Factors affecting adolescent nutritional status in Banda Aceh, Indonesia

Zurnila Marli Kesuma¹, Asep Rusyana², Latifah Rahayu³

¹,²,³Statistics Department, Faculty of Mathematics and Natural Science
Syiah Kuala University Banda Aceh 23111,
Indonesia

E-mail: kesumaku@yahoo.com

Abstract. The purpose of this study was to examine factors affecting nutritional status, anemia status, and adolescent blood pressure in Banda Aceh, Indonesia. A cross-sectional study was designed using interviewed questionnaire on nutritional status, knowledge about nutrition, and anemia status. National Health Research Data (Riskesdas) in 2013 shows the prevalence of obesity in adolescents 13-15 years based on Body Mass Index (BMI) was 8.3% and for very fat was reaches 2.5%. Although these prevalences are still lower than the national prevalence, which is 8.2% and 1.7%, there are still many other issues that need serious attention related to nutrition in adolescents. For this reason, we use adolescent BMI to see adolescent nutritional status. Moreover, to find out the factors affecting adolescent nutritional status, multivariate regression analysis is used. Before estimating parameters by multivariate regression, predictor variables were reduced using factor analysis and five new factors were generated from the reduction. The results shows that factors affecting the nutritional status are dietary factors, pocket money and age factors, and knowledge about nutrition. Factors affecting anemia Status are allowance and age, income, and a number of family member. While the factors that affect the blood pressure is allowance and age and the number of family member.

1. Introduction

Young people are long-term human resources and future generations. According to the Central Statistics Agency (BPS), Indonesia’s youth population in 2015 reached 18.69% of the total Indonesian population. For this reason, health problems experienced by adolescents is a serious matter because it involves the quality of human resources in country. [1]

Much of the attention has been on adolescent health. Adequacy of nutrition in adolescents will also affect the future towards the level of intelligence. Knowledge about nutritional adequacy for adolescent health is needed. Nutritional status, anemia status, and blood pressure are some examples of health indicators that need more attention to achieve optimal health. Basic Health Research Data (Riskesdas) of Indonesia shows the prevalence of fat in adolescents 13-15 years were 8.3% and the prevalence of obesity reached 2.5%. Although the prevalence of obese and overweight Aceh province is still under the national prevalence, which is 8.2% and 1.7%, many other issues need serious attention related to nutrition problems in adolescents. [2]

At this time, important nutrients should be given to prevent the occurrence of chronic diseases associated with nutrition in later adulthood. [3]

Another common problem in adolescents is anemia. Anemia is a condition where the hemoglobin (Hb) level is lower than normal. Anemia is influenced by diet, social condition, economy, culture, environmental health, immune system, health facilities, and others. [4] Anemia can cause various effects in adolescents, such as, lower body resistance that is so easily affected by the disease, decreased activity, and learning achievement. The prevalence of anemia in Indonesia reached 21.7%. Blood pressure is also a problem in adolescents. The occurrence causes several things of hypertension in adolescents. They are poor sleep patterns, unhealthy diet, obesity, and stress. WHO says the normal limit of a person’s blood pressure is between 120-140 mmHg for systolic pressure and 80-90 mmHg for diastolic pressure. A person will have hypertension problem if his blood pressure exceeds 140/90 mmHg. The results of
showed that the prevalence of hypertension in the age group of 15-24 years reached 8.7%. Based on this background, a study was conducted to find out the factors influencing nutritional status, anemia, and blood pressure of adolescents in Banda Aceh, Indonesia using multivariate regression.

2. Literature Review

2.1 Factor Analysis

Factor analysis is one of the multiple variable analyses that can be used to reduce data. [4] In this study, factor analysis is used to reduce the predictor variables that are considered to be correlated with each other. The basic principle of factor analysis is to extract a number of factors from the original variables X1, X2, X3... Xp into a number of factors that are much smaller than the number of origin variables X. There are several stages that are carried out to conduct a factor analysis, namely the Kaiser Meyer Olkin test (KMO) and multivariate correlation between variables that can be done with the Bartlett’s Test of Sphericity test.

Conducting correlation test between predictor variables. To test whether the predictor variables are correlated with each other. Testing the independence of response variables were done using the Bartlett Sphericity test, with the following hypothesis.

