## Variables

### Dependent variables

The three anthropometric variables are measured through z-scores for height-for-age, weight-for height and weight-for-age and are defined as: $Z_i = \frac{AI_i - \mu}{\sigma}$, where $AI_i$ is the child anthropometric indicator, $\mu$ and $\sigma$ refer respectively to median and standard deviation of the reference population.

The correlations between the three anthropometric variables are statistically significant (Table 1) and it is reasonable that a single composite index was created from HAZ, WHZ and WAZ using principal component analysis [17-19].The first component alone explains 65.7% (Table 1) of the total variation of all anthropometric indices and this is a significantly high enough figure to create a single index of under-nutrition as 0.81HAZ+0.55WAZ+0.15WHZ [20]. Therefore, the first component of principal component analysis was taken as a new composite index of under-nutrition which was classified as nutrition status (severely undernourished if z-score <-3, moderately undernourished if -3≤ z-score <-2.0 and nourished if z-score ≥ -2.0). Then, the transformed variable was re-coded into ordinary outcome as '1' = nourished, '2'= moderately undernourished and '3' =severely undernourished. The methodology for computing the indicators was based on the 2006 WHO Child Growth Standards [2].

### Explanatory variables

The selection of explanatory variables are theoretically driven that draw support from prior research with regard to factors affecting children's nutritional status. Previous researches are referenced in creating categories for naturally continuous and discrete variables [21-25] (Table 2).

## Statistics Analysis

## Ordinal Logistic Regression Model

Logistic regression serves to model a categorical dependent variable as a function of one or more independent variables. The dependent variable may have two or more categories. When it has more than two categories, it may be ordered or unordered. Proportional Odds Model is instrumental to model the ordinal dependent variable through defining the cumulative probabilities rather than the probability of an individual event. The proportional odds model estimates the odds of being at or below a particular level of the response variable. It considers the probability of that event and all events before it. The proportional odds model is the default ordinal logistic regression type provided by statistical software [17, 18]. The proportional odds model with the logit or log-odds of the first $i$ cumulative probabilities is modeled as a linear function of the explanatory variables as:

$$Logit\,[Y_j \le i | x_j] = \log\left[\frac{\pi_i(X_j)}{1-\pi_i(X_j)}\right] = \log it(\pi_i) = \alpha_i - X'_j \beta, i = 1,2,...3C-1; \; j = 1,2,..., n$$

If the proportional odds assumption is not met, then different models are needed to describe the relationship between each pair of outcome groups [20].

## Generalized Ordered Logit Model

The proportional odds assumption (**β** is independent of response level) may be too strict and needs testing. The generalized ordered logit model (GOLM) adopts the parallel lines assumption for all C outcome categories. It allows the slope coefficients to differ for each of C-1 binary regressions [20]. The GOLM retains the nature of the proportional odds model (POM) by considering simultaneously the effects of a set of independent variables across successive dichotomizations of the outcome [26].

## Partial Proportional Odds Model

A partial proportional odds or non-proportional odds model relaxes the assumption of proportional odds. When the proportional odds assumption applies does not apply to all of the covariates, the partial proportional odds model may be used. This model allows some co-variables to be modeled with the proportional odds assumption, but for variables where the assumption is not satisfied, the effect associated with each $i^{th}$ cumulative logit is increased by a coefficient $(\gamma)$ adjusted by the other co-variables. The PPOM as formulated by Peterson and Harrell [27] imposes constraints for parallel lines only where they are needed. Thus, in this study the partial proportional odds model (PPOM) was employed which is defined as:

$\log\left(\frac{pr(Y \le i \mid x)}{pr(Y > i \mid x)}\right) = \alpha_i - (X\beta_i + \tau\gamma_i), i = 1,...C-1,$ where x is vector containing the full set of independent variables, $\tau$ is a vector of subset independent variables that violate the parallel line assumption, and $\gamma_i$ the regression coefficients associated with $\tau$. The predicted probabilities of belonging to a certain category are then defined by taking the exponential and rearranging the above equation in both sides .

**Parameter Estimation**

All the models described above were fitted to the data set using STATA (version 14). The variable selection was purposeful and the first step is used to evaluate the parallel assumption. Firstly, a POM was fitted with command "ologit" and then the "Brant" test was performed to evaluate the parallel assumption. This test compares the beta coefficients from *C-1* binary logits and gives a list of variables violating the parallel assumption. The command "**gologit2**" was used to estimate the GOLM. The PPOM was estimated using the **gologit2** command with the **autofit** option to impose constraints on the variables where the parallel assumption was not violated.

All the ordinal logistic models are estimated through the procedure of maximum likelihood estimation. Maximum likelihood estimates are values of the parameters that have the "maximum likelihood" of generating the observed sample. The likelihood equations lead to the unknown parameters in a non-linear function. . The ordinal logistic regression model is fitted to the observed responses using the maximum likelihood approach. In general, the method of maximum likelihood produces values of the unknown parameters that best match the predicted and observed probability values.

**Model selection**

To compare the ordinal logistic models, the log-likelihood were calculated, and the model with a higher log-likelihood is taken as a better-fitting one. Models are compared based on the Akaike Information Criterion (AIC) and Baye's Information Criterion (BIC) [28] and the models with the smaller absolute AIC and/or BIC values are preferred. The best model is selected from a set of competing models that has lowest value of AIC and BIC,

**Test of overall model fit**

The null hypothesis for an overall model fit test may be stated as “all the regression parameters are zero” while the alternative hypothesis is “at least one regression coefficient (parameter) is not zero”. To keep use of the selected mode, significance of the individual parameters is examined; the null hypothesis must be rejected. In addition, the goodness of fit of a model is tested by the deviance statistic [29].