AN ORDINAL LOGISTIC REGRESSION MODEL FOR ANALYZING RISK ZONE STATUS OF COVID-19 SPREAD

Tessya Mutiara Dewi [1*]**, Rosita Kusumawati**[2]

[1]*Department of Mathematics, FMIPA, Universitas Negeri Yogyakarta*
Colombo St., No. 1, Sleman, Yogyakarta, 55281, Indonesia

[2]*Department of Statistics, FMIPA, Universitas Negeri Yogyakarta*
Colombo St., No. 1, Sleman, Yogyakarta, 55281, Indonesia

Corresponding author's e-mail: ¹* tessyamutiaradewi99@gmail.com

Abstract. *Coronavirus disease 2019 (COVID-19) is a new type of virus that has been found to have infected human since it first appeared in Wuhan, China, in December 2019. This study aims to determine the factors that influence the risk zone status of COVID-19 spread in Indonesia using ordinal logistic regression. The ordinal logistic regression model in this study uses proportional odds model because the researcher assumes probability of predictor variable coefficients is the same for each respond category. The response variable is secondary data from the COVID-19 Handling Task Force, namely the status of the risk zone for the spread of COVID-19 who has 4 categorical levels, namely high, medium, low, and no cases. Predictor variables are elderly population, COVID-19 referral hospital, diabetes mellitus, hypertension, hand washing behavior, male population, and smoking habits. Based on results of the analysis, variables that significantly affect the risk zone status of COVID-19 spread in Indonesia are elderly population and diabetes mellitus. The Odds proportional figure shows that the higher percentage of the elderly population, the higher chance of an area with high-risk zone status (OR=1.171). The higher percentage of comorbidities diabetes mellitus, the higher chance of an area with high-risk zone status (OR=1.569).*

Keywords: *Ordinal Logistic Regression, Risk Zone COVID-19.*

https://ojs3.unpatti.ac.id/index.php/barekeng/   barekeng.math@yahoo.com

1. INTRODUCTION

Coronavirus disease 2019 (COVID-19) is a new type of virus that has been found to have infected humans since it first appeared in Wuhan, China, in December 2019 [1]. The World Health Organization (WHO) officially declared COVID-19 as a pandemic on March 9, 2020 [2]. COVID-19 is spreading rapidly to various countries including Indonesia. The spread of Coronavirus Disease in the world tends to continue to increase from time to time and causes greater casualties and material losses. Based on data from the COVID-19 Handling Task Force shows that the number of COVID-19 cases as of December 30, 2020 nationally was 735,124 people with positive cases, 603,741 people with recovered cases and 21,944 people with death cases [3].

In an effort to break the chain of spread of COVID-19, the COVID-19 Handling Task Force established the status of risk zones for the spread of COVID-19 for each region. The risk zone status is divided into four regional zoning levels based on color, namely red for high risk, orange for moderate risk, yellow for low risk and green for areas with no cases [4]. The determination of the risk zone uses public health indicators which are divided into 10 epidemiological indicators, two public health surveillance indicators and two health service indicators.

Number of research that study factors that influence the spread of COVID-19 have been done, research from Atmojo, et al [5] said that COVID-19 patients who had a smoking habit before the pandemic had the potential to experience 2 times worsening of the situation. This is due to a decrease in lung function and exacerbated by COVID-19 infection. According to Muflihah et al [6], infection prevention and control is an effort to prevent infection to minimize the transmission of infectious diseases such as COVID-19. One of these preventive behaviors is to wash hands. According to Satria, Tutupoho and Chalidyanto [7], elderly COVID-19 patients have a 2,097 times greater risk of dying from COVID-19, male COVID-19 patients have a 1.087 times greater risk of dying than women and COVID-19 with comorbid diabetes has a 4.384 times greater risk of dying from COVID-19 than patients without comorbid diabetes mellitus. According to Gunawan et al [8], hypertension is a comorbid that is often found in COVID-19 patients and can worsen the condition of COVID-19 patients up to 2.5 times. The need for health services and facilities will increase along with the increasing number of cases. Health facilities that need to be considered are the availability of hospitals, bedrooms and PPE [9]. Population mobility is also a factor in the spread of COVID-19 disease caused by the stages of the spread of this disease that have taken place through local transmission [10]. Therefore, it is necessary to conduct an analysis to determine the factors that influence the status of the risk zone for the spread of COVID-19 in Indonesia.

