Analysis of Factors Affecting Adolescent’s Hypertension in Banda Aceh, Indonesia

L Rahayu¹, Munawar², Z M Kesuma³ and D Aini⁴

¹,²,³,⁴Statistics Department, Syiah Kuala University, Banda Aceh, Indonesia

latifah.rahayu@unsyiah.ac.id

Abstract. Hypertension is a heart and blood vessels disorder characterized by an increase of blood pressure. Hypertension is also known as the silent disease or the silent killer because the patients do not know they had hypertension or not before checking their blood pressure. This research aims to identify any factors that affect hypertension in adolescents and analyze the probability of hypertension in adolescents. The data in this study were obtained from the adolescent nutrition status survey in 2017 in Kota Banda Aceh. The method used to analyzed the data is binary logistic regression method. This method was chosen because the response variable is dichotomous, that is the events of hypertension for adolescents. The results of this research shows that the odds ratio of predictor variables that influenced the incidence of adolescent hypertension was 0.893 (0.835-0.955), 1.052 (1.015-1.091), and nutritional status 0.255 (0.069-0.942) with 95% confidence interval. Based on the model obtained can be known that the probability of hypertension for adolescents and the accuracy level of the model is 96.8%.

1. Introduction

Public health improvement could be one success indicators of the government in national development. The health change level will trigger transition epidemiology disease that is, a disease degenerative or noninfectious disease. One of the noninfectious disease is hypertension or high blood pressure [1]. Hypertension is degenerative diseases which becoming a serious problem that time. Hypertension are classified as the silent disease or the silent killer, because patients do not know they had incurable hypertension or not before they checked their blood pressure. In Indonesia, hypertension is the third biggest cause of death of all ages after stroke (15.4%) and tuberculosis (7.5%), with the number of reaches 6.8% [2].

Hypertension generally occurs in the elderly. But some research shows that hypertension can emerge since adolescents. Adolescent with hypertension is also a problem, because the hypertension can hold continues to adulthood and risk having higher morbidity and mortality. Although the prevalence of clinically very few for children and adolescents than adult, but hypertension in adults can be started in the infancy and adolescents. Research in [3], investigated the prevalence of hypertension with high school students in Sivas Province, Turkey using cross sectional. The result of the research is 4.4 % adolescents had hypertension. Hypertension can be influenced by several factors. The high blood pressure also deals with individual characteristics. Significantly, hypertension more experienced by adolescents between boys compared to girls.

Other research about adolescent nutritional status is identification of adolescent Nutritional Status in Banda Aceh, Indonesia. The result shows that there are 3.48% adolescent have malnutrition and 15.59% obesity in Banda Aceh, Indonesia. [4]
In this paper, we analyze the factors affecting adolescent’s hypertension with binary logistic regression. Binary logistic regression used to describe the nature of the relations between dichotomous predictor variables. The aim of this research is to identify the factors that affect hypertension in adolescents and analyze probability of hypertension in adolescents in Banda Aceh.

2. Methods
The data used are from adolescent nutrition status survey 2017 in Banda Aceh. The data consist of 29 schools in 9 districts in Banda Aceh with three levels school, they are 9 elementery schools, 9 junior high schools, and 11 senior high schools. The number of respondents to each school are 20-25 students, consist of boys and girls with the total of respondents are 600 students. The variables used are dependent and independent variables, detail in Table 1.

| Dependent variable | Measurement |
|--------------------|-------------|
| Y = Hypertension (mmHg) | 0= Non hypertension 1= hypertension |

| Independent variable | Measurement |
|----------------------|-------------|
| X₁ = Age (years) | Numeric : 10 – 19 |
| X₂ = Gender | 0 = Woman 1 = Boy |
| X₃ = Education Levels | 1 = Elementery School 2 = Junior high school 3 = Senior high school |
| X₄ = Allowance (Rupiah/month) | Numeric : (<150.000 – >1000.000) |
| X₅ = Height (cm) | Numeric : 114 – 190 |
| X₆ = Body Weight (kg) | Numeric : 20 – 109 |
| X₇ = Anemia (g/dl) | 0 = Non Anemia 1 = Anemia |
| X₈ = Nutritional Status (Kg/m²) | 0 = Normal 1 = Ubnormal |

