Demographics Characteristics of Smoker in Poor Households in Riau Islands Province

D D Saputra

Badan Pusat Statistik Kabupaten Bintan, Toapaya, Bintan, Indonesia

*Corresponding author's e-mail: dio.saputra@bps.go.id

Abstract. Smoking habits in Indonesia have been formed since the colonial era. Smoking habits that need attention are in poor households. In 2020, Riau Islands Province as the one of youngest provinces in Indonesia has a smoking prevalence of 26.16% and the percentage of poor people is 5.92%. This condition is the basis for researchers to conduct a study that aims to determine the demographics characteristics of smokers. This study uses raw data from the National Socio-Economic Survey (SUSENAS) in Riau Islands Province in March 2020. The variables used are smoking status, gender, age group, education level, region, and recent migrant. The output of the processing stage is that the prevalence of smoking will be greater in the male population (OR = 132.04), the age group of 46-65 (OR = 4.77), the age group of 66 and over (OR = 2.11), the junior high school level (OR = 4.66), the senior high school level (OR = 5.98), the college level (OR = 3.13), living in the urban area (OR = 1.22) and the recent migrant (OR = 3.12). Thus, it is necessary to make a specific policy following the above characteristics in reducing smoking habits among poor households.

1. Preface

Cigarettes have become a health problem for almost all countries in the world, including Indonesia. The content of additives in it, makes smokers addicted and difficult to escape from smoking habits. Tobacco which is the main ingredient of cigarettes has quite a lot of substances that are harmful to health. Based on the results of previous research, the chemical content of tobacco that has been identified reaches 2,500 components, while in cigarette smoke there are 4,800 kinds of components [1]. From these chemical components that have been identified that are harmful to health are tar, nicotine, CO, and NO gases produced by tobacco plants, and some residual materials formed during planting, processing, and serving in trade, namely fertilizer and pesticide residues, TSNA (tobacco-specific nitrosamine), Bap (benzo-a-pyrene) and NTRM (non-tobacco related material).

In Indonesia, smoking was brought by the colonizers from Europe. The people of the archipelago, who previously had a tradition of chewing betel nut, considered the tradition of smoking tobacco as a new drug adopted from western society and then localized by adding various kinds of sauces and cloves to produce completely new products and customs [2]. This is not found anywhere, including in European communities and indigenous people of the Caribbean Islands and mainland America as the origin of this custom.

The description of smoking habits in Indonesia has been widely published by researchers, both in the form of books and scientific journals. One of the publications related to Indonesia's smoking habit is *Atlas Tembakau Indonesia 2020.* Some data presented in the book show that in 2018 the average prevalence of smokers of the population 15 years old and overreached 32.2% and as many as 52.1% of

smokers started smoking habits at the age of 15-19 years old. In addition, there has been a shift in the causes of death in Indonesia. Since 2017, the highest contributor to death has been non-communicable disease, in which the risk factor is behavior or lifestyle, one of which is smoking [3].

In addition to causing health problems, smoking habits also have an impact on the level of the economy. Bazotti *et al* examine the characteristics of the Brazilian population who spend their money on tobacco products. The variable used in their study were gender, education level, age (above 14 years), income level, region, and ethnicity [4]. Using a dataset from Brazil's Family Budgets Survey 2008-2009, it was found that on average, 10% of Brazilians spend their money on tobacco products. In addition, the characteristics of this population are older, lower-income, and less educated when compared to those who do not consume tobacco. This population also spends 1.5% of its family budget on buying tobacco products.

Haustein's research examines smoking and poverty. The results of his research stated that poor people spend up to 20% of their income on tobacco consumption (cigarettes) [5]. In addition, in many industrialized countries, the age at starting smoking is younger, which will increase the risk of developing smoking-related diseases. Research by Hosseinpoor *et al* analyzes the magnitude and pattern of socioeconomic inequality of smokers in low and middle-income countries. The results show that in many countries the poorest male population is 2.5 times more likely to smoke than the richest male population [6]. A smoker will also have the opportunity to be poor 8 times if the surrounding environment supports him to smoke [7].

Then, a study from Flint & Novotny uses the variables of sex, age, education, race, employment status, marital status, and geographic area in analyzing the independent relationship between poverty status and prevalence for smoking or quitting smoking in the United States in 1983-1993 resulted in the conclusion, people who are below the poverty line will be more likely to become current smokers than those who are in the poverty line or above. Poverty can be an indicator of lack of participation in changing social norms related to smoking behavior today [8].