Ordinal logistic regression analysis is a statistical method that describes the relationship between an ordered response variable (Y) and one or more predictor variables (Y) [11]. Study about factors that significantly influence the status of the risk zones for the spread of COVID-19 in districts/cities in Indonesia can be done using ordinal logistic regression. The response variable is a status of the COVID-19 risk zone with four categories, high risk, medium risk, low risk, and no affected/no cases. The predictor variables include a referral hospital of COVID-19, age > 65 years old, the habit of washing hands properly, diabetes comorbidity, hypertension comorbidity, male population and smoking habit. We assumes that the probability of the predictor variable coefficients is the same for each response category, therefore the proportional odds model is used in this study.

2. RESEARCH METHODS

Ordinal logistic regression will be used to determine the factors that influence the status of the risk zone for the spread of COVID-19 in Indonesia. In this study, data of response and predictor variables of year 2020 obtained from the Covid-19 Handling Task Force (Satgas COVID-19) [12] [13], the Central Statistics Agency (BPS) [14], and the Indonesian Health Research and Development Agency (Litbangkes RI) [15]. Based on districts/cities throughout Indonesia is used. However, data on diabetes mellitus, hypertension, hand washing habits, and smoking habits for year 2020 are not available. Therefore, predictions for year 2020 are made using data from year 2013 to 2018 data with constant growth assumption. The data used are 166 districts/cities. The response variable used is the status of the COVID-19 risk zone which consists of 4 levels, namely high risk, medium risk, low risk, and no case/no case. The predictor variables used were COVID-19

referral hospital, age > 65 years, correct hand washing habits, comorbidities with diabetes, comorbidities with hypertension, male population, and smoking habits.

Steps to analysis of the factors that influence the status of the risk zone for the spread of COVID-19 using ordinal logistic regression are: 1) Build a predictive ordinal logistic regression model. In the study using the proportional odds model, 2) Examine the significance of the overall regression parameters using the Likelihood Ratio Test (LRT), 3) Examine the significance of the parameters partially using the Wald test, 4) Define the best ordinal logistic model that have a significant effect with the Wald test, 5) Calculate odds ratios and interprets them, 6) Calculate the estimated probability for each category level, 7) Visualize the category level of each district/city based on the estimated probability in the form of a map, 8) Test the assumption of proportional odds to find out the proportional odds model is feasible to use, 9) Conduct a fitness test model (Goodness of Fit) to find out whether the model used is in accordance with the data.

2.1. Ordinal Logistic Regression

Ordinal logistic regression is a statistical method that describes the relationship between a response variable (Y) and more than one predictor variable (X) where the response variable is more than two categories and the measurement scale is level [11]. One of the model approaches used to analyze ordinal logistic regression is the cumulative logit model. Ordinal logistic regression uses the logit model for cumulative probability, namely the cumulative logit model. The cumulative logit model is a model obtained by comparing the cumulative probability, namely the probability of less than or equal to the j-th response category on the predictor variable X is $P(Y \leq j|X)$. Cumulative probability $P(Y \leq j|X)$, defined by

$$P(Y \leq j|X) = \pi(X) = \frac{e^{\beta_{j0}+\beta_1 x_1+\beta_2 x_2+\cdots+\beta_p x_p}}{1+e^{\beta_{j0}+\beta_1 x_1+\beta_2 x_2+\cdots+\beta_p x_p}} \tag{1}$$

where j is the categorical level the response variable, $j = 1,2,..,J-1$.
Cumulative logit models are defined by:

$$Logit\ P(Y \leq j|X) = \ln\left(\frac{P(Y \leq j|X)}{1-P(Y \leq j|X)}\right) \tag{2}$$

$$Logit\ P(Y \leq j|X) = \ln\left(e^{\beta_{j0}+\beta_1 x_1+\beta_2 x_2+\cdots+\beta_p x_p}\right) \tag{3}$$

$$Logit\ P(Y \leq j|X) = \beta_{j0} + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p \tag{4}$$

where j is the categorical level on the response variable, $j = 1,2,..,J-1$.
The probability of observing the response category j can be written as [16]:

$$\pi_j(x) = P(Y = j|x) = P(Y \leq j|x) - P(Y \leq j-1|x) \tag{5}$$

with

$$P(Y \leq 0) = 0,\ P(Y \leq J|x) = 1 \text{ and } P(Y \leq j|x) = \frac{e^{\beta_{j0}+\beta_1 x_1+\beta_2 x_2+\cdots+\beta_p x_p}}{1+e^{\beta_{j0}+\beta_1 x_1+\beta_2 x_2+\cdots+\beta_p x_p}}$$

and the probability for category J is $\pi_J(x) = P(Y \leq J|x) - P(Y \leq J-1|x)$

$$= 1 - \frac{e^{\beta_{J-1,0}+\beta_1 x_1+\beta_2 x_2+\cdots+\beta_p x_p}}{1+e^{\beta_{J-1,0}+\beta_1 x_1+\beta_2 x_2+\cdots+\beta_p x_p}} \tag{6}$$

2.2 Parameter Significant Test

Likelihood Ratio Test (LRT) is used to test the fitness of the model with the predictor variables as a whole [11]. The test statistic used is the Likelihood Ratio Test (LR). The hypothesis used is $H_0: \beta_1 = \beta_2 = \cdots = \beta_p = 0$ and $H_1: \exists\ \beta_k \neq 0$, with k = 1,2,…p. The test statistic used is the Likelihood Ratio Test (LRT).

$$G = -2 \ln \left[\frac{maximum\ likelihood\ fuction\ on\ H_0}{aximum\ likelihood\ fuction\ on\ H_0\ and\ H_1} \right] \tag{7}$$

The rejection criteria for H_0 are $G > \chi^2(\alpha, df)$ or $p - value < \alpha$ with df = k, k is the number of predictor variables.

Wald test was conducted to examine the effect of each coefficient β_i partially. The test results partially show whether a predictor variable is eligible to be included in the model or not [17]. The hypothesis used is $H_0: \beta_k = 0$ and $H_1: \beta_k \neq 0$; with k = 1,2,…,p.
The test statistic used is the Wald test.

$$W = \frac{\widehat{\beta_k}}{SE(\hat{\beta}_k)} \text{ and } SE(\hat{\beta}_k) = \sqrt{\widehat{Var}(\hat{\beta}_k)} \tag{8}$$

The rejection criteria for H_0 are $|W| > Z_{\alpha/2}$ or $W^2 > \chi^2(\alpha, db)$ with df = k, k is the number of predictor variables or $p - value < \alpha$.

2.3 Assumption of Proportional Odds

An important assumption of proportional odds model is parallel lines, which means that the odds ratio is not affected by where the dependent variable is dichotomized [18]. The proportional odds model is a special model that assumes that the logit of this cumulative probability changes linearly as the predictor variable changes [19]. This assumption states that the categories of predictor variables are parallel to each other.

The hypothesis used is H_0: the model produces the same regression coefficient (slope) and H_1: the model does not produces the same regression coefficient (slope).
The test statistics used are :

$$PL = -2ln \left[\frac{l_0}{l_1} \right] \sim X^2{}_{\alpha;k\ (j-2)} \tag{9}$$

with $PL = parallel\ lines$
        l_0 : likelihood function with predictor variables that assume parallel lines.
        l_1 : likelihood function with predictor variables that do not assume parallel lines.