The method used to analysis the data is binary logistic regression. The general form of regression logistic is

\[ \pi(x) = \frac{\exp(\hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \cdots + \hat{\beta}_k x_k)}{1 + \exp(\hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2 + \cdots + \hat{\beta}_k x_k)} \]  \hspace{1cm} (1)

The steps of binary logistic regression analysis are as follows:

2.1 Parameter estimation of logistic regression
The most commonly used method of parameter estimation in logistic regression is the Maximum Likelihood Estimation (MLE) method for obtaining maximum completion of the likelihood function. The principle of Maximum Likelihood method is to estimate \( \beta \) to maximize likelihood function.

\[ l(\hat{\beta}) = \prod_{i=1}^{N} \pi (x_i)^{y_i}[1 - (x_i)]^{n_i-y_i} \]  \hspace{1cm} (2)

where \( y_i \) are observations on \( i \)-th variables and \( \pi_i \) are Probability for \( i \)-th predictor variables. [5]

2.2 Significance parameter test
The test used to test the significant \( \beta \) coefficients of the model can be tested partially or simultaneously. Simultaneous test is done by using the likelihood ratio test. The hypothesis is:
H₀ : β₁ = β₂ = ⋯ = βₖ = 0, (simultaneously the predictor variable has no significant effect on the response variable)

H₁ : βₚ ≠ 0 (p = 1, 2, 3, …, k), (at least one predictor variable has a significant effect on the response variable)

Test Statistics used is G likelihood ratio test:

\[ G = -2 \ln \left( \frac{L_0}{L_1} \right) = -2 \ln \left( \frac{n_1}{n} \right)^{n_1} \left( \frac{n_0}{n} \right)^{n_0} \prod_{j=1}^{k} P_j^{y_j}(1 - P_j)^{1-y_j} \]  \hspace{1cm} (3)

where,

\( L_0 = \) Maximum likelihood function without predictor variable

\( L_1 = \) Maximum likelihood function predictor variable

\( n_1 = \) number of categorical observations 1

\( n_0 = \) number of categorical observations 0

\( n = \) the number of observation

\( y_i = \) the value of response variable in the \( i \)-th observation, with \( i = 1, 2, 3, \ldots, n \)

\( p_i = \) probability of observation \( j \)

The test statistic \( G \) follows the chi-square distribution with the rejection criterion \( G > \chi^2_{(df, \alpha)} \). [6]

On the other hand, Wald test is used to partial test with the following hypothesis

H₀ : \( \beta_p = 0 \), for \( p = 1, 2, \ldots, k \) (predictor variable \( p \) not significant with dependent variable)

H₁ : \( \beta_p \neq 0 \), for \( p = 1, 2, \ldots, k \) (predictor variable \( p \) significant with dependent variable)

with statistic test:

\[ W = \frac{\hat{\beta}_p}{se(\hat{\beta}_p)} \]  \hspace{1cm} (4)

To obtain a result, a comparison is made with the standard normal distribution (Z). The criterion to reject H₀ is \( W > Z_{\alpha/2} \).

2.3 Fit test of the model

The Hosmer-Lemeshow test (HL test) is a goodness of fit test for logistic regression, especially for risk prediction models. The Hosmer-Lemeshow test calculates if the observed event rates match the expected event rates in population subgroups. The hypothesis for model conformity test is formulated as follows.

H₀ : The model fits the data

H₁ : The model not fits the data

The statistic test if \( y = 1 \) is

\[ \hat{C} = \sum_{p=1}^{k} \frac{o_p - n_p \tilde{p}_{1p}}{n_p \tilde{p}_{1p}(1 - \tilde{p}_{1p})} \]  \hspace{1cm} (5)

Where \( \tilde{p}_{1p} \) is mean of probability success estimation \( p \), \( O_p \) is the number of successful event samples in the group \( p \). \( n_p \) is the number of sample group \( p \), and \( \sum_{p=1}^{k} n_p = n \) with \( p = 1, 2, \ldots, g \). Reject H₀ if \( \hat{C} > \chi^2_{g-2} \). [6]

2.4 Interpretation

The parameters in binary logistic regression can be interpreted using the odds ratio, each addition of one unit of \( x \) then the odds will increase by \( e^{\beta} \) times. In addition, model interpretation can also be done by analyzing the probability of influential variables in the model. The odds ratio symbolized by \( \psi \) [6]

\[ \psi = \frac{\pi(1)}{\pi(0)} \]  \hspace{1cm} (6)

Where odds ratio is comparison odds value for category \( x = 1 \) to category \( x = 0 \).
3. Results

**Figure 1(a).** Percentage of hypertension adolescent in Banda Aceh

**Figure 1(b).** Percentage of nutritional status in adolescent in Banda Aceh

From Figure 1. a. We can see that percentage of hypertension of adolescents in Banda Aceh only 3.33%, but there are 26.67% have prehypertension. Whereas, b. shows that percentage of adolescents with normal nutritional status (66%) is more than those with abnormal nutritional status (34%).