Researchers have also studied the prevalence of smoking on migrant status. Hou *et al* stated that migrants returning to rural areas and migrants returning to urban areas were more likely to start smoking and less likely to quit smoking compared to non-migrant groups [9]. Meanwhile, research by Tong *et al* found that migrants tend not to start smoking and are less likely to become persistent smokers [10]. Bosdriesz *et al* also mention that the prevalence of smoking among migrants is lower than the group born in the United States and their country of origin [11].

Riau Islands Province (Kepri) as one of the youngest provinces in Indonesia has a fairly low average prevalence of smokers aged 15 years old and over. Based on data from Badan Pusat Statistik (BPS), in 2020 Riau Islands Province occupies the lowest position of the prevalence of smokers aged 15 years old and over among other provinces on the Sumatera Island, which is 26.16%. When viewed from the side of spending on cigarette consumption (in this case grouped into tobacco and betel nut), households in Riau Islands Province spend an average of 9.6% of their income on buying cigarettes. Meanwhile, if you look at the poverty level in the Riau Islands Province, the conditions are also better than other provinces on Sumatera Island. In the first semester of 2020, the percentage of poor people in Riau Islands Province was 5.92%, which put Riau Islands Province in the second-lowest position among other provinces on the Sumatera Island.

Based on the description above, this study aims to describe the demographic and economic characteristics of smokers in poor households in the Riau Islands Province. What are the characteristics of smokers in poor households in the province which generally have a low smoking prevalence and poverty rate.

2. Method

The data used in this study is raw data from the processing of the National Socio-Economic Survey (SUSENAS) in Riau Islands Province in March 2020. SUSENAS is a survey conducted by Badan Pusat Statistik (BPS) and is designed to collect population social data that is relatively very broad. The data collected in SUSENAS concerns the fields of education, health/nutrition, housing, other socio-economic aspects, socio-cultural activities, household consumption/expenditure and income, travel, and public opinion regarding the welfare of their household. The sampling design used in the SUSENAS is divided

into two, namely for district/municipality estimates and province estimates. Estimation of district/municipality using the two stages one phase stratified sampling and estimation of the province using two stages stratified sampling. The number of samples in the SUSENAS March 2020 are 345,000 households and for the Riau Islands Province, the number is 4,250 households. But from the target of households sample, only 4,114 households were declared clean and can be processed by the weighting factor using the population of March 2020 for the estimation of the district/municipality.

This study uses two analytical methods, namely descriptive and inferential analysis methods. Descriptive analysis is used to briefly describe the variables used in the study. While inferential analysis is used in binary logistic regression. This analysis is used to see the relationship between the dependent variable and a group of independent variables [12]. The difference between the linear regression model and the logistic regression model is that the dependent variable in the logistic regression is dichotomous/binary. The binary logistic regression model used is as follows.

$$g(x) = ln\left[\frac{\pi(x)}{1 - \pi(x)}\right]$$
$$= \beta_0 + \beta_i x_i$$
(1)

Description:
$g(x)$ = logit n(x)
$\pi(x)$ = probability to smoke
β_0 = constant
β_i = parameters to be estimated
x = independent variables
i = number of independent variables

The general method of binary logistic regression estimation is maximum likelihood estimation. In a general sense, the method of maximum likelihood yields values for the unknown parameters that maximize the probability of obtaining the observed set of data [12]. As the observations are assumed to be independent, the likelihood function is obtained as follows.

$$l(\beta) = \prod_{i=1}^{n} \pi(x_i)^{y_i}[1 - \pi(x_i)]^{1-y_i}$$
(2)

Description:
$l(\beta)$ = likelihood function for β parameter
$\pi(x_i)$ = probability to smoke
y_i = value of dependent variables (0 or 1)
x_i = independent variables
i = number of independent variables

The principle of maximum likelihood states that we use as our estimate of β the value that maximizes the expression in equation (2). However, it is easier to work with the log. So, the log-likelihood is defined as follows.

$$L(\beta) = ln[l(\beta)] = \sum_{i=1}^{n} \{y_i ln[\pi(x_i)] + (1 - y_i)ln[1 - \pi(x_i)]\}$$
(3)

