3082, 0.7880)
Average	0.3134	0.1004	1.368	0.0020	(0.1161, 0.5106)
Mother's current age	-0.0203	0.0064	0.9798	0.0020	(-0.0331,-0.0075)
Mother's BMI	0.0479	0.0136	1.0491	0.0000	(0.0212,0.0747)
Mother work status (ref. No)					
Yes	-0.0331	0.0963	0.9673	0.7310	(-0.2224, 0.1561)
Educational attainment of mother (ref. No education)					
Primary	-0.065	0.0966	0.937	0.5010	(-0.2549,0.1248)
Secondary	0.0919	0.1645	1.0963	0.5770	(-0.2312,0.4151)
Higher	0.0701	0.2332	1.0726	0.7640	(-0.3879,0.5282)
Current marital status (ref. Married)					
Not married	-0.3423	0.1519	0.7101	0.0250	(-0.6406,-0.0440)
Religion (ref. Orthodox)					
Catholic	-0.2474	0.3597	0.7808	0.4920	(-0.9538, 0.4590)
Muslim	0.2161	0.1384	1.2412	0.1190	(-0.0557, 0.4879)
Protestant	0.0635	0.1179	1.0655	0.5910	(-0.1681,0.2951)
Other	0.6863	0.6262	1.9864	0.2740	(-0.5434, 1.9160)
Region (ref. Oromia)					
Addis Ababa	0.2395	0.2165	1.2706	0.2690	(-0.1858, 0.6648)
Affar	-0.4445	0.1481	0.6411	0.0030	(-0.7355,-0.1535)
Amhara	-0.0663	0.1339	0.9357	0.6210	(-0.3295,0.1967)
Benishangul-Gumuz	-0.2702	0.1582	0.7632	0.0880	(-0.5809, 0.0404)
Dire Dawa	-0.5879	0.1522	0.5554	0.0000	(-0.8868, -0.2889)
Gambela	-0.612	0.1845	0.5422	0.0010	(-0.9745, -0.2496)
Harari	-0.3224	0.1557	0.7243	0.0390	(-0.6282, -0.0167)
SNNP	0.1665	0.1381	1.1811	0.2280	(-0.1047, 0.4378)
Somali	-0.9143	0.1651	0.4007	0.0000	(-1.2387, -0.5899)
Tigray	-0.1783	0.1445	0.8367	0.2180	(-0.4622,0.1056)
Place of residence (ref. Rural)					
Urban	0.2044	0.2022	1.2268	0.3120	(-0.1927, 0.6016)
Wealth index (ref. Poor)					
Middle	-0.017	0.1101	0.983	0.8770	(-0.2332,0.1991)
Rich	0.1908	0.1152	1.2102	0.0980	(-0.0355, 0.4171)
F(28, 588) = 13.01			Prob> F= 0.0000		

Table 1 reports three cut-points: cut¹, cut² and cut³. Cut¹ is the estimated cut point on the latent variable used to differentiate underweight status from normal, overweight, and obese status when values of the predictor variables are evaluated at zero. When the ordinal outcome category is 1 given significant predictor variables (for categorical variables the reference variable evaluated at zero) and

had zero value for all other predictor variables, the latent variable falls at or below the first cut point, -2.7741. Cut² is the estimated cut point on the latent variable used to differentiate underweight and normal weight status from overweight and obese weight status when values of the predictor variables are evaluated at zero. When the ordinal outcome category is 2 given significant predictor

variable and controlling for all other predictor variables in the model, the latent variable falls between the first cut point, -2.7741 and the second cut point, 2.2885. Cut³ is the estimated cut point on the latent variable used to differentiate underweight, normal, and overweight status from obese status when values of the predictor variables are evaluated at zero. When the ordinal outcome category is 3 given significant predictor variable (reference variable evaluated at zero) and controlling for all other predictor variables in the model, the latent variable falls between cut², 2.2885 and cut³, 3.5703; classified as overweight. When the ordinal outcome category reaches 4, if the latent variable had a value at or beyond the third cut point, 3.5703, controlling for all other predictor variables in the model would be classified as child with obese nutrition status.