The rejection criteria for H_0 is $p - value < \alpha$ or $PL > X^2{}_{\alpha;k\ (j-2)}$.

2.4 Lipsitz Test for Goodness of Fit

Fitness test can be done by comparing the observed value for a subject with the predicted value for that subject [18]. Based on the Hosmer-Lemeshow [20] approach, the Lipsitz test is suitable for the ordinal logistic regression model.

The hypothesis used is H_0: the model is in accordance with the data (there is no difference between the results of the observations and the results of the predictions of the model) and H_1: the model does not match the data (there is a difference between the results of the observations and the results of the predictions of the model). The test statistics used are:

$$LR = -2(L_1 - L_0) \sim X^2{}_{g-1} \tag{10}$$

with

        l_1 : log likelihood of the model without indicator variables
        l_1 : log likelihood of the model with indicator variables
        g : number of groups

The rejection criteria for H_0 is $LR > X^2{}_{g-1}$ or p-value $< \alpha$.

2.5 Odds Ratio

The odds ratio is a measure that estimates how much the predictor variable tends to respond to the response variable. The odds ratio is denoted by OR, defined by the comparison of the odds value for $x = 1$ with the odds value for $x = 0$, and the following equation can be formed [17]:

$$OR = \frac{\pi(1)/[1-\pi(1)]}{\pi(0)/[1-\pi(0)]} = \frac{\left(\frac{e^{\beta_0+\beta_1}}{1+e^{\beta_0+\beta_1}}\right)/\left(\frac{1}{1+e^{\beta_0+\beta_1}}\right)}{\left(\frac{e^{\beta_0}}{1+e^{\beta_0}}\right)/\left(\frac{1}{1+e^{\beta_0}}\right)} = e^{(\beta_0+\beta_1)-\beta_0} = e^{\beta_1} \tag{11}$$

Interpretation of odds ratio (OR):
1. OR> 1 indicates that there is an effect because the value of the probability of success is greater than the value of the event of not being successful.
2. OR=1 indicates that there is no effect because the probability of success is the same as failure.
3. OR < 1 indicates that the probability of success is smaller than the value of the probability of not being successful.

3. RESULTS AND DISCUSSION

Figure 1 shows COVID-19 status of 166 cities, the majority are medium risk zone status, which is 93 cities. 44 cities have high risk zone status, 27 cities have low risk zone status and 2 cities have no affected/no cases risk zone status. It is clearly seen in table 3 that the referral hospital of COVID-19 has a minimum and maximum percentage of 0 and 0.01, the mean and median percentage of referral hospital of COVID-19 are 0.001 and 0.001 with a standard deviation of 0.001.
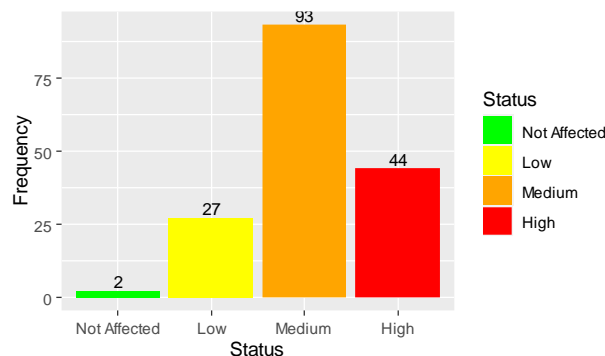

Figure 1. Frequency Of The Risk Zone Status Of Covid-19 Spread In Indonesia 2020

Table 3. Descriptive Statistics for Predictor Variables

Variables	Min	Max	Median	Mean	Sd
Referral hospital of COVID-19	0	0.01	0.001	0.001	0.001
Age > 65 years old	1.34	20.38	8.305	8.78	2.84
Correct hand washing behavior	0.1	95.98	49.33	46.26	20.72
Diabetes	0.01	5.35	2.01	2.12	1.17
Hypertension	2.23	20.11	9.34	9.65	3.43
Male population	48.72	55.14	50.9	50.93	1.01
Smoking habits	13.03	32.08	23.40	23.21	3.68