This study only limits the analysis to poor households, namely households that have an average monthly expenditure per capita below the poverty line. The poverty line is calculated by adding up the value of the minimum need for food which is equivalent to 2,100-kilocalories per capita per day (Food Poverty Line) and the minimum need for housing, clothing, and health (Non-Food Poverty Line) [13]. The variables used are smoking status as the dependent variable, then gender, age group, education level, region, and recent migrant as independent variables. Age group variable in this study using the

classification of the Indonesian Department of Health (5-11 is child, 12-25 is teenager, 26-45 is adult, 46-65 is elderly, and 66+ is senior). Table 1 shows a breakdown of the categories of each variable.

**Table 1.** Categories of dependent and independent variables

Variables	Question on Form	Categories
Smoking status (Y)	P. 1206	1 = Smoking every day or sometimes
		0 = Others
Gender (X1)	P. 405	1 = Male
		0 = Female
Age Group (X2)	P. 407	1 = 5 – 11 years old
		2 = 12 – 25 years old
		3 = 26 – 45 years old
		4 = 46 – 65 years old
		5 = 66 years old and above
Education Level (X3)	P. 615	1 = Primary level
		2 = Junior high level
		3 = Senior high level
		4 = College level
Region (X4)	P. 105	1 = Urban
		0 = Rural
Recent Migrant Status (X5)	P. 604 & P. 605	1 = Recent Migrant
		0 = Others

The data processing stage starts from sorting the entire sample, namely by selecting only poor households. The number of samples of poor households is 359 households. While the number of individuals is 1,697 people. Furthermore, from the sample that has been selected, deletion is carried out for individuals under 5 years old. This is because point 615 on the SUSENAS questionnaire is only asked for household members aged 5 years and over. And the final sample used in this study is 1,508 people.

3. Discussion

3.1. Socio-economic characteristics of smokers in the Riau Islands Province

Before analyzing the binary logistic regression model, the research data was first carried out with descriptive analysis. Weighting is also applied before processing, both descriptive and inferential. The researcher uses a weighting because to obtain the results of the analysis at the population level. Figure 1 shows the percentage of smokers aged 5 years and over in poor households in the Riau Islands Province.
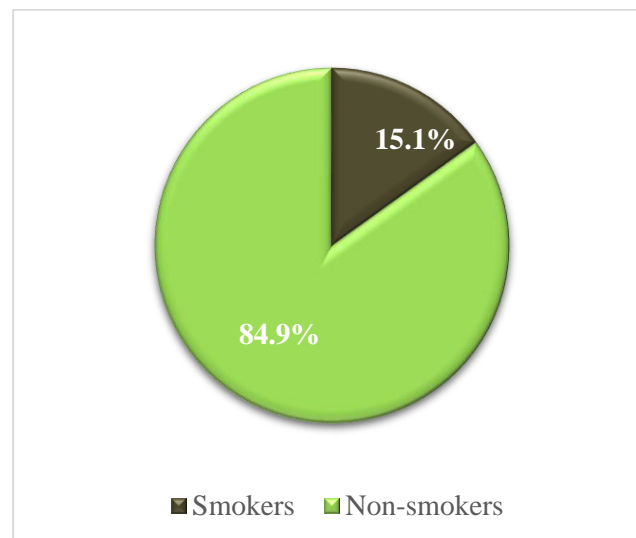
Source: Data processing results of SUSENAS March 2020

Figure 1. The percentage of smokers aged 5 years and over in poor households in the Riau Islands Province, March 2020

Based on Figure 1, it can be seen that in poor households in the Riau Islands Province, the percentage of the population who smokes and is aged 5 years and over is 15.1%. In addition, the characteristics of the population in the poor household aged 5 years and over can be seen in Table 2.

Table 2. Characteristics of the population in the poor household aged 5 years and over in Riau Islands Province, March 2020

Variables	Categories	Percentage	Total
Gender (X1)	Male	51.8	100
	Female	48.2	
Age Group (X2)	5 – 11 years old	23.7	
	12 – 25 years old	23.5	
	26 – 45 years old	33.4	100
	46 – 65 years old	14.5	
	66 years old and above	4.9	
Education Level (X3)	Primary level	67.5	
	Junior high level	12.4	100
	Senior high level	18.8	
	College level	1.4	
Region (X4)	Urban	79.0	100
	Rural	21.0	
Recent Migrant Status (X5)	Recent migrant	1.4	100
	Others	98.6	

Source: Data processing results of SUSENAS March 2020

Based on Table 2, it can be seen that the population aged 5 years old and over in poor households is dominated by the male population (51.8%). Then, the population in poor households is on average 26-

45 years old (33.4%), low education (67.5%), and living in urban areas (79%). In addition, the population in poor households is not recent migrants (98.6%).