Ho: There is no correlation between response variables
H1: There is a correlation between response variables

Kaiser Meyer Olkin (KMO) Test were done using this formula:

\[
KMO = \frac{\sum_{i=1}^{p} \sum_{j=1}^{p} r_{ij}^2}{\sum_{i=1}^{p} \sum_{j=1}^{p} r_{ij}^2 + \sum_{i=1}^{p} \sum_{j=1}^{p} a_{ij}^2}
\]  

(1)

Where the notations are:
- \( I \): 1,2,3,...., p
- \( J \): 1,2,3,...., p
- \( r_{ij} \): The correlation coefficient between variables i and j
- \( a_{ij} \): The partial correlation coefficient between variables i and j

2.2 Bartlett’s Test of Sphericity

Bartlett’s Test of Sphericity is a statistical test used to test whether variables are mutually independent or not. The hypothesis used is.

Ho: R = I (the predictor variable is independent)
H1: R ≠ I (the predictor variable is dependent)

With test statistics as follows:

\[
\chi^2 = \left( n - 1 - \frac{2q+5}{6} \right) \ln |R|
\]

(2)

Where the notations are:
- \( q \): Number of predictor variables (X)
- \(|R|\): Determinant value of the correlation matrix of each predictor variable
- \( n \): Number of observations

Furthermore factor analysis will reduce the variables to several factors. Predictions or allegations of shared factor values that correspond to observations with a certain originator value are referred to as factor scores for these observations. The factor score estimation method in this study uses the regression method. This method predicts factor scores for observations that match xi using the same formula as the linear regression coefficient formula. [5]
2.3 Multivariate Regression Analysis

The multivariate linear regression model is a linear regression model consisting of more than one response variable (Y) and consists of one or more predictor variables (X). Suppose that there is a response variable of q and a predictor variable of p, then a multivariate regression model for observing the j-th response where \( j = 1, 2, ..., q \), can be represented by the following equation.

\[
\begin{align*}
Y_{i1} &= \beta_{01} + \beta_{11}X_{i1} + \cdots + \beta_{p1}X_{ip} + \varepsilon_{i1} \\
Y_{i2} &= \beta_{02} + \beta_{12}X_{i1} + \cdots + \beta_{p2}X_{ip} + \varepsilon_{i2} \\
&\vdots \\
Y_{iq} &= \beta_{0q} + \beta_{1q}X_{i1} + \cdots + \beta_{pq}X_{ip} + \varepsilon_{iq}
\end{align*}
\]

(3)

Where the notations are:
- \( Y_q \): The q-response variable
- \( X_p \): The p-predictor variable
- \( \beta_p, \beta_{p1}, \beta_{p2}, ..., \beta_{pq} \): The regression parameter for \( X_p \)
- \( \varepsilon_{iq} \): The error

A multivariate regression model consisting of q linear model can be written in matrix form like the following equation.

\[
Y_{(n\times q)} = X_{n\times(p+1)}B_{(p+1)\times q} + \varepsilon_{(n\times q)}
\]

(4)

Where the error in the above equation is assumed to have the following properties:
1. \( \text{E}(\varepsilon_i) = 0; \)
2. \( \text{Cov}(\varepsilon_i, \varepsilon_j) = \sigma_{ij}I \), where \( ij = 1, 2, ..., q \).

3. Research Methodology

The cross-sectional study was conducted during April-May 2017. The 289 male and 307 female adolescents aged between 10-19 years, who were studying in 30 schools in Banda Aceh, and agreed to participate were recruited. We employed a cluster sampling technique. The process of sampling was conducted using the following steps. The list of all schools from 3 types (elementary school (SD), junior high school (SMP), senior high school (SMA)) and the total students were obtained from the Central Bureau of Statistics (BPS). The cumulative frequency was computed. For feasibility reasons, 30 schools were chosen using probability proportional to size. The 20 students were randomly selected from the school until the sample size was achieved.

The variables used in the study are BMI as nutritional status, anemia and blood pressure. While the predictor variables are knowledge, income of parents, pocket money, eating patterns (morning, afternoon, night), age, and number of dependents. The explanation of the variables used in this study can be seen in the following table.
### Table 1. Variables used in the study