Table 1 shows the effect of socio-economic, demographic and geographic factors that have influence in fitting the proportional odds model for underfive children ordinal nutritional status. The estimated logit regression coefficients of current age of child is, $\beta = -0.3233$ (P-value=0.000). This is the ordered log-odds estimate for a one-unit increase in age of a child on the expected nutritional status level given the other variables are held constant in the model. The estimated coefficients of female child is ($\beta = -0.2681$, P-value=0.000). The estimated coefficients of weight of child at birth are: large ($\beta = 0.5547$, P-value=0.000), average ($\beta = 0.3776$, P-value=0.000). For mother's BMI, ($\beta = 0.0604$, P-value=0.000), which is the ordered log-odds estimate for one-unit increase in mother's BMI keeping other variables constant. The estimated coefficients for regions are: Affar, $\beta = -0.3951$ (P-value=0.001), DireDawa, $\beta = -0.6004$ (P-value=0.000), Gambela, $\beta = -0.4453$ (P-value=0.002), Harari, $\beta = -0.3353$ (P-value=0.016), SNNP, $\beta = 0.2134$ (P-value=0.047), Somali, $\beta = -0.8988$ (P-value=0.000) were found to be significant determinants of underfive children's nutritional status.

Substituting the values of the estimated logit coefficients into the equation (4) resulted in $\text{logit}[P(Y \leq j | x)] = \beta_{-j0} + (-\beta_{-jp} x)$. By exponentiating the negative logit coefficients ($e^{(-\beta)}$) the odds of being at or below a particular ordinal nutritional status category, that is obese versus being below that category (overweight, normal and underweight), were obtained. Therefore, to estimate the cumulative odds of being at or below a particular underfive ordinal nutritional status variable (based on weight) category j , for the first predictor, current age

of child, the logit form of proportional odds model was used, $\text{logit}[P(Y \leq j | x_{-1})] = \beta_{-j0} - (-0.3233(\text{age}))$. $OR = e^{((0.3233))} = 1.3817$, indicating that the odds of being at or below a particular underfive ordinal nutritional status variable (based on weight) increased by 38.17% with a one unit increase in the value of current age of a child, holding other variables constant. The estimated cumulative odds of being at or below an ordinal nutritional status (based on weight) category j , for female child, we calculated $\text{logit}[P(Y \leq j | x_{-1})] = \beta_{-j0} + (-0.2681(\text{female}))$. $OR = e^{((0.2681))} = 1.3075$, suggesting that the odds of female child being at or below a particular underfive ordinal nutritional status (based on weight) increased by 30.75%. The estimated cumulative odds of being at or below an ordinal nutritional status (based on weight) category j , for child who had large weight at birth, we calculated $\text{logit}[P(Y \leq j | x_{-1})] = \beta_{-j0} + (0.5547(\text{large}))$. $OR = e^{((-0.5547))} = 0.5743$, suggesting that a child who had large weight at birth, the odds of being at or below a particular underfive ordinal nutritional status (based on weight) decreased by $(1 - 0.5743) \times 100\% = 42.57\%$ as compared to small weight of child at birth, controlling for all other independent variables in the model. The estimated cumulative odds of being at or below an ordinal nutritional status (based on weight) category j , for a child who had average weight at birth, we calculated $\text{logit}[P(Y \leq j | x_{-1})] = \beta_{-j0} + (0.3776(\text{average}))$. $OR = e^{((-0.3776))} = 0.6856$, suggesting that a child who had average weight at birth, the odds of being at or below a particular underfive ordinal nutritional status (based on weight) decreased by $(1 - 0.6856) \times 100\% = 31.44\%$ as compared to small weight of child at birth, controlling for all other independent variables in the model. The odds of being at or below a particular underfive ordinal nutritional status for the other significant effects were computed in the same way as above. It was found that for a one-unit increase in the value of mother's BMI, holding other variables constant, the odds of being at or below a particular underfive ordinal nutritional status decreased by $(1 - 0.9414) \times 100\% = 5.86\%$ ($OR = 0.9414$). The odds of being at or below a particular underfive ordinal nutritional status variable for children from Affar region was 1.4845 (P-value= 0.001) times the odds of children from Oromia region. The odds of being at or below a particular underfive ordinal nutritional status variable for children from Dire Dawa region was 1.8228 (P-value= 0.000) times the odds of children from Oromia region. The odds of being at or below a particular underfive ordinal nutritional status variable for