Binary logistic regression model is obtained from the $\beta$ coefficient value (Table 3.2). The model for hypertension data in adolescents in Kota Banda Aceh is as follows:

$$g(x) = 12.502 - 0.045x_1 - 0.648x_2(1) + 0.515x_3(1) + 0.542x_3(2) - 0.113x_5 + 0.051x_6 + 0.789x_7(1) - 1.368x_8(1)$$

**Annotation:**

- $X_1$ = Ages
- $X_8(1)$ = Nutritional status (unnormal)
- $X_2(1)$ = Gender (girl)
- $X_7(1)$ = Anemia
- $X_3(1)$ = School level (junior high school)
- $X_6$ = Body weight
- $X_3(2)$ = School level (senior high school)
- $X_5$ = High

The next step is to carry out a number of analysis test model with significance test parameters and the consistency of the test model.

The parameter significance test is simultaneously performed by likelihood ratio test. The Likelihood Ratio (G) test is used to determine if there are predictor variables that affect the model. Likelihood Ratio (G) distribution statistic is distributed $\chi^2$. The G values are presented in Table 2.

| Model     | Db | Chi-Square | P value |
|-----------|----|------------|---------|
| Hypertension | 9  | 30.711     | 0.000   |

The hypothesis test is as follows.

$H_0$: $\beta_1 = \beta_2 = \cdots = \beta_8 = 0$, (together the predictor variable has no significant effect on the hypertension variable)

$H_1$: There is at least one $\beta_p \neq 0$ ($p = 1, 2, 3..., 8$) (at least one predictor variable has a significant effect on hypertension variables)

Based on Table 2, obtained $p$ value of 0.000. By using significance level ($\alpha$) equal to 0.05 then $p$-value $<\alpha$ or 0.000 $<0.05$ so it can be concluded that there is at least one predictor variable that has significant effect on hypertension variable.
Table 3. Partial significant test parameters

| No | Variable          | \( \beta \) | Wald    | \( P \)-value | \( \text{Exp} (\beta) \) | Interval 95%          |
|----|-------------------|-------------|----------|---------------|-------------------|------------------------|
|    |                   |             |          |               |                   | Lower | Upper    |
| 1  | Ages              | -0.045      | 0.027    | 0.869         | 0.956             | 0.556 | 1.641    |
| 2  | Gender (1)        | -0.648      | 1.620    | 0.203         | 0.523             | 0.193 | 1.419    |
| 3  | School level      |             | 0.389    | 0.823         |                   |        |          |
|    | School level (1)  | 0.515       | 0.333    | 0.564         | 1.674             | 0.291 | 9.635    |
|    | School level (2)  | 0.542       | 0.117    | 0.732         | 1.720             | 0.077 | 38.416   |
| 4  | Allowance         | 0.000       | 2.597    | 0.107         | 1.00              | 1.00  | 1.00     |
| 5  | High              | -0.113      | 10.805   | 0.001*        | 0.893             | 0.835 | 0.955    |
| 6  | Body weight       | 0.051       | 7.624    | 0.006*        | 1.052             | 1.015 | 1.091    |
| 7  | Anemia (1)        | 0.789       | 2.094    | 0.148         | 2.201             | 0.756 | 6.408    |
| 8  | Nutritional status| -1.368      | 4.197    | 0.040*        | 0.255             | 0.069 | 0.942    |
|    | Constant          | 12.052      | 5.623    | 0.018         | 268,874,4         |        |          |

From Table 3 can be seen by looking at \( P \)-value at significance level \( \alpha = 0.05 \) there are 3 variables that have significant effect on hypertension variable that is height \( X_5 \), body weight \( X_6 \), and nutrient status \( X_8 \). For nutritional status variables significantly affect the category of nutrients abnormal. The result test of suitability model of hypertension data in adolescents in Banda Aceh detail in Table 4.

