Ordinal Logistic Regression Analyses on Anemia for Children Aged 6 Months to 5 Years Old in the Philippines

Junero Gorospe[1]*, Simon Antoni Bismonte[1], Rechel G. Arcilla[1], Glen Gironella[2]
[1] *Mathematics Department, De La Salle University*
[2] *Food and Nutrition Research Institute, Department of Science and Technology*
**Corresponding Author: junero.gorospe@yahoo.com*

Abstract: Anemia is one of the most prevalent nutritional deficiencies in the world with one out of four people being considered anemic. Anemia is considered one of the factors that negatively affect preschool children and prevalence of anemia among these children is second to the highest here in the Philippines. It is well known that anemia leads to damaging consequences such as hindered physical and cognitive development and weakened immune system. In this study, hemoglobin level was analyzed through ordinal logistic regression models using the data from the 2008 National Nutrition Survey. The estimates generated strongly suggested that malnourished children were more likely to be anemic. Moreover, the results showed that children whose age is between 6-11 months required more attention since this age group was at most risk with anemia. Nutrient intake was found to be significant and the likelihood of developing anemia was lessened when proper diet was implemented. The ordinal logistic models were also compared to the multiple linear regression models and the multinomial logistic regression models. Finally, for assessing moderate to severe anemia level it was found that multinomial logistic model is more appropriate while an ordinal logistic regression was found better for mild anemia level.

Key Words: Anemia; Hemoglobin Level; Ordinal Logistic Regression

1. INTRODUCTION

One major concern in a child's health is the presence of anemia. Anemia is formally defined by World Health Organization (WHO) as a hemoglobin level below that which is considered normal for age, sex, physiological state, and altitude, without considering the cause of the deficiency. Note that altitude is considered in WHO definition since hemoglobin concentration increases at high altitude. This is one of the notable features of acclimation to high altitude (Baillie, 2007).

The South-East Asian region was found to have the greatest number of pre-school age children affected by anemia where 115.3 million or 65.5% of young children are positive to anemia. Being part of the South-East Asian region, anemia is likewise an ongoing public health concern in the Philippines. Despite the decrease of its prevalence from 1998 to

2008, data from the Food and Nutrition Research Institute (FNRI) shows that the overall anemia prevalence is still 19.5%. With respect to age, anemia is most prevalent among infants (6-11 months) at 55.7%. Anemic infants are considered to be a significant public concern under the epidemiological criteria for assessing severity and magnitude of anemia. Based on the 2008 Projected Population, there are 0.74 million anemic children among 6 months to <1 year and 2.10 million anemic children among 1-5 years old. For children with ages 6 months to 5 years old, anemia prevalence is at 23.6% and it is considered moderate (10-39%) in all regions. For ages 1 to 5, the anemia prevalence is 41.0%, 26.6%, 13.7%, 13.6% and 10.6%, respectively. Despite the decrease in prevalence as children get older, anemia is still considered a major problem (FNRI, 2008).

Hemoglobin is considered the primary factor used to determine anemic level. It is derived from the level of hemoglobin present in the blood, which is a metalloprotein found in the red blood cells. It accounts for about two-thirds of the iron found in the body. In addition, it is known for its role in carrying oxygen to organs and tissues. One-sixth of the remaining iron is stored as ferritin and is only used when dietary intake is insufficient. The remaining irons are found in protein tissues that help one's body to function (Center for Disease Control and Prevention, 2012). A test usually used by doctors is by measuring the hemoglobin level with machines running several different tests on a child's blood. It is considered relatively easy and inexpensive. The hemoglobin level of a child is expressed as the amount of hemoglobin in grams per deciliter of blood (gm/dl).

The anemia limits showed in Table 1 were published in 1968 by a WHO study group on nutritional anemia. The limits that classifies mild, moderate and severe anemia were first introduced in the 1989 book *Preventing and controlling anemia through primary health care*. Then adjusted for age groups of children and pregnant women which is presented in *The management of nutrition in major emergencies*. Overall limits remained unchanged since 1968 and had been validated through various publications.