children from Dire Dawa region was 1.8228 (P-value=0.000) times the odds of children from Oromia region. The odds of being at or below a particular ordinal nutritional status category for children from Gambela, Harari and Somali were respectively 1.5609 (P-value=0.002), 1.3984 (P-value=0.016), and 2.4567 (P-value=0.000) times the odds of children from Oromia region. However, the odds of being at or below a particular ordinal nutritional status category for children from SNNP was 0.8078 (P-value=0.047) times the odds for children from Oromia region (see Table 1).

The odds of being beyond a particular category of ordinal nutritional status are the inverse of those of being at or below a category [48], equation (4) can be transformed to $\text{logit}[P(Y > j|x)] = -\beta_0 + \beta_j x$. Odds ratios (Table 1) can be used directly for the analysis. In terms of odds ratio (Table 1), it was found that the odds of being beyond a particular category of ordinal nutritional status was increased by $(1 - 0.7237) \times 100\% = 27.63\%$ (P-value=0.000) with a one-year increase in current age of child, holding other variables constant. Similarly, the odds of being beyond a particular under five ordinal nutritional status for female child was 0.7647 times the odds of male child. The odds of being beyond a particular category of ordinal nutritional status for children who had large weight at birth was 1.7414 times the odds of children who had small weight at birth. The odds of being at or beyond a particular category of ordinal nutritional status for other significant effects can be interpreted in the same way as above.

Application of complex survey design for ordinal logistic regression

In the subsequent section, the same variables from the previous section are used for data analysis with reference to the Ethiopian DHS (2016) data. Here we investigate the relationship (association) between the response variable and the explanatory variables by the method of proportional odds (PO) model with complex survey design using the `stata: svy: ologit` prefix command. Stata's survey data `svy` prefix command is used to fit the PO model when taking all the elements of survey design features such as strata, cluster, and weight variables into account [48].

The result of the `svy: ologit` is indicated in Table 1 below. The `svy: ologit` for PO model that considers sampling design, reports the adjusted Wald test for all parameters rather than the log likelihood ratio Chi-square test for

the ordinal PO model [47]. $F(28, 588) = 13.01$, $\text{Prob} > F = 0.0000$ indicates that the full model with all parameters was significant in fitting the PO model with complex survey design. The logit coefficients and odds ratios in the PO model with complex survey design can be interpreted in the same way as those in the standard PO model.

The three cut points, when estimating the odds of being at or below a particular ordinal nutritional status category (based on weight), are used to differentiate the adjacent categories of the response variable (ordinal nutritional status). $\alpha_1 = -3.2418$, which is the first cut point for the cumulative logit model for $Y \leq 1$ that is level 1 versus levels 2-4; $\alpha_2 = 1.7371$ is the cut point for the cumulative logit model for $Y \leq 2$ that is levels 1 and 2 versus 3 and 4; $\alpha_3 = 2.9871$ is used as the cut point for the cumulative logit model when $Y \leq 3$, that is levels 1-3 versus level 4.