| Variables          | Information                                                      |
|--------------------|------------------------------------------------------------------|
| **Response Variables** |                                                                 |
| $Y_1$ Body Mass Index (BMI) | The results of the calculation of body weight and height         |
| $Y_2$ Anemia Status (g/dL)  | Adolescent Hemoglobin Levels (g/dL)                             |
| $Y_3$ Blood Pressure (mmHg) | Systolic blood pressure (mmHg)                                  |
| **Predictor Variables** |                                                                 |
| $X_1$ Knowledge about nutrition | Score from knowledge (score 1-10)                               |
| $X_2$ Parental income (Rp)    | Parents income in one month                                     |
| $X_3$ Pocket money (Rp)       | Pocket money given in one month                                 |
| $X_4$ Breakfast Pattern      | Frequency of breakfast in one week (0-7 times)                  |
| $X_5$ Lunch Pattern          | Frequency of lunch in one week (0-7 times)                      |
| $X_6$ Dinner Pattern         | Frequency of dinner in one week (0-7 times)                     |
| $X_7$ Age (year)             | Adolescent (10-19 years old)                                    |
| $X_8$ The number of dependents| Number of dependents in one family                             |

The research assistants viewed the questionnaires and gave their approval for the students to participate in the study. The research instrument was a self-administered, semi-structured questionnaire. Information was collected from respondents on socio-demographic characteristics, their eating habits and physical activity. Anthropometric measurements were carried out on the respondents (weight and height) by four experienced trained research assistants. Variables used in this study consisted of 3 response variables and 8 predictor variables that are considered scale. Response variables are such as BMI, blood pressure, and anemia. While, the predictor variables of knowledge, income parents, pocket money, diet (morning, afternoon, night), age, and the number of a family member were analyzed using bivariate and multivariate regression. The degree of association between dependent and independent variables were assessed.

For the next stage of data processing, the data used is the result of transformation. Using this data, we find out the factors affecting adolescent nutritional status with the following steps:

- Testing the correlation between the predictor variables to ascertain whether the predictor variables are correlated with each other.
- Reducing the predictor variables that are correlated with each other to new factors by factor Analysis.
- Estimating the parameters and testing the significance of the parameters of the multivariate regression model.

### 4. Results and Discussion

#### 4.1. Variable Reduction Process

Before conducting a multivariate regression analysis, the correlation test between the response variables is first performed. This test is carried out to see whether the response variables are correlated with each other or not. If the response variables are correlated with each other, then the data in the study can be analyzed using multivariate regression analysis.

Testing the freedom between response variables was done using the Barlett Spericity test, with the following hypothesis.

- $H_0$: There is no correlation between response variables
- $H_1$: There is a correlation between response variables

The result is showed in Table 2.
Based on the results of Barlett’s Sphericity Test in Table 2, we reject Ho which means there is a correlation between the response variables so that multivariate regression analysis can be performed. The results of the factor analysis can be seen in the following table.

### Table 3. Variable reduction results

| Variable                | Factor 1 | Factor 2 | Factor 3 | Factor 4 | Factor 5 |
|-------------------------|----------|----------|----------|----------|----------|
| Dinner (X₆)             | 0.735    | 0.129    | -0.067   | 0.105    | 0.309    |
| Lunch (X₅)              | 0.698    | 0.104    | -0.285   | -0.082   | 0.088    |
| Breakfast (X₄)          | 0.596    | 0.159    | 0.068    | 0.083    | -0.059   |
| Allowance (X₃)          | -0.236   | 0.663    | 0.254    | -0.057   | 0.481    |
| Age (X₇)                | -0.344   | 0.546    | -0.450   | 0.337    | 0.271    |
| Income (X₂)             | 0.134    | 0.532    | **0.640**| -0.297   | -0.171   |
| Number of dependents (X₈)| 0.069    | -0.295   | 0.503    | **0.774**| 0.131    |
| Knowledge (X₁)          | 0.052    | 0.594    | -0.146   | 0.373    | **-0.635**|

Table 3 shows the results of the factor analysis data reduction. From the results of the factor analysis 5 new factors were produced. Factor 1 is formed by the variable breakfast, lunch, and dinner with the largest contribution being the dinner variable. Factor 1 is named the **diet**. Factor 2 is formed by variable allowance and age with the largest contribution being allowance. Factor 3 is formed by the income variable. Factor 4 is formed by the variable number of dependents. And factor 5 is formed by the knowledge variable. After reducing the variables, a factor score or a new value will appear for each factor.

#### 4.2. Parameter Estimation

After analysing the factors and producing 5 new factors, the parameter testing will then be performed. Parameter testing was performed using factor scores. Table 4 shows the of estimation results of multivariate regression parameters using 5 factors.