Based on the analysis using the R program, the estimation results and parameter tests is given in Table 4. Table 4 consists of coefficients, standard errors (S.E), Wald test values, and p-value. Overall parameter significance testing is carried out using the likelihood ratio test statistic (LR). The hypothesis used is $H_0: \beta_1 = \beta_2 = \cdots = \beta_7 = 0$ and $H_1: \exists\, \beta_k \neq 0$. H_0 is rejected if $-2\log(\Lambda) > X^2(\alpha, db)$ or $p - value < \alpha$. In this study, the LR statistic value was obtained 170.18 and p < 0.000 which indicated that there is strong evidence of the importance of referral hospital of COVID-19, age > 65 years old, correct hand washing behavior, diabetes, hypertension, male population, and smoking habits are in the model. To test each predictor variable that has a significant effect on the model, Wald test is carried out. The Wald test is used to test the set of

hypotheses ($H_0: \beta_k = 0$ vs $H_1 : \beta_k \neq 0$) for individual regression slope coefficients. The null hypothesis is rejected if $|W| > Z_{\alpha/2}$ or p-value is less than the significance level ($\alpha = 0.05$). Based Table 2, variable age > 65 years has a Wald value $|W| = 2.275$ then p-value is $0.003 < 0.05$ and the diabetes variable has a Wald $|W| = 3.135$ then p-value is $0.001 < 0.05$. Therefore H_0 is rejected and there is strong evidence that variable age > 65 years and diabetes is important to include in the model given the other predictor variables in the model. The Wald test showed that age > 65 years, and diabetes were statistically significant at the significance level 0.05.

Table 4. Estimation and test results for parameters

Variable	$\widehat{\beta}$	S.E.	Wald |W|	p-value
Referral hospital of COVID-19	51.696	0.075	0.273	1.000
Age > 65 years old	-0.153	0.067	2.275	0.013*
Correct hand washing behavior	-0.011	0.008	1.42	0.064
Diabetes comorbidity	-0.472	0.150	3.135	0.001**
Hypertension comorbidity	0.051	0.050	1.017	0.164
Male population	-0.051	0.188	0.274	1.000
Smoking habits	0.041	0.047	0.667	0.413

The estimation results and parameter tests for the best model is given in Table 5. Based on Table 5, the best ordinal logistic regression model is $logit\left(\hat{P}(Y \leq j)\right) = \hat{\beta}_{j0} + -0.159X_2 - 0.451X_4$ with $\hat{\beta}_{10} = -2.283$, $\hat{\beta}_{20} = 0.655$ and $\hat{\beta}_{30} = 3.491$. Description of predictor variables, namely elderly population (X_2) and diabetes mellitus (X_4). To produce an easier interpretation, the odds ratio calculation is carried out. Based on the results of the processing, the odds ratio for the elderly population variable was 1.171. Because OR> 1, it can be interpreted that the increase in status is 1.171 times higher for every increase in the elderly population. This may be because the most vulnerable groups are affected by the COVID-19 disease is elderly. The vulnerability of the elderly during the COVID-19 pandemic is caused by a decrease in body resistance and comorbid diseases in the elderly that can increase the risk of death. Information on the impact of COVID-19 and restrictions on physical social interaction also affect mental health for the elderly, which can lower the immune system. The diabetes variable has an odds ratio of 1.569. Because OR > 1, it can be interpreted that the increase in status is 1.569 times higher for each increase in diabetics. This may be because the decreased immunity of diabetics is one of the risk factors for being infected with COVID-19.