Table 1. Hemoglobin levels to diagnose anemia of Children 6 months to 5 years old

Anemic Level	Hemoglobin Count
Non Anemia*	11 or higher
Mild Anemia*	10 to 10.9
Moderate to Severe Anemia*	9.9 or lower

SOURCE: WHO,1968 and WHO, 1989
*Hemoglobin in grams per deciliter

The main objective of the study is to assess the effects of the aforementioned determinants to hemoglobin level via several regression models. These models consist of a linear, multinomial logistic, and an ordinal logistic model. Comparisons of these models were also done.

2. METHODOLOGY

2.1 Data

The data used in this study was from the 7th National Nutrition Survey (NNS): Philippines, 2008. The survey was conducted by FNRI of the Department of Science and Technology. The 7th NNS sampling design utilized the 2003 Master Sample of the National Statistics Office. The NNS is composed of anthropometric, biochemical, clinical, dietary, socio-economic, food insecurity, government nutrition/health program participation and health surveys. The survey covered the 17 regions and the 79 provinces. The data consisted of 36,634 households from 3,377 enumeration areas with 191,316 individual observations.

From the complete data, only several determinants were considered in the study. The variables applied in the study were decided through nutritional literatures (WHO, NHLBI, Tengco, L et. al.), and recommendations of FNRI. The study considered 2484 observations for children aged 6 months to 71 months. Table 2 shows the description of the variables which were found significant among the fitted models. Some variables were omitted since the variables are sum of other variables which causes multicollinearity. Take note that categorical counter-parts of nutrient and food intake were included in the study and these variables are defined whether a child has met the nutrient intake requirement or not or if a child has eaten certain foods in the past two days or not.

2.2 Analysis

Ordinal logistic regression models were fitted on the data. The dependent variable was classified into three groups: severe anemia (2), mild anemia (1), and normal level (0). Stepwise selection procedure is used to form the fitted model. Contrasting the models was done by assessing which models are statistically significant and checking the diagnostics. For the logistic regression models, Pseudo R-squares were computed to compare the models. Variance inflation factor was used for both logistic models to test for multicollinearity. Finally, significance of the model was also checked using Chi-square test for logistic regression. Furthermore, two models were fitted for each regression model. As mentioned earlier, some numerical variables have categorical counterparts leading to multicollinearity. Thus, the first model included the categorical covariates while the second model consisted of the numerical covariates. Statistical inferences were done at 5% level of significance. The statistical software used in this study was Stata/S.E. v12.0.

3. RESULTS AND DISCUSSION

3.1 Ordinal Logistic Regression

Results were shown to provide an overview of hemoglobin count. Two ordinal logistic regression results will be presented and interpretations among the coefficients are deemed valid as long as other variables were held constant or controlled.

Table 2. Variable Names of the Determinants of Anemia

Variable Name	Description
Anthropometric Variables	
HAZ_CGS	Height for Age Z-scores
HEIGHTCAT1	Stunted Growth 0 – No 1 – Stunted
WEIGHTCAT1	Underweight Child 0 – No 1 – Underweight
Biochemical Variables	
DEFLOW	Vitamin A Deficient 0 – No 1 – Deficient
FEVER1	Fever 0 – Did not get sick 1 – Got sick last 30 days
URTI1	Coughs and Colds 0 – Did not get sick 1 – Got sick last 30 days
Dietary Variables	
NIAC_EAR	Meeting 80% Niacin RENI 0 – No 1 – Yes
RIBO_EAR	Meeting 80% Riboflavin RENI 0 – No 1 – Yes
THIA_EAR	Meeting 80% Thiamin RENI 0 – No 1 – Yes
VITC_EAR	Meeting 80% Vitamin C RENI 0 – No 1 – Yes
Mean One-day per Capita Nutrient Intake	
TOT_IRON	Iron Intake(mg)
TOT_NIAC	Niacin Intake(mg)
Mean One-day per Capita Food Consumption	
SF3AP	Corn and Product Intake(g)
Socio-Economic Variables	
AGE	Age of child in months
COMPU	Computer In Household
WSMCHNE	Washing Machine in Household

Estimates of the fitted ordinal logistic model are shown in Table 3. Children tend to be more anemic due to being stunted and underweight with odds of 1.3865 and 1.2300 respectively. Vitamin A deficient children also have a higher likelihood of being anemic. Moreover, sufficient riboflavin intake and upper respiratory tract infection are other factors that contribute to increased chances of having anemia. On the other hand, presence of appliances such as computer and washing machine in a household deflates the possibility of having childhood anemia. As the age of a child increases by a month, the odds of having anemia are decreased multiplicatively by 0.9574.. Multicollinearity among covariates is non-existent and the fitted model is considered significant with p-value<0.0001.