### Table 4. Parameter estimation

| Predictor Variables | Parameter | B   | Std. Error | t     | p-value |
|---------------------|-----------|-----|------------|-------|---------|
| Y₁ (Nutritional Status) | Intercept | 4.561 | 0.021     | 217.571 | 0.000   |
|                     | F₁        | -0.091 | 0.021    | -4.356 | 0.000   |
|                     | F₂        | 0.110  | 0.021    | 5.247  | 0.000   |
|                     | F₃        | 0.020  | 0.021    | 0.963  | 0.336   |
From Table 4 a multivariate regression model is obtained:

\[
Y_1^* = 4.561 - 0.091 F_1 + 0.110 F_2 + 0.020 F_3 + 0.030 F_4 - 0.076 F_5 \\
Y_2^* = 3.592 + 0.009 F_1 + 0.052 F_2 - 0.032 F_3 - 0.025 F_4 - 0.002 F_5 \\
Y_3^* = 10.644 + 0.025 F_1 + 0.095 F_2 - 0.006 F_3 + 0.055 F_4 - 0.022 F_5
\]

Table 4 shows the significant factor to the model. For the nutritional status a significant factor was factor 1, namely diet, factor 2, namely pocket money and age, and factor 5, namely knowledge. For anemia the significant factor is factor 2, namely pocket money and age, factor 3 is income, and factor 4 is the number of dependents. As for the blood pressure variable the significant factor is factor 2, which is pocket money and age, and factor 4, which is the number of dependents.

From the significant factors it can be said that to increase the value of nutritional status of adolescents is by improving adolescent eating patterns, managing snack money for adolescents, and by improving adolescent knowledge about nutrition. To increase the value of adolescent anaemia status is to manage pocket money, manage income, and reduce the number of family dependents. Whereas to improve the adolescent blood pressure is to regulate allowances and family dependents.

4.3. Significant regression parameters

There are two tests conducted to determine the level of significance, namely the significant regression parameters to the model simultaneously and the significant regression parameters to the model partially.

4.3.1. Model simultaneously test

\[ H_0 : \beta_{11} = \beta_{12} = \beta_{13} = \ldots = \beta_{53} = 0 \text{ (model is not significant)} \]
\[ H_1 \text{ at least there is one } \beta_{ij} \neq 0 \text{ (model is significant), where } i = 1, 2, \ldots, 5 \text{ ; } j = 1, 2, 3 \]

Testing is done by comparing the value of Wilk’s Lambda.

\[
\Lambda = \frac{|E|}{|E + H|} = \frac{|Y^TY - \hat{\beta}^TX^TY|}{|Y^TY - \bar{y}^T\bar{y}|}
\]

\[
\Lambda = \frac{\begin{vmatrix} 156667 & 3732 & 14801 \\ 3732 & 47339 & 8440 \\ 14801 & 8440 & 260960 \end{vmatrix}}{\begin{vmatrix} 170911 & 5943 & 21514 \\ 5943 & 50028 & 10879 \\ 21514 & 10879 & 168804 \end{vmatrix}} = 0.839
\]

Based on the calculation, we obtained \( \Lambda_{\text{count}} = 0.839 \). Then we compare with \( \Lambda_{0.05,3,5,590} = 0.959 \), and we reject the hypothesis, and conclude there is at least one significant parameter to the model.

4.3.2. Model partially test
Partial parameter testing is also performed by comparing the value of Wilk’s Lambda, while the value of $\Lambda$ (count) for each predictor variable is shown in Table 5.

Table 5. Wilk’s Lambda value for partial testing

| Variable | P-value | $\Lambda_{\text{count}}$ | $\Lambda_{\text{table}}$ |
|----------|---------|--------------------------|--------------------------|
| F1       | 0.000   | 0.966                    | 0.987                    |
| F2       | 0.000   | 0.918                    | 0.987                    |
| F3       | 0.030   | 0.986                    | 0.987                    |
| F4       | 0.008   | 0.982                    | 0.987                    |
| F5       | 0.016   | 0.980                    | 0.987                    |

Tests are carried out for each predictor variable. The hypothesis for Factor 1 (F1) is as follows.

$H_0$: $\beta_{11} = \beta_{12} = \beta_{13} = 0$ (Regression parameter F1 no significant effect on the response variable)

$H_1$: at least one $\beta_{ij} \neq 0$ (Regression parameter F1 significantly influence the response variable)

Then we reject $H_0$ which means that the F1 regression parameters are significantly influence the response variable.

The partial testing of other predictor variables is carried out with the same criteria and test statistics. Wilk’s Lambda values for other predictor variables are also seen in Table 5. Based on Wilk’s Lambda values and p-values as shown in Table 5, it can be concluded that each predictor variable significantly influences the response variable.