Tabel 5. Estimation Results and Parameter Tests For The Best Model

Variable	$\widehat{\beta}$	S.E.	p-value	Odds Ratio
Intercept 1	-2.283	0.869		
Intercept 2	0.655	0.558		
Intercept 3	3.491	0.624		
Age > 65 years old	-0.159	0.056	0.004**	1.171
Diabetes comorbidity	-0.451	0.138	0.000***	1.569

From the logistic regression equation, the best guess can be seen for the status for each district/city using probability estimates. The estimated probability of success can be calculated by substituting the predictor values in the logistic regression equation or the best guess. The results of the calculation of the estimated value of 166 districts/cities have been presented on a map of the status of the risk zone for the spread of COVID-19 in Indonesia based on the calculation of the estimated probability value in Figure 2.
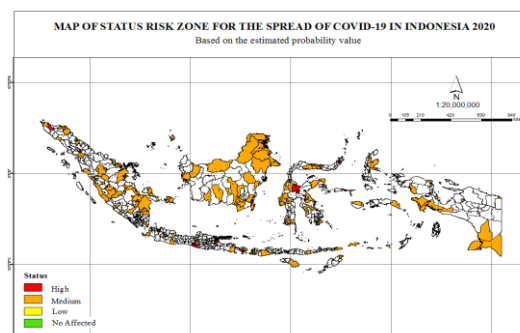

Figure 2. Map of the risk zone status of COVID-19 spread in Indonesia 2020

Based on Figure 2, 18 districts/cities have high status, namely: Pidie, South Tangerang, North Jakarta, Central Jakarta, Bantul, Sleman, Yogyakarta City, Gunung Kidul, Kulon Progo, Blitar City, Sidoarjo, Tabanan, Tomohon City, Manado City, Minahasa, Sangihe Islands, Poso, Gorontalo City. There are 148 districts/cities with moderate status, namely North Aceh, East Aceh, Aceh Besar, Southwest Aceh, Medan City, Deli Serdan, North Tapanuli, Pariaman City, Padang, Padang Pariaman, Pesisir Selatan, West Pasaman, Pekanbaru City, Kampar, Indragiri Hilir, Meranti Islands, Kuantan Singingi, Jambi City, Muaro Jambi, Merangin, Bungo, Sarolangun, Palembang City, Banyuasin, Ogan Komering Ilir, Ogan Ilir, Ogan Komering Ulu, Bengkulu City, South Bengkulu, Muko-muko, Bengkulu Central, Kepahiang, Bandar Lampung City, South Lampung, Tulang Bawang, Mesuji, Pangkal Pinang City, Bangka, Central Bangka, Belitung, East Belitung, Batam, Karimun, Tanjung Pinang, Anambas Islands, Natuna, Tangerang City, Tangerang, Pandeglang, City Serang, South Jakarta, East Jakarta, West Jakarta, West Bandung, Depok City, Bandung City, Majalengka, Tasikmalaya, Surakarta City, Tegal City, Boolali, Semarang City, Magelang City, Malang City, Mojokerto, Banyuwangi, Gianyar, Badung, Karangasem , Kot a Denpasar, Bima City, Sumbawa, Mataram City, West Lombok, Central Lombok, Kupang City, East Sumba, Belu, Alor, Ende, Pontianak City, Singkawang City, Kubu Raya, Kapuas Hulu, Sanggau, West Kotawaringin, Kapuas, Palangkaraya City , East Kotawaringin, South Barito, Tanah Laut, Kotabaru, Banjar, Barito Kuala, Tapin, Balikpapan, North Penajem Paser, East Kutai, Bulungan, Malinau, Nunukan, Tana Tidung, Tarakan City, Bitung City, Morowali, Palu City, Banggai, Sigi, Bulukumba, Pare-Pare City, Tana Toraja, Maros, Gowa, Bau-Bau City, Wakatobi, Buton, Konawe, Kolaka, North Gorontalo, Gorontalo, Pahuwato, Boalemo, Majene, Olewali Mandar, Masama, North Mamuju, Mumuju, Southwest Maluku, Baru, Ambon City, Tual City, Eastern Seram, Ternae City, North Halmahera, Central Halmahera, East Halmahera, Sula Islands, Manokwari, Sorong City, Fakfak, Kaimana, South Sorong, Jayapura City, Mappi, Dogiyai, Merauke. Low status and unaffected/no cases, no districts/cities have both statuses.