The results (Table 2) revealed that estimated logit coefficients of current age of child, female children, large and average weight of a child at birth, mother's current age, mother's BMI, mothers who are not married and Affar, Dire Dawa, Gambela, Harari and Somali regions were significant. Therefore, for the predictor, current age of child ($\beta = -0.3186$, $\text{OR} = 0.7271$) indicates that the odds of being at or beyond a particular ordinal nutritional status category decreased by $(1 - 0.7271) \times 100\% = 27.29\%$ with a one year increase in current age of child, holding other variables constant; female child ($\beta = -0.2417$, $\text{OR} = 0.7852$) suggesting that the odds of female child being at or beyond a particular under five ordinal nutritional status (based on weight) decreased by $(1 - 0.7852) \times 100\% = 21.48\%$. The odds of being at or beyond a particular ordinal nutritional status for weight of child at birth: large ($\beta = 0.5481$), and average ($\beta = 0.3134$) were 1.7301, and 1.3680, respectively times the odds of small weight of a child at birth; for the predictor mother's age ($\beta = -0.0203$, $\text{OR} = 0.9798$) indicates that the odds of being at or beyond a particular ordinal nutritional status category decreased by $(1 - 0.9798) \times 100\% = 2.02\%$ with a one year increase in mother's age; for the predictor mother's BMI ($\beta = 0.0479$, $\text{OR} = 1.0491$) indicates that a one-unit increase in mother's BMI, holding other variables constant, the odds of being at or beyond a particular under five ordinal nutritional status increased by 4.91%. It was found that the odds of being at or beyond a particular ordinal nutritional status for children born to unmarried mother was 0.7101 ($\beta = -0.3423$) times the odds for children born

tomarried mother. The odds of being at or above a particular ordinal nutritional status for children from Affar, Dire Dawa, Gambela Harari and Somali regions were re-

spectively OR = 0.6411 ($\beta = -0.4445$), OR = 0.5554 ($\beta = -0.5879$), OR = 0.5422 ($\beta = -0.6120$), OR = 0.7243 ($\beta = -0.3224$) and OR = 0.4007 ($\beta = -0.9143$) times the odds for children from Oromia region (see Table 2).

Table 2: Comparison of the PO models without and with complex survey design (CSD)

Parameters	PO model without CSD				PO model with CSD			
	Coeff.	SE	OR	P-val	Coeff.	SE	OR	P-val
cut1	-2.774	0.238			-3.242	0.357		
cut2	2.289	0.235			1.737	0.350		
cut3	3.570	0.239			2.987	0.363		
Current age of child	-0.323	0.020	0.724	0.0000	-0.319	0.027	0.727	0.0040
Sex of child (ref. Male)								
Female	-0.268	0.054	0.765	0.0000	-0.242	0.083	0.785	0.0000
Weight of child at birth (ref. Small)								
Large	0.555	0.074	1.741	0.0000	0.548	0.122	1.730	0.0000
Average	0.378	0.069	1.459	0.0000	0.313	0.100	1.368	0.0020
Mother's age	-0.008	0.005	0.992	0.0660	-0.020	0.006	0.980	0.0020
Mother's BMI	0.060	0.008	1.062	0.0000	0.048	0.014	1.049	0.0000
Mother work status (ref. No)								
Yes	0.042	0.062	1.042	0.5060	-0.033	0.096	0.967	0.7310
Educational attainment of mother (ref. No education)								
Primary	-0.007	0.069	0.993	0.9180	-0.065	0.097	0.937	0.5010
Secondary	0.120	0.117	1.128	0.3040	0.092	0.165	1.096	0.5770
Higher	0.268	0.149	1.307	0.0730	0.070	0.233	1.073	0.7640
Current marital status (ref. Married)								
Not married	-0.140	0.112	0.869	0.2100	-0.342	0.152	0.710	0.0250
Religion (ref. Orthodox)								
Catholic	-0.145	0.344	0.865	0.6740	-0.247	0.360	0.781	0.4920
Muslim	0.038	0.105	1.038	0.7200	0.216	0.138	1.241	0.1190
Protestant	0.008	0.091	1.008	0.9260	0.064	0.118	1.066	0.5910
Other	0.036	0.221	1.037	0.8700	0.686	0.626	1.986	0.2740
Region (ref. Oromia)								
Addis Ababa	0.228	0.155	1.256	0.1410	0.240	0.217	1.271	0.2690
Affar	-0.395	0.122	0.674	0.0010	-0.445	0.148	0.641	0.0030
Amhara	-0.135	0.124	0.873	0.2730	-0.066	0.134	0.936	0.6210
Benishangul-Gumuz	-0.224	0.118	0.799	0.0570	-0.270	0.158	0.763	0.0880
Dire Dawa	-0.600	0.151	0.549	0.0000	-0.588	0.152	0.555	0.0000
Gambela	-0.445	0.143	0.641	0.0020	-0.612	0.185	0.542	0.0010
Harari	-0.335	0.140	0.715	0.0160	-0.322	0.156	0.724	0.0390
SNNP	0.213	0.107	1.238	0.0470	0.167	0.138	1.181	0.2280
Somali	-0.899	0.112	0.407	0.0000	-0.914	0.165	0.401	0.0000
Tigray	-0.239	0.128	0.787	0.0620	-0.178	0.145	0.837	0.2180
Place of residence (ref. Rural)								
Urban	0.080	0.098	1.083	0.4150	0.204	0.202	1.227	0.3120
Wealth index (ref. Poor)								
Middle	-0.036	0.083	0.965	0.6690	-0.017	0.110	0.983	0.8770
Rich	0.128	0.078	1.136	0.1010	0.191	0.115	1.210	0.0980
Iteration 0: log likelihood = -6362.3952 LR χ^2 (28) = 706.45						F(28, 588) = 13.01		
Iteration 4: log likelihood = -6009.1723 Prob > χ^2 = 0.0000						Prob > F = 0.0000		