Table 3. The percentage of smokers on the characteristics of the population in the poor household aged 5 years and over in Riau Islands Province, March 2020

Variables	Categories	Percentage		Total
		Smokers	Non-smokers	
Gender (X1)	Male	28.7	71.3	100
	Female	0.5	99.5	100
Age Group (X2)	5 – 11 years old	0.0	100.0	100
	12 – 25 years old	7.0	93.0	100
	26 – 45 years old	30.1	69.9	100
	46 – 65 years old	19.2	80.8	100
	66 years old and above	12.9	87.1	100
Education Level (X3)	Primary level	13.1	86.9	100
	Junior high level	19.0	81.0	100
	Senior high level	20.1	79.9	100
	College level	9.3	90.7	100
Region (X4)	Urban	15.9	84.1	100
	Rural	12.1	87.9	100
Recent Migrant Status (X5)	Recent migrant	21.7	78.3	100
	Others	15.0	85.0	100

Source: Data processing results of SUSENAS March 2020

Table 3 shows the percentage of smokers on the characteristics of the population in poor households aged 5 years old and over. From Table 3, it can be seen that in poor households, the male population is more likely to smoke. Then, the population in the range of 26-45 years old (adults), education in senior high school, and living in urban areas tend to be smokers. In addition, recent migrants in poor households also tend to smoke.

3.2. Binary-logistic regression analysis for socio-economic characteristics of smokers in the Riau Islands Province

After conducting a descriptive analysis, the next step is to analyze a binary logistic regression to find out more about the tendency of each independent variable to influence the decision to smoke. However, first, the parameter estimator will be tested simultaneously. The test used is the Omnibus Test. the purpose of this test is to determine the effect of all independent variables on the dependent variables simultaneously. The null hypothesis for this test is that there is no independent variable that can explain the prevalence to smoke in poor households in Riau Islands Province.

Table 4. The output of Omnibus Tests of Model Coefficients

	Step	Chi-square	df	Sig.
Step 1	Step	45291.495	10	0.000
	Block	45291.495	10	0.000
	Model	45291.495	10	0.000

Source: Data processing results of SUSENAS March 2020

Based on Table 4, it can be seen that the significance value is smaller than alpha (0.000 < 0.05), so the decision is rejecting H0, meaning that there is at least one independent variable that can explain the prevalence to smoke in poor households in Riau Islands Province.

After that, the Goodness-of-Fit test was carried out using the Hosmer and Lemeshow Test. This test aims to see whether the model formed is fit or not. The null hypothesis for this test is the model formed is fit.

Table 5. The output of Hosmer and Lemeshow Test

Step	Chi-square	Df	Sig.
1	1133.929	8	0.000

Source: Data processing results of SUSENAS March 2020

The output of the Hosmer and Lemeshow test as shown in Table 5 shows that the significance value is smaller than alpha (0.000 < 0.05), so the decision is rejecting H0, meaning that the model formed does not fit. Some researchers stated that there were several conditions where the binary logistic regression model tested was not fit, one of which was a large number of samples as in this study. When the sample size is very large, however, even minuscule discrepancies could lead to the rejection of the null hypothesis in the Hosmer-Lemeshow (HL) goodness-of-fit test for the logistic regression model [14]. The goodness-of-fit test should be considered as just one of several tools for assessing goodness-of-fit [15].

Therefore, the researchers tried to look at other alternatives in assessing the binary logistic regression model that was formed, namely by using a classification table. This table is used because it has similarities with the Hosmer and Lemeshow test. The Hosmer-Lemeshow goodness-of-fit test is used to assess whether the number of expected events from the logistic regression model reflects the number of observed events in the data [16]. While the classification table presents the degree to which predicted probabilities agree with actual outcomes [17].