To find out that the proportional odds model has met the parallel lines assumption, the odds proportional assumption test is was used. This test is conducted to determine whether the proportional odds model is feasible to use. The hypothesis used in the proportional odds assumption test is H_0: the model produces the same regression coefficient (slope) vs H_1 : the model does not produces the same regression coefficient (slope). The result of this test is the PL value of 3.02 and p-value of 0.55. Because p-value $> \alpha$ and PL< Chi-square table (9.487) so H_0 accepted. Thus, at the 5% confidence level, it can be said that the resulting model has the same regression coefficient (slope).

Table 6. Test The Assumption Of Proportional Odds

Variable	PL	Df	Chi-square table	p-value
Omnibus	3.02	4	9.487	0.55

The Lipsitz test is used to determine the resulting model is in accordance with the data. The hypothesis used in the Lipsitz test is H_0: model according to data and H_1: model does not according to data. The result of this test is the LR value of 3.841 and p-value of 0.846. Because p-value $> \alpha$ and LR< Chi-square table (16.918) so H_0 accepted. Thus, at the 5% confidence level, it can be said that the resulting model according to data.

Table 7. Lipsitz Test (Goodness Of Fit)

LR	Chi-square table	p-value
3.841	16.918	0.846

From the two tests above, namely the assumption of proportional odds and goodness of fit, it can be concluded that the resulting model fits the data.

This study describes the factors that influence the status of the risk zone for the spread of COVID-19 in Indonesia. The model obtained shows that 2 factors, namely elderly and diabetes co-morbidities, significantly affect the status of the risk zone for the spread of COVID-19. Based on the results of the processing, the odds ratio for the elderly population variable was 1.171. It can be interpreted that the increase in status is 1.171 times higher for district/ city with higher percentage the elderly population. This is in accordance with Indarwati's research [21] which states that the elderly population is the group most vulnerable to being affected by the COVID-19 disease. The vulnerability of the elderly during the COVID-19 pandemic is caused by a decrease in body resistance and comorbid diseases in the elderly that can increase the risk of death. Information on the impact of COVID-19 and restrictions on physical social interaction also

affect the mental health of the elderly, which can lower the immune system [21]. The diabetes variable has an odds ratio of 1.569. It means that the increase in status is 1.569 times higher for higher diabetics value. This is in accordance with research conducted by Roeroe et al [22] which states that diabetes is one of the main risk factors for COVID-19. The severity and mortality of COVID-19 is higher in patients with diabetes than in non-diabetics. This is due to the decreased immunity of diabetics, which is one of the risk factors for being infected with COVID-19 [22].

4. CONCLUSIONS

The ordinal logistic regression model shows that the elderly population and comorbid diabetes have a significant influence on the status of the risk zone for the spread of COVID-19. The increasing percentage of the elderly population and the population suffering from diabetes mellitus in a district/city will increase the chance of the status of the risk zone for the spread of COVID-19 to a higher level in that district/city.

REFERENCES

[1]   Kemenkes RI, "FAQ," 3 2020. [Online]. Available: https://www.kemkes.go.id/folder/view/full-content/structure-faq.html. [Accessed 15 Desember 2020].

[2]   Satgas COVID-19, "Apa yang dimaksud dengan pandemi?," 2020. [Online]. Available: https://covid19.go.id/tanya-jawab?.

[3]   Kemenkes RI, "Perkembangan COVID-19 di Indonesia update 5 Juli," Kemenkes RI, 2020.

[4]   BNPB, "Pengelompokan Kriteria Risiko COVID-19 di Daerah Berdasarkan Zonasi Warna," 20 Juni 2020. [Online]. Available: https://bnpb.go.id/berita/pengelompokan-kriteria-risiko-covid19-di-daerah-berdasarkan-zonasi-warna. [Accessed 16 Desember 2020].