To estimate the odds of being at or below a particular ordinal nutritional status category compared with being at or above that category, we need to reverse the signs before the cut points and the logit coefficients into equation (4) resulted in $\text{logit}[P(Y \leq y_{-j} | x)] = \beta_{-j}0 + (-\beta_{-j} x)$. Odds ratios (Table 2) can be used directly to the analysis of the odds of being beyond a particular ordinal nutritional status category for significant effects.

Comparison of results

Table 2 provides the results of the two models, the fitted classical PO model and thereafter PO model with complex sampling design. After complex sampling design was applied to the PO model, the estimated logit coefficients and their standard errors were different from those in the PO model under the simple random sampling assumption. The logit coefficient of the predictors current age of child, female child, Dire Dawa and Harari regions were increased and those of the other significant predictors (large and average weight of child at birth, mother's age, mother's BMI, not married mothers and Affar, Gambela and Somali region) were decreased.

Compared to the PO model without complex survey design, the estimated logit coefficient for current age of child in the PO model with complex survey design increased by 1.48%, and its standard error increased by 37.7%; the logit coefficient for female child increased by 10.92%, and its standard error increased by 55.31%; the logit coefficient for Dire Dawa and Harari region were respectively increased by 2.12% and 4%, with their standard error increased by 0.99% and 11.53%; the logit coefficient for large and average weight of child at birth were respectively decreased by 1.2% and 20.48%, with their standard error increased by 65.2% and 46.35%; the logit coefficient for mother's age, mother's BMI and not married mothers were respectively decreased by 40.9%, 79.3% and 41.01%, with standard error increased by 42.2%, 65.8% and 35.6%; and the logit coefficient for Affar, Gambela and Somali region were respectively decreased by 11.1%, 27.2% and 1.8%, with their standard error increased by 21.3%, 29.2% and 47.27%.

Further, the standard errors of the significant coefficients in the PO model with complex sampling design were higher as compared to the corresponding standard errors of the significant coefficients in the conventional PO model

indicating that standard errors were underestimated when we considered the conventional PO model^{48,3}. This is an important distinguishing feature between the models. Analyses ignoring the complex sampling design will lead to a false increased precision and should be avoided.