Table 6. Classification Table

Observed			Predicted		
			Smoking status		Percentage Correct
			Smokers	Non-smokers	
Step 1	Smoking status	Smokers	11521	6533	63.8
		Non-smokers	7731	93895	92.4
	Overall Percentage				88.1

Source: Data processing results of SUSENAS March 2020

Table 6 shows a good overall correct prediction percentage, which is 88.1%. With the classification table, sensitivity, specificity, false positive, and false negative can be measured [17]. Sensitivity measures the proportion of correctly classified events, whereas specificity measures the proportion of correctly classified non-events.

After that, the next step is to do a partial test, namely by using the Wald test. This test is used to see which independent variables have a significant effect on smokers in poor households in the Riau Islands Province, as a dependent variable. These partial test results can be seen in Table 7.

Table 7. The output of Partial Test

Independent Variables	β Coefficient	S.E.	Wald	df	Sig.
Gender*	4.883	0.061	6335.721	1	0.000

Independent Variables	β Coefficient	S.E.	Wald	df	Sig.
Age Group (5-11)*	-	-	7923.506	4	0.000
Age Group (12-25)	-19.154	210.776	0.008	1	0.928
Age Group (26-45)*	-0.964	0.050	378.702	1	0.000
Age Group (46-65)*	1.563	0.046	1177.919	1	0.000
Age Group (>=66)*	0.746	0.048	239.095	1	0.000
Education Level (Primary)*	-	-	612.760	3	0.000
Education Level (Junior high)*	1.539	0.097	252.691	1	0.000
Education Level (Senior high)*	1.788	0.100	321.193	1	0.000
Education Level (College)*	1.140	0.098	136.496	1	0.000
Region*	0.201	0.028	51.630	1	0.000
Recent Migrant Status*	1.138	0.099	132.157	1	0.000
Constant*	-7.678	0.125	3795.779	1	0.000

* = significant variables

Source: Data processing results of SUSENAS March 2020

Based on Table 7, it can be seen which independent variables significantly affect the dependent variable. This can be seen from the significance value is smaller than alpha (0.000 < 0.05). So, that it can be concluded that the independent variables that have a significant effect on smokers in poor households in the Riau Islands ate gender, age group (categories: 5-11, 26-45, 46-65, and >=66), an education level (all categories), region and recent migrant status. It can be said that almost all independent variables have a significant effect, except the age group variable with categories 12-25 years old.

Then, the most important of logistic regression analysis is the interpretation of the odds ratio. The odds ratio is a measure of association between an exposure and an outcome [18]. Table 8 shows the odds ratio of each significant independent variable along with the value of the confidence interval.

Table 8. The output of Odds Ratio and Confidence Interval

Significant Independent Variables	Odds Ratio/Exp (β)	95% CI for Exp (β)	
		Lower	Upper
Gender	132.041	117.081	148.911
Age Group (26-45)	0.381	0.346	0.420
Age Group (46-65)	4.771	4.364	5.216
Age Group (>=66)	2.109	1.919	2.318
Education Level (Junior high)	4.661	3.855	5.635
Education Level (Senior high)	5.976	4.915	7.266
Education Level (College)	3.127	2.583	3.787
Region	1.223	1.158	1.292
Recent Migrant Status	3.121	2.570	3.789

Source: Data processing results of SUSENAS March 2020

From the output of odds ratio, it can be seen that the risk of the male population in a poor household in the prevalence of smoke is 132 times greater than the female population. However, the value of this odds ratio can be said to be very large. The cause of the large value of the odds ratio is from the sample taken. There is 750 female population in the sample, but only 4 women smoked. This condition is in line with research from Syamlal et al which states that women had a lower prevalence of smoking than

men, especially among working adults [19]. Bazotti et al also stated that in the Brazilian smoking population, men are predominant in all the age groups [4].

Then, the risk of the population 46 years old and over in poor households in the prevalence to smoke is greater than population 5-11 years old (child). 4.77 times to age group 46-65 years old and 2.11 times to age group 66 years old and over. This is following the condition in Indonesia in general. People will smoke more often as they get older. And start smoking habits in the age group of children or adolescents. However, the risk for smoking in the age group 26-45 years old is smaller than the age group 5-11 years old, which is 0.38 times.