Tabel 4. Hosmer dan Lemeshow Test

| Hypertension | Chi-Square | db | \( P \)-value |
|--------------|------------|----|--------------|
|              | 14.505     | 8  | 0.245        |

Table 4. shows that with the value of degrees of freedom of 8, obtained Chi-Square value for model conformity test of 14.505. With \( \alpha = 0.05 \), then \( P \)-value > \( \alpha \) where 0.245 > 0.05, so it can be concluded that the model accordingly or model able to predict the value of observation. In addition to the model conformance test, there are other models of feasibility tools used in logistic regression, is Pseudo \( R^2 \) Test (Negelkerke \( R^2 \)). In this study the value Negelkerke \( R^2 \) is 0.197. That is, the predictor variable can explain the response variable of 19.7%.

The next step is to look at the accuracy of the classification of the model. Accuracy of classification is used to determine the level of goodness of the model in predicting a condition with real conditions. Percentage of classification accuracy of model that is equal to 96.8%. This means that overall binary logistic regression model that has been obtained can classify respondents correctly that is equal to 96.8%.

After performing the test of significance and fit test, the next step is to interpret the model. One of the ways to interpret coefficients in binary logistic regression is the odds ratio. Based on Table 3, we can see that value of odds ratio (\( \text{Exp} (\beta) \)) of each predictor variable. If the height of the respondent increases one centimeter (cm) it will give a multiplicative effect of 0.893 at odds \( Y = 1 \) (hypertension) and if the weight of the respondent increases one kilogram (kg) it will provide a multiplicative effect of 1.052 at odds \( Y = 1 \) (hypertension). Meanwhile, respondents who have abnormal nutritional status tend to have hypertension of 0.225 times compared with normal nutritional status.

Based on the RLB model can also be known probability of respondents who experience hypertension based on several factors. For example, want to know the chances of respondents experiencing hypertension when the age of 19 years, male sex, junior high school level / equivalent, have a monthly allowance of Rp. 500,000, 140 cm height, 60 kg weight, anemia, and abnormal nutritional status. The probability of respondents experiencing hypertension with the condition is as follows:

\[
g(\chi) = 12.502 - 0.045 (12) - 0.648 + 0.515 - 0.113(157) + 0.051(60) + 0.789 - 1.368 = -1.51
\]
So

\[ \pi(x) = \frac{e^{-1.51}}{1 + e^{-1.51}} = \frac{1.208}{2.208} = 0.54 \]

So respondents aged 12 years, male sex, junior high school level / equivalent, have a monthly allowance of Rp 500,000, height 140 cm, weight 60 kg, anemia, and abnormal nutritional status has a chance of experiencing hypertension is equal to 54%.

4. Conclusion
Based on the results, the conclusions of this research are:
1. Predicting variables affecting hypertension in adolescents in Kota Banda Aceh at the 0.05 significance level were height (X5), body weight (X6), and nutritional status (X8).
2. Based on the value of odds, if the height (X5) of the respondent increases one centimeter (cm) it will give a multiplicative effect of 0.893 and if the weight (X6) of respondents increases one kilogram (kg) it will give a multiplicative effect of 1.052 at odds Y = 1 (hypertension).
3. Based on the value of odds on nutritional status (X8) it can be concluded that respondents who have abnormal nutritional status have a risk of hypertension 0.255 times compared with normal nutritional status and nutritional status have a negative relationship with hypertension.

5. Acknowledgment
The author would like to thanks to Hibah Laboratorium 2019, Syiah Kuala University funding for supporting this research and publication.

6. References
[1] Indonesian Department of Health 2017 Profil Kesehatan 2016 Jakarta.
[2] Indonesian Department of Health 2013 Laporan Nasional Riset Kesehatan Dasar Jakarta.
[3] N Nur, S Cetinkaya, A Yilmaz, A Ayvaz, M O, Bulut and H Sumer 2008 Prevalence of Hypertension among High School Students in a Middle Anatolian Province of Turkey Journal Health Popul nutr vol. 26 no. 1 pp. 88–94.
[4] Z M Kesuma and L Rahayu 2017 Identifikasi status gizi pada remaja di kota Banda Aceh Statistika vol. 17, no. 2, pp. 63–69
[5] A Agresti 2002 Categorical data analysis New York: John Wiley and Sons. Inc.
[6] D W Hosmer and S Lemeshow 2000 Apllied Logistic Regression Canada: John Wiley and Sons. Inc.