Conclusion

Therefore, policymakers need to focus on the influence of these significant factors to develop strategies that enhance the normal or healthy weight status of under-five children in Ethiopia. This study also suggests that improving the nutritional status of mothers will consequently improve the nutritional status of their children. Improving the work status of the mothers will enhance the mother's economic status and consequently improve the basic needs of their children. To change weight-related disorders, changes related to children, environmental and social intervention is required to promote and support weight-related change in mothers. The government of Ethiopia needs urgent implementation of programs targeted to the regions of Affar, Dire Dawa, Gambela, Harari and Somali to develop the strategies of enhancing the good nutritional status of under-five children in Ethiopia.

Future direction

It must be borne in mind that this study was conducted based on certain socioeconomic and environmental factors. Further research is hence needed to unravel the specific socio-economic and environmental factors and determine whether they serve as an influential factor that affects the malnutrition status of under-five children and enhance the findings in this study. In a further study, we will extend this study by considering non-parametric and semi-parametric approaches to ordinal logistic regression, Spatial-temporal analysis, and other advanced statistical models. In addition, we will try to identify the trends of malnutrition status of the under-five children using the available EDHS survey results.

Ethics approval and consent to participate

Ethical clearance for the survey was provided by the EHNRI Review Board, the National Research Ethics Review Committee (NRERC) at the Ministry of Science and Technology, the Institutional Review Board of ICF International, and the CDC. Publisher's Note: Spring-

er Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Availability of data and material

The data used for this study can be obtained by requesting the ORC macro and DHS.

Competing interests

The authors declare that they have no competing interests.

Funding

This work was supported through the DELTAS Africa Initiative. The DELTAS Africa Initiative is an independent funding scheme of the African Academy of Sciences (AAS)'s Alliance for Accelerating Excellence in Science in Africa (AESA) and supported by the New Partnership for Africa's Development Planning and Coordinating Agency (NEPAD Agency) with funding from the Wellcome Trust [grant 107754/Z/15/Z], DELTAS Africa Sub-Saharan Africa Consortium for Advanced Biostatistics (SSACAB) programme] and the UK government. The views expressed in this publication are those of the author(s) and not necessarily those of AAS, NEPAD Agency, WellcomeTrust or the UK government.

Acknowledgements

We thank, with deep appreciation, ORC Macro and Measure DHS for giving us access to the data file.

Authors' contributions

AAY acquired the data, performed the analysis, and drafted the manuscript. AAY, HGM, DGA, and SFM designed the research problem. All authors discussed the results and implications and commented on the manuscript at all stages. All authors contributed extensively to the work presented in this paper. All authors read and approved the final manuscript.

References

1. Black, R.E., et al., Maternal and child undernutrition and overweight in low-income and middle-income countries. *The Lancet*, 2013. 382(9890): p. 427-451.
2. De Onis, M., et al., Levels and trends in child malnutrition. UNICEF-WHO-The World Bank Joint Child Malnutrition Estimates. 2012.
3. Habyarimana, F., et al., Spatial Distribution of Deter-