It is quite interesting in the education level variable. The risk of a population in a poor household with a minimum education level of junior high school in the prevalence of smoke is greater than those with low-level education. 4.66 times to junior high school level, 5.98 times to senior high school level, and 3.13 times to college level. Whereas some researchers state that higher education is always associated with the prevalence of not smoking [20-23]. However, other researchers found a novelty that is in line with the results of this study, where the prevalence of smoking in the highly educated population is higher than the low educated population [24-26]. The increasing prevalence at the high education level can be caused, one of which is a lack of education and health literacy [26]. Although the education obtained is high, but still lacking in health education, it can increase the risk for smoking.

Next, the risk of the population in poor households living in urban areas in the prevalence of smoke is 1.22 times greater than living in rural areas. A study from Idris et al also concluded that in most countries, smoking prevalence was highest in urban areas, and increased with urbanization [27].

The risk of the population in poor households and their status as a recent migrant in the prevalence to smoke is 3.12 times greater than non-recent migrants. This result is in line with a study from Liu et al who state that the migratory history was positively associated with current smoking behavior [28].

4. Conclusion

Smoking habits in Indonesia have been formed in the past. Starting from the colonial era until now, smoking seems to have become a culture in society. Even though there have been so many studies related to the adverse effects of smoking, both from the social, economic, and health aspects.

Riau Islands Province as one of the youngest provinces in Indonesia has a smoking prevalence in the population aged 15 years and over which is 26.16%. Although this condition is considered to be better than other provinces on Sumatera Island, it is still necessary to conduct research related to the characteristics of smokers in the Riau Islands province. This study focuses on the smoking habit of the population in poor households for cigarette consumption is still quite high. On the other hand, the budget used for smoking should be diverted to better nutrition so that they can get prosperity.

This research method uses descriptive and inferential analysis. The result of the descriptive analysis shows that the percentage of smokers is greater in males, aged 26-45 years old, with a senior high school level, living in urban areas and as a recent migrant. While the results of inferential analysis using binary logistic regression analysis showed that of all the independent variables used, there was only one variable that was not significant, namely the age group variable (12-25 years old). From the results of binary logistic regression analysis, the prevalence for smoking in the population aged 5 years old and over in poor households will be greater in the male population (OR = 132.04), the age group of 46-65 years old (OR = 4.77), the age group of 66 years old and over (OR = 2.11), the junior high school level (OR = 4.66), the senior high school level (OR = 5.98), the college level (OR = 3.13), living in the urban area (OR = 1.22) and the recent migrant (OR = 3.12).

The results of this study can be used as recommendations for policymakers to provide special treatment to poor households in reducing their cigarette consumption and divert them to meet their nutritional needs. Then, more specific policies can be made for each character that has a greater prevalence of smoking. For example, by increasing health education for the population with higher education. Then, for the young age group to be massively socialized about the negative effects of smoking, so that it can reduce the opportunity to smoke in old age.

Suggestion for further study is the other independent variables can be added, especially economic variables/characteristics, so the information on the characteristics of smokers obtained can be more

complete. In addition, comparisons can be made with other provinces on the Sumatera Island or compared with different years.

References

[1]   Tirtosastro S and Murdiyati A S 2010 *Buletin Tanaman Tembakau, Serat dan Minyak Industri* Kandungan kimia tembakau dan rokok **2** 33

[2]   Sunaryo T 2013 *Kretek Pusaka Nusantara* (Jakarta: Serikat Kerakyatan Indonesia) p 31

[3]   TCSC I 2020 *Atlas Tembakau Indonesia 2020* (Jakarta Pusat: TCSC IAKMI) p 10

[4]   Bazotti A, Finokiet M, Conti I L, Franca M T A and Waquil P D 2016 *Ciencia & Saude Coletiva* Smoking and poverty in Brazil: an analysis of the profile of the smoking population **21** 45

[5]   Haustein K 2006 *Euro. J. of Cardio. Preven. and Rehab.* Smoking and poverty **13** 312

[6]   Hosseinpoor A R, Parker L A, Espaignet E T and Chatterji S 2012 *Plos One* Socioeconomic inequality in smoking in low-income and middle-income countries: results from the world health survey **7** 1

[7]   Sari P K and Seftarita C 2018 J. *Ilmiah Mahasiswa Ekonomi Pembangunan FEB Unsyiah* Analisis konsumsi rokok pada rumah tangga miskin dan tidak miskin di Kabupaten Aceh Besar **3** 315