[5]   J. T. Atmojo, D. Arradini, A. T. Darmayanti, A. Widiyanto and R. T. Handayani, "Dampak Merokok Terhadap COVID-19," *Jurnal Ilmiah Permas,* vol. 11, no. 1, pp. 169-176, 2021.

[6]   I. S. Muflihah, A. Noviyana, Purwati and W. Margiana, "Edukasi Cuci Tangan sebagai Salah Satu Pencegahan dan Penanganan Infeksi dalam Upaya Mawas Diri dari Covid," *Jurnal Publikasi Kebidanan,* vol. 11, no. 2, pp. 109-117, 2020.

[7]   M. A. Satria, R. V. Tutupoho and D. Chalidyanto, "Analisis Faktor Risiko Kematian dengan Penyakit Komorbid COVID-19," *Jurnal Keperawatan Silampari,* vol. 4, no. 1, pp. 2581-1975, 2020.

[8]   A. Gunawan, K. Prahasanti, M. R. Utama and M. P. Airlangga, "Pengaruh Komorbid Hipertensi terhadap Severitas Pasien Coronavirus Disease 2019," *Jurnal Implementa Husada,* vol. 2, no. 1, 2020.

[9]   S. Setiati and M. K. Azwar, "COVID 19 AND INDONESIA," *Acta Medica Indonesia,* pp. 84-89, 2020.

[10] R. A. Ghiffari, "Dampak Populasi dan Mobilitas Perkotaan terhadap Penyebaran Pandemi COVID-19 di Jakarta," *Jurnal Tunas Geografi,* vol. 9, no. 1, pp. 81-88, 2020.

[11] D. W. Hosmer and S. Lemeshow, Applied Logistic Regression, New York: John Wiley & Sons, Inc, 2000.

[12] Satgas COVID-19;, "Peta Zona Risiko COVID-19," 2020. [Online]. Available: https://covid19.go.id/peta-risiko. [Accessed 5 January 2021].

[13] Satgas COVD-19, "Rumah Sakit Rujukan COVID-19 Update Mei," [Online]. Available: https://bnpb-inacovid19.hub.arcgis.com/datasets/inacovid19::rs-rujukan-update-may-2020/about. [Accessed 30 December 2020].

[14] Badan Pusat Statistik, "Jumlah Penduduk Menurut Wilayah, Kelompok Umur, dan Jenis Kelamin," [Online]. Available: https://sensus.bps.go.id/topik/tabular/sp2020/86. [Accessed 5 January 2021].

[15] Litbangkes RI, "Riskesdas," [Online]. Available: https://www.litbang.kemkes.go.id/laporan-riset-kesehatan-dasar-riskesdas/ . [Accessed 2021 Februari 2021].

[16] C. R. Bilder and T. M. Loughin, Analysis of Categorical Data With R, Boca Raton: Taylor & Francis Group, 2015.

[17] A. Agresti, Categorical Data Analysis, New York: John Wiley & Sons, Inc., 1990.

[18] D. G. Kleinbaum and M. Klein, Logistic Regression: A Self Learning Text (edisi ketiga), New York: Springer, 2010.

[19] N. I. Putri and Budyanra, Penerapan Regresi Logistik Ordinal dengan Proportional Odds Model pada Determinan Tingkat Stres Akademik Mahasiswa, Jakarta Timur: Seminar Nasional Official Statistics, 2019.

[20] M. W. Fagerland, "Tests for Goodness of Fit in Ordinal Logistic Regression Models," *Journal of Statistical Computation and Simulation,* vol. 86, no. 17, pp. 3398-3418, Maret 2016.

[21] R. Indarwati, "Lindungi Lansia dari COVID-19," *Indonesian Journal of Community Health Nursing,* vol. 5, 1 April 2020.

[22] P. A. Roeroe, B. P. Sedli and O. Umboh, "Faktor Risiko Terjadinya Coronavirus Disease 2019 (COVID-19) pada Penyandang Diabetes Mellitus," *e-Clinic,* vol. 9, no. 1, pp. 154-169, 2021.