Table 3. Ordinal Logistic Regression Estimates (Model 1)

Variable	Coefficient	Odds Ratio
AGE	-0.0435 (0.0034)	0.9574
WEIGHTCAT1	0.3268 (0.1407)	1.3865
HEIGHTCAT1	0.2070 (0.1338)	1.2300
URTI1	0.2819	1.3256

	(0.1113)	
RIBO_EAR	0.3680	1.4448
	(0.1361)	
COMPU	-0.7294	0.4822
	(0.3538)	
WSMCHNE	-0.3632	0.6594
	(0.1717)	
FEVER1	-0.3711	0.6900
	(0.1317)	
VITC_EAR	-0.5785	0.5607
	(0.1385)	
NIAC_EAR	-0.4840	0.6163
	(0.1314)	
DEFLOW	0.3172	1.3733
	(0.1442)	
THIA_EAR	-0.3863	0.6796
	(0.1422)	
CONSTANT1	-0.6977	0.4977
	(0.1564)	
CONSTANT2	0.9199	2.5090
	(0.1621)	

NOTE: Numbers inside parentheses are standard errors of the estimates.

Table 4 presents another array of ordinal logistic estimates. For every month increase in age, it is found that the likelihood of having anemia is decreased. But this time, the multiplicative decrease in odds is 0.9585. Almost all of the appliances were dropped except for washing machine. Its presence reduces the odds of having anemia by almost 36%. Moreover, for every milligram increase in niacin and vitamin C the odds of having anemia are decreased multiplicatively by 0.7829 and 0.9734, respectively. Other factors that deflate the odds of anemia are increase in height z-score, having fever, and total intake of corn. Finally, having URTI1 is the only determinant that increases the risk of having anemia among children. Moreover, the model was found significant with p-value<0.0001 and no multicollinearity was found among the covariates.

Table 4. Ordinal Logistic Regression Estimates (Model 2)

Variable	Coefficient	Odds Ratio
AGE	-0.0424	0.9585
	(0.0037)	
HAZ_CGS	-0.1913	0.8259
	(0.0431)	
FEVER1	-0.3379	0.7133
	(0.1291)	
TOT_IRON	-0.0702	0.9322
	(0.0227)	
SF3AP	-0.0075	0.9925
	(0.0029)	
TOT_NIAC	-0.0408	0.9600
	(0.0183)	
WSMCHNE	-0.4454	0.6406
	(0.1641)	
URTI1	0.3137	1.3685
	(0.1096)	
CONSTANT1	-0.7436	0.4754
	(0.1558)	
CONSTANT2	0.8709	2.3891
	(0.1611)	

NOTE: Numbers inside parentheses are standard errors of the estimates.

3.2 Linear and Logistic Regression Models

The Table 5 and Table 6 below present the results of the fitted multiple linear regression models.

Table 5. Linear Regression Estimates (Model 1)

Variable	Coefficient	Standard Error
AGE	0.0206	0.0013
RIBO_EAR	-0.2050	0.0572
VITC_EAR	0.2613	0.0527
R_ZN	-0.1198	0.0556
HEIGHTCAT1	-0.2287	0.0498
WSMCHNE	0.1537	0.0655
VCD	0.1388	0.0531
MILK_PDS	0.1235	0.0551
NIAC_EAR	0.3565	0.0584
DEFLOW	-0.2447	0.0641
CONSTANT	10.6190	0.0705

Table 6. Linear Regression Estimates (Model 2)

Variable	Coefficient	Standard Error
AGE	0.0231	0.0015
SF18AP	0.0033	0.0016
SF27AP	-0.0028	0.0014
HAZ_CGS	0.0522	0.0244

TOT_VITA	-0.0001	<0.0001
TELEPN	0.1408	0.0484
SF19AP	0.0002	0.0001
SF3AP	-0.0187	0.0003
TOT_NIAC	0.0401	0.0096
WAZ_CGS	0.0984	0.0285
TOT_PROT	-0.0136	0.0033
TOT_IRON	0.0265	0.0092
CONSTANT	10.7237	0.0714