[8]   Flint A J and Novotny T E 1997 *Tobacco Control* Poverty status and cigarette smoking prevalence and cessation in United States, 1983-1993: the independent risk of being poor **6** 14

[9]   Hou B, Nazroo J, Banks J and Marshall A 2018 *Frontiers in Public Health* Migration status and smoking behaviors in later-life in China – evidence from the China health and retirement longitudinal study **6** 1

[10] Tong E, Saito N, Tancredi D J, Borges G, Kravitz R L, Hinton L, Aguilar-Gaxiola S, Medina-Mora M E and Breslau J 2012 Ame. *J. of Public Health* A transnational study of migration and smoking behavior in the Mexican-origin population **102** 2116

[11] Bosdriesz J R, Lichthart N, Witvliet M I, Busschers W B, Stronks K and Kunst A E 2013 *Plos One* Smoking prevalence among migrants in the US compared to the US-born and the population in countries of origin **8** 1

[12] Hosmer D W, Lemeshow S and Sturdivant R X 2013 *Applied Logistic Regression Third Edition* (New Jersey: John Wiley & Inc) p 1

[13] Badan Pusat Statistik 2020 *Welfare Indicators 2020* (Jakarta: BPS RI) p 150

[14] Chen L C and Wang J Y 2020 *Wiley Biometrics* Discussion of "assessing the goodness-of-fit of logistic regression models in large samples: a modification of the Hosmer-Lemeshowtest," by Giovanni Nattino, Michael L. Pennell, and Stanley Lemeshow **76** 569

[15] Fagerland M W and Hosmer D W 2012 *The Stata J.* A generalized Hosmer-Lemeshow goodness-of-fit test for multinomial logistic regression models **12** 452

[16] Guffey D 2012 *Hosmer-Lemeshow goodness-of-fit test: Translations to the Cox Proportional Hazard Model* (Thesis) (Washington: University of Washington) p 2

[17] Boateng E Y and Abaye D A 2019 *J. of Data Analys. and Inform. Process.* A review of the logistic regression model with emphasis on medical research **7** 202

[18] Szumilas M 2010 *J. Can. Acad. Child Adolesc. Psychiatry* Explaining odds ratios **19** 227

[19] Syamlal G, Mazurek J M and Dube S R 2014 *Am. J. Prev. Med.* Gender differences in smoking among U.S. working adults **47** 474

[20] Assari S and Mistry R 2018 *Int. J. Environ. Res. Public Health* Educational attainment and smoking status in a national sample of American adults; evidence for the blacks' diminished return **15** 767

[21]   Urban R, Kugler G, Olah A and Szilagyi Z 2006 *Nicotine & Tobacco Research* Smoking and education: do psychosocial variables explain the relationship between education and smoking behavior? **8** 571

[22]   Tomioka K, Kurumatani N and Saeki K 2020 *J. Epidemiol.* The association between education and smoking prevalence, independent of occupation: a nationally representative survey in Japan **30** 140

[23]   Gilman S E, Martin L T, Abrams D B, Kawachi I, Kubzansky L, Loucks E B, Rende R, Rudd R and Buka S L 2008 *Int. J. Epidemiol.* Educational attainment and cigarette smoking: a causal association? **37** 621

[24]   Zhu B P, Giovino G A, Mowery P D and Eriksen M P 1996 Ame. *J. of Public Health* The relationship between cigarette smoking and education revisited: implications for categorizing persons' educational status **86** 1587

[25]   Xu X, Liu L, Sharma M and Zhao Y 2015 *Int. J. Environ. Res. Public Health* Smoking-related knowledge, attitudes, behaviors, smoking cessation idea and education level among young adult male smokers in Chongqing, China **12** 2146

[26]   Xu X, Rao Y, Wang L, Liu S, Guo J J, Sharma M and Zhao Y 2017 *Tobacco Induced Diseases* Smoking in pregnancy: a cross-sectional study in China **15** 40

[27]   Idris B 1 et al 2007 *Health & Place* Higher smoking prevalence in urban compared to non-urban areas: time trends in six European countries **13** 708

[28]   Liu Y et al 2015 BMC *Public Health* Determinants of tobacco smoking among rural-to-urban migrant workers: a cross-sectional survey in Shanghai **15** 138