- minants of Malnutrition of Children under Five Years in Rwanda: Simultaneous Measurement of Three Anthropometric Indices. *J Hum Ecol*, 2016. 54(3): p. 138-149.
4. Getahun, Z., et al., Review of the status of malnutrition and trends in Ethiopia. *The Ethiopian Journal of Health Development (EJHD)*, 2017. 15(2).
5. Hammer, L.D., et al., Standardized percentile curves of body-mass index for children and adolescents. *American Journal of Diseases of Children*, 1991. 145(3): p. 259-263.
6. Pietrobelli, A., et al., Body mass index as a measure of adiposity among children and adolescents: A validation study. *The Journal of Pediatrics*, 1998. 132(2): p. 204-210.
7. Himes, J.H. and W.H. Dietz, Guidelines for overweight in adolescent preventive services: recommendations from an expert committee. The Expert Committee on Clinical Guidelines for Overweight in Adolescent Preventive Services. *The American Journal of Clinical Nutrition*, 1994. 59(2): p. 307-316.
8. WHO, Cancer pain relief: with a guide to opioid availability. 1996: World Health Organization.
9. Ratzan, S.C., G.L. Filerman, and J.W. LeSar, Attaining global health: Challenges and opportunities. Vol. 55. 2000: Population Reference Bureau Washington, DC.
10. Asfaw, S.T. and L. Giotom, Malnutrition and enteric parasitoses among underfive children in Aynalem Village, Tigray. *Ethiopian Journal of Health Development*, 2000. 14(1): p. 67-75.
11. Genebo, T., et al., Factors contributing to positive and negative deviances in child nutrition. *The Ethiopian Journal of Health Development (EJHD)*, 2017. 12(2).
12. Genebo, T., et al., The association of children's nutritional status to maternal education in Zigbaboto, Guragie Zone, Ethiopia. *Ethiop. J. Health Dev.*, 1999. 13(1): p. 55-61.
13. Teller, C.H. and G. Yimer, Levels and determinants of malnutrition in adolescent and adult women in Southern Ethiopia. *Ethiopian Journal of Health Development*, 2000. 14(1): p. 57-66.
14. Kulasekaran, R.A., Influence of mothers' chronic energy deficiency on the nutritional status of preschool children in Empowered Action Group states in India. *International Journal of Nutrition, Pharmacology, Neurological Diseases*, 2012. 2(3): p. 198.
15. Yimer, G., Malnutrition among children in southern Ethiopia: Levels and risk factors. *Ethiopian Journal of Health Development*, 2000. 14(3).
16. Prah, M., Insights into gender equity, equality and