Both models fitted using multiple linear regression were significant with p-values <0.0001. However, both models violated the assumption of normality (Shapiro-Wilk's test, p-value <0.0001 for both models), homogeneity of variances (Breusch-Pagan/ Cook-Weisberg test, p-value <0.0001 for both models) and, independence (Durbin Watson test min(p_-,p_+)<0.0001 for both models). Only 20.32% of the variation of hemoglobin count was explained by the first model while the second model only accounts for 19.88% of the hemoglobin count variation. Moreover, neither of the models was found to violate the assumption of multicollinearity.

Table 7 and Table 8 present the results of the fitted multinomial logistic regression models.

Table 7. Multinomial Logistic Regression Estimates (Model 1)

Variable	Coefficient	Odds Ratio
1. Mild vs. Normal		
AGE	-0.0393 (0.0036)	0.9615
COMPU	-0.7206 (0.3558)	0.4865
TV	0.3869 (0.1703)	1.4724
WEIGHTCAT1	0.4519 (0.1417)	1.5713
URTI1	0.2864 (0.1274)	1.3316
ELEC1	-0.5670 (0.1855)	0.5672
VITC_EAR	-0.3314 (0.1463)	0.7179
FEVER1	-0.3409 (0.1518)	0.7111
THIA_EAR	-0.4167 (0.1518)	0.6592
Table 7. (continuation)		
CONSTANT	0.2711 (0.1980)	1.3114
2. Moderate to Severe vs. Normal		
AGE	-0.0643 (0.0069)	0.9377
DEFLOW	0.5612 (0.2281)	1.7528
WEIGHTCAT2	0.5794 (0.2141)	1.7850
ENER_EAR	0.7001 (0.2907)	2.0140
TELEPN	0.6437 (0.2004)	1.9035
VITC_EAR	-1.4306 (0.3096)	0.2392
NIAC_EAR	-0.8912 (0.2115)	0.4102
CONSTANT	1.0509 (0.2749)	2.8602

NOTE: Numbers inside parentheses are standard errors of the estimates.

Table 8. Multinomial Logistic Regression Estimates (Model 2)

Variable	Coefficient	Odds Ratio
1. Mild vs. Normal		
AGE	-0.0361 (0.0037)	0.9645
SF2AP	-0.0017 (0.0007)	0.9983
SF3AP	-0.0073 (0.0032)	0.9927
TOT_IRON	-0.0598 (0.0188)	0.9420
URTI1	0.3010 (0.1257)	1.3512
FEVER1	-0.3236 (0.1495)	0.7235
WAZ_CGS	-0.2163 (0.0585)	0.8055
ELEC1	-0.3722 (0.1449)	0.6892
CONSTANT	0.3035 (0.2166)	1.3546
2. Moderate to Severe vs. Normal		
AGE	-0.0692 (0.0081)	0.9331
TELEPN	-0.5044	0.6039

	(0.2072)	
SF10AP	-0.0156 (0.0053)	0.9845

Table 8. (continuation)

HAZ_CGS	-0.1982 (0.0710)	0.8202
TOT_NIAC	-0.2448 (0.0729)	0.7829
SF3AP	-0.0184 (0.0068)	0.9818
TOT_VITC	-0.0270 (0.0094)	0.9734
SF19AP	-0.0044 (0.0012)	0.9956
TOT_RIBO	1.3828 (0.4167)	3.9860
TOT_PROT	0.0653 (0.0244)	1.0675
TOT_IRON	-0.1440 (0.0583)	0.8659
CONSTANT	0.6901 (0.1621)	1.9939

NOTE: Numbers inside parentheses are standard errors of the estimates.

Estimates of the Pseudo R-squares are presented in Table 9. For the multinomial logistic model, it can be seen that Model 2 is better compared to Model 1 for both categories since Pseudo R-squares were higher. In other words, for modeling a multinomial category variable for anemia, it is better to use the numerical counterparts of food intake, vitamin intake, and anthropometric measurements. In comparison to ordinal logistic regression, categorical counterparts of food intake, vitamin intake, and anthropometric measurements are more preferred. Finally, if the best models identified for the two logistic models are compared, multinomial logistic model is preferred when assessing moderate to severe anemia level while ordinal logistic model is preferred when assessing mild anemia level.

