

Final Project (1653 words)

Di Zhen (1717719)

June 17, 2020

Introduction

Suicide has drawn a primary public health concern in China. There are about 13 deaths per 100,000 people commit suicide each year. The features of suicide in China are important and unique. They are different from the features in the western world. In China, women are more likely to commit suicide than men [2]. The elderly have a higher suicide rate than the young [3]. People living in rural areas commit more suicide than those in urban [5]. And pesticide self-poisoning has been reported to be the main suicide method [1]. In addition, suicide actions not only result in loss of lives but also cause hospitals to suffer from a tremendous burden on resources, since the majority of people who committed suicide are survived and hospital-admitted.

The study of fatal suicide attempts in China would indicate factors leading to a fatal outcome after a suicide attempt, and also give suggestions about the prevention of suicide acts. In this report, we replicate part of the surveillance study [4] on the incidence and fatality of serious suicide attempts (SSA) in Shandong, China.

Methods

The data are from two public health surveillance system, which are the Disease Surveillance Points (DSP) system and Shandong Injury Surveillance System (ISS). The authors in the surveillance study [4] extracted data of suicide deaths and hospitalizations in three selected cities in Shandong, China and ascertained the cases by the 10th revision of the International Classification of Diseases and Related Health Problems.

After categorizing and cleaning the data, there are 2289 people and 8 major variables including whether dead or not, living in urban or rural areas, sex, age group, education level, occupation, method of suicide, and season of suicide. The age group includes 10-34, 35-49, 50-64, and no less than 65. The education level includes illiterate, primary, and no less than secondary level (secondary and tertiary). The occupation includes farming and all others (household, professional, unemployed, business, student, worker, retiree, others). The methods of suicide include hanging, pesticide, other poison (other poison, unspecified poison), and all others (cutting, drowning, jumping, others). The season is defined as spring (March-May), summer (June-August), autumn (September-November) and winter (December-February).

Fatality rate(%) is the proportion of died cases among all SSAs. χ^2 test was conducted to test the differences in fatality between groups and p-values were reported. Multiple logistic regression was also applied to test the associations between main factors and fatality, and the ORs, 95% CIs and p-values were reported. In order to extend the analysis of the authors', the assumption of multiple logistic regression mentioned above was tested. In addition, the data were grouped by residency, sex, age group, education, occupation, method, and season. Then, the number of death for each group was counted. Using this grouped data, a Poisson regression was conducted to analyze the associations between main factors and fatality, and the assumption of the Poisson regression was tested subsequently. All statistical analyses were conducted in R (Version 3.6.3).

Results and Discussion

As shown in Table 1, people who are living in a urban area, who are male, who are older, who are with low education, or with a farmer occupation have higher fatality rates (all chi-square test, p-value < 0.05). As for the suicide methods, the fatality of using other poison (10.0%) is the lowest, while the fatality of hanging (97.4%) is the highest. Attempts in seasons other than summer have higher fatality rates. The p-values of chi-square test are all less than 0.05, indicating that those independent variables are significantly associated with the fatality.

	Bivariate analysis			Multiple logistic regression	
	Number of SSAs)	Fatality(%)	p Value ¹	OR(95% CI)	p Value ²
Residency					
<i>Rural</i>	2064	1125 (54.5)	0.047	1.00	Reference
<i>Urban</i>	225	107 (47.6)		1.18 (0.82, 1.68)	NS
Sex					
<i>Female</i>	1184	579 (49.0)	< 0.001	1.00	Reference
<i>Male</i>	1105	653 (59.1)		1.30 (1.04, 1.62)	0.02
Age group					
<i>10-34</i>	482	117 (24.3)	< 0.001	1.00	Reference
<i>35-49</i>	488	187 (38.3)		1.33 (0.97, 1.83)	NS
<i>50-64</i>	569	337 (59.2)		1.59 (1.14, 2.22)	0.007
<i>65 and above</i>	750	591 (78.8)		2.13 (1.47, 3.07)	<0.001
Education					
<i>Secondary and above</i>	1136	317 (27.9)	< 0.001	1.00	Reference
<i>Primary</i>	633	463 (73.1)		4.77 (3.63, 6.31)	<0.001
<i>Illiterate</i>	520	452 (86.9)		10.93 (7.65, 15.83)	<0.001
Occupation					
<i>All Others</i>	355	150 (42.3)	< 0.001	1.00	Reference
<i>Farming</i>	1934	1082 (55.9)		1.35 (0.98, 1.86)	NS
Method					
<i>Pesticide</i>	1645	761 (46.3)	<0.001	1.00	Reference
<i>Other poison</i>	160	16(10.0)		0.09 (0.05, 0.16)	<0.001
<i>Hanging</i>	420	409 (97.4)		26.79 (14.99, 53.17)	<0.001
<i>All others</i>	64	46 (71.9)		3.70 (2.00, 7.04)	<0.001
Season					
<i>Summer</i>	669	316 (47.2)	<0.001	1.00	Reference
<i>Autumn</i>	558	287 (51.4)		1.11 (0.83, 1.49)	NS
<i>Winter</i>	473	278 (58.8)		1.65 (1.20, 2.26)	0.002
<i>Spring</i>	589	351 (59.6)		1.65 (1.23, 2.21)	<0.001

¹Test the differences in fatality between groups using Chi-square test.

²Test for regression coefficients in the multiple logistic model.

NS, not significant (p>0.05); SSA, serious suicide attempts.

Table 1: Fatality (% and 95% CIs) of SSAs and results of multiple logistic regression analysis (ORs and 95% CIs).

The result of multiple logistic regression is summarized in Table 1. People living in urban have a 18