- power relations in sub-Saharan Africa. 2013: African Books Collective.
17. Ferro-Luzzi, A., et al., Seasonal energy deficiency in Ethiopian rural women. *European Journal of Clinical Nutrition*, 1990. 44(Supp.1).
 18. Taddese, Z., C.P. Larson, and J.A. Hanley, Anthropometric status of Oromo women of childbearing age in rural south western Ethiopia. *The Ethiopian Journal of Health Development (EJHD)*, 2017. 12(1).
 19. CSA, Ethiopia Demographic and Health Survey 2011. Addis Abeba, Ethiopia, and Rockville, Maryland, USA: CSA and ICF, 2012.
 20. CSA, Central Statistics Agency of Ethiopia and ORC Macro: Ethiopia demographic and health survey 2000. Addis Abeba and Calverton, MD: Central Statistics Agency and ORC Macro. 2000.
 21. CSA, Central Statistics Agency of Ethiopia and ORC Macro: Ethiopia demographic and health survey 2005. Addis Abeba and Calverton, MD: Central Statistics Agency and ORC Macro. [<http://www.measuredhs.com/pubs/pdf/FR179/FR179.pdf>]. 2006.
 22. CSA, Central Statistics Agency of Ethiopia and ORC Macro: Ethiopia demographic and health survey 2011. Addis Abeba and Calverton, MD: Central Statistics Agency and ORC Macro. 2012.
 23. Liu, X., Fitting stereotype logistic regression models for ordinal response variables in educational research (Stata). *Journal of Modern Applied Statistical Methods*, 2014. 13(2).
 24. Agresti, A., Wiley series in probability and statistics. Analysis of Ordinal Categorical Data, Second Edition. 2002: p. 397-405.
 25. Kleinbaum, D.G., Klein, M. and Pryor, E.R., Logistic regression: a self-learning text. 2002.
 26. Kleinbaum, D.G.a.K., M., Polytomous logistic regression. Springer New York, 2010: p. 429-462.
 27. Cliff, N.a.K., J.A., Ordinal measurement in the behavioral sciences. Psychology Press., 2003.
 28. Clogg, C.C.a.S., E.S. , Statistical models for ordinal variables Sage Publications, Inc., 1994. 4.
 29. Ishii-Kuntz, M., Ordinal log-linear models (Quantitative Applications in the Social Sciences). 1994. 97.
 30. Agresti, A., Analysis of ordinal categorical data. John Wiley & Sons., 2010. 656.
 31. Agresti, A.a.K., M. , Categorical data analysis. In International Encyclopedia of Statistical Science. *Springer Berlin Heidelberg*, 2011.
 32. Allison, P.D., Logistic regression using SAS: Theory and application. SAS Institute., 2012.
 33. Miller, P.W.a.V., P.A., On the determination of occupational attainment and mobility. *Journal of Human Resources*, 1985: p. 197-213.
 34. Ananth, C.V.a.K., D.G., Regression models for ordinal responses: a review of methods and applications. *International Journal of Epidemiology*, 1997. 26(6): p. 1323-1333.
 35. Long, J.S., Regression models for limited and categorical dependent variables. 1997.
 36. McCullagh, P., Regression models for ordinal data. . *Journal of the Royal Statistical Society. Series B (Methodological)*, 1980: p. 109-142.
 37. McKelvey, R.D.a.Z., W., A statistical model for the analysis of ordinal level dependent variables. *Journal of Mathematical Sociology*, 1975. 4(1): p. 103-120.
 38. Winship, C.a.M., R.D., Regression models with ordinal variables. *American Sociological Review*, 1984: p. 512-525.
 39. Long, J.S.a.F., J., Regression models for categorical dependent variables using Stata. Stata press., 2006.
 40. Liao, T.F., Interpreting probability models: Logit, probit, and other generalized linear models Sage., 1994(No. 101).
 41. McCullagh, P.a.N., J.A., Generalised linear modelling. Chapman and Hall: New York, 1989.
 42. O'Connell, A.A., Logistic regression models for ordinal response variables. *Sage.*, 2006. 146.
 43. Anderson, J. A. Regression and ordered categorical variables. *Journal of the Royal Statistical Society. Series B. (Methodological)*, 1984: 1-30
 44. Bender, R. a. B., A. Calculating ordinal regression models in SAS and S-Plus. *Biometrical Journal*, 2000. 42(6): 677-699.
 45. Brant, R. Assessing proportionality in the proportional odds model for ordinal logistic regression. *Biometrics*, 1990: 1171-1178.
 46. Lall, R., Campbell, M.J., Walters, S.J., Morgan, K. and MRC CFAS Co-operative. A review of ordinal regression models applied on health-related quality of life assessments. *Statistical Methods in Medical Research*, 2002. 11(1): 49-67.
 47. Liu, X. Applied Ordinal Logistic Regression Using Stata: From Single-level to Multilevel Modeling. *Sage Publications*, 2015.

48. Liu, X. a. K., H. Fitting proportional odds models to educational data with complex sampling designs in ordinal logistic regression. *Journal of Modern Applied Statistical Methods*, 2013. 12(1): 26.
49. Long, J. S. Regression models for limited and categorical dependent variables, 1997.
50. Sharon, L. L. Sampling: Design and analysis, 1999. Duxbury Press 190.
51. Sommerfelt, A. E. a. S., M.K. Childrens nutritional status, 1994.
52. Teller, C. H. and G. Yimer. Levels and determinants of malnutrition in adolescent and adult women in Southern Ethiopia. *Ethiopian Journal of Health Development*, 2000.14(1):57-66.
53. Hien, N. N. and N. N. Hoa. Nutritional status and determinants of malnutrition in children under three years of age in Nghean, Vietnam. *Pak J Nutr*, 2009. 8(7): 958-964.
54. Achadi, E., et al., Global nutrition report 2016: From promise to impact: Ending malnutrition by 2030. 2016: *International Food Policy Research Institute*.