4.4. Residual Assumption Test

In multivariate regression there are three assumptions that must be fulfilled, Identical Residual Assumptions, The assumption of identical residuals is carried out to see whether the residual has a homogeneous variance-covariance matrix. Assumption testing can be done using Box’s M test statistics with a hypothesis.

$H_0$: $\Sigma_1 = \Sigma_2 = \Sigma_3 = \Sigma$

$H_1$: at least one of $\Sigma_k \neq \Sigma_j$ untuk $k \neq j$

Test Statistics: $u = -2 (1-c_1) \ln M$

$u = -2 (1-0, 01220074) (-12.64737)$

$u = 24.98613$

$X_{table}^2 = X_{0.05,30}^2 = 43.77$

Because of the value $u < X_{table}^2$, it can be decided failing to reject $H_0$. This means that the residual has a homogeneous or identical variance-covariance matrix.

Testing residual assumptions is done using the Barlett Sphericity test with the following hypothesis.

$H_0$: Residual data is independent

$H_1$: Residual data is dependent

Table 6. Test results for independent residual assumptions

| Bartlett’s Test Sphericity | Chi-Square | df | P-value   |
|---------------------------|------------|----|-----------|
|                           | 7.452      | 3  | 0.059     |
Based on the test results shown in Table 6 it can be seen if the value of $X_{count}^2$ is 7.452. While the value of $X_{tab}^2$ for $X_{0.05,3}^2$ is 7.81 then it was decided to fail to reject H0, which means residual data is independent.

Testing the assumption of multivariate normally distributed residuals is carried out with the following hypothesis.

Ho: Residuals are normally distributed multivariate
H1: Residuals are not multivariate normal distribution

Test Statistics

\[ X_{tab}^2 = X_{0.05,3}^2 = 2.366 \]

| $i$ | $d_i^2$ | $i$ | $d_i^2$ | $i$ | $d_i^2$ |
|-----|---------|-----|---------|-----|---------|
| 1   | 3.520   | 10  | 0.553   | 19  | 2.620   |
| 2   | 2.786   | 11  | 0.963   | 20  | 2.289   |
| 3   | 1.352   | 12  | 0.585   | 21  | 15.775  |
| 4   | 8.570   | 13  | 0.215   | 22  | 1.538   |
| 5   | 2.209   | 14  | 0.394   | 23  | 6.105   |
| 6   | 0.866   | 15  | 4.632   | 24  | 1.245   |
| 7   | 0.859   | 16  | 9.680   | 25  | 0.460   |
| 8   | 1.318   | 17  | 1.479   | :   | :       |
| 9   | 1.082   | 18  | 1.609   | 596 | 3.154   |

Table 7 shows the value of $d_i^2$. Based on hypothesis test, we have the value of $d_i^2 \leq X_{0.05,3}^2$ for 326 observations or 54.70%. Therefore it was decided to fail to reject Ho, which means that the residuals were normally distributed multivariate.

Based on the result, we conclude that for multivariate model, all of the assumptions are met.

5. Conclusion

Based on multivariate regression model, we obtain the significant factor for adolescent nutritional status, anaemia and blood pressure are: For adolescent nutritional status a significant factor was diet, pocket money and age, and knowledge. For anaemia the significant factor is pocket money and age, income, and the number of dependents. As for the blood pressure the significant factor is pocket money and age, and number of dependents.

Acknowledgments

We would like to thank to Hibah Laboratorium Grant 2019 Syiah Kuala University for the financial support of his research and publication.

References

[1] Kesuma ZM and Rahayu L 2017 Identification of Adolescents Nutritional Status in Banda Aceh Statistika 17 63–69

[2] Health Department of Republic Indonesia 2013 Basic Health Research (Riskesdas) Health Department of Republic Indonesia 88-94

[3] Rusdiana S Rahayu L and Fradinata E 2018 Spatial Stochastic Modeling of Adolescent and Under-Five Children Nutritional Status: A Case Study from Aceh, Sumatera Island, Indonesia Proceedings of MICoMS (Emerald Reach Proceedings Series Emerald Publishing Limited, Bingley, 1 273-281

[4] World Health Organization 2004 Focusing on Anaemia: Towards an Integrated Approach for Effective Anaemia Control World Health Organization, Geneva 1-2
[5] Johnson RA, Wichern DW 2007 *Applied Multivariate Statistical Analysis, Sixth Edition* Prentice Hall. New Jersey 144-151