Table 9. Psuedo R-square values of the Logistic Regression Models

Model	Multinomial	Ordinal
Model 1	Mild vs. Normal 0.1004 Moderate to severe vs. Normal 0.2607	0.1299
Model 2	Mild vs. Normal 0.1015 Moderate to severe vs. Normal 0.2980	0.1253

4. CONCLUSIONS

In this study, hemoglobin is measured numerically and then categorized into normal, mild, and moderate to severe anemia levels. Several regression models were utilized to assess the effect of several determinants on preschool children anemia. Consistent determinants of anemic level are age, nutrient intake, and anthropometric measures. From these results, younger preschool children easily develop anemia, whether it is mild or moderate to severe. Furthermore, stunted and underweight children tend to acquire anemia faster compared to children who are considered normal, tall, or overweight. Proper nutrient intake is also recommended as this may hamper development of anemia. However, all vitamin supplement and majority of the food intake covariates are regarded as insignificant. Thus, vitamins supplementation and daily food intake does not essentially affect the anemic level of a child. For instance, a study has shown that the impact of vitamin supplementation regarding anemia is not clear. This includes possible interaction among multiple vitamins (Fishman et al., 2000). Contradictions among models such as presence of fever or a normal-weight child may be present due to lack of interaction variables. Including interaction among variables may arise to new possibilities and conclusions not yet discovered (Box, 1990). Overall, it can be derived from these results that malnourished children, especially those who belong to the lower age group, are at risk with respect to anemia.

The results of the study suggest a number of implications of interest to both policymakers and health planners. They should focus on giving proper education on parents regarding proper diet as nutrients play a vital role on preventing anemia. Besides educating parents about nutrients, they should also be given additional knowledge about anemia. The government, on the other hand, should provide better food supplementation and food fortification. For health care services, nutritional assistance among children on a regular basis may also pave way to a better prevention of anemia. However, these ideas are easier to accomplish if the accessibility and quality of the programs were enhanced.

6. REFERENCES

Baillie, K. (2007). *Haemoglobin carries oxygen in the blood*. Retrieved August 28, 2013, from http://www.altitude.org/haemoglobin.php.

Begg, C.B., & Gray, R. (1984). Calculation of polychotomous logistic regression parameters using individualized regressions. *Biometrika*, 11-18.

Box, G. (1990). “Do interactions matter?”.*Quality Engineering*.Vol.2. 365-369.

Fishman, S., Christian P., & West K.P. (2000). The role of vitiamins in the prevention and control of anaemia. *Public Health Nutrition*. Vol.3. 125-148.

Food and Nutrition Research Institute-Department of Science and Technology (FNRI-DOST). 2010 .Philippine Nutrition Facts and Figures 2008. DOST Complex, FNRI Bldg., Bicutan, Taguig City, Metro Manila, Philippines.

National Heart, Lung, and Blood Institute. (2012). *What are the signs and symptoms of anemia?.* Rerieved April 16, 2013, from www.nhlbi.nih.gov/health/healthtopics/topics/anemia/signs.html.

Osorio, M. (2002). Determinant factors of anemia in children. *Jornal de Pediatria*. Vol 78. 269-275.

Tengco, L., Solon P., Solon J., Sarol Jr. J., & Solon, F. (2007). Determinants of anemia among preschool children in the Philippines. *Journal of the American College of Nutrition*. Vol. 27. 229-241.

World Health Organization. (1968). Nutritional anaemias. *World health organization technical report series*, 8-19.

World Health Organization. (2000). Major nutritional deficiency diseases in emergencies. *The management of nutrition in major emergencies*, 9-16.

World Health Organization. (2011). Haemoglobin concentrations for the diagnosis of anaemia and assessment of severity. *VMINS/ vitamin and mineral nutrition information system*, 1-5.

World Health Organization. (2013). *The determinants of health.* Retrieved April 23, 2013, from World health organization: http://www.who.int/hia/evidence/doh/en/index.html

World Health Organization, & Centers for Disease Control and Prevention Atlanta. (2008). WHO global database on Anemia. *Worldwide prevalence of anaemia 1993-2005*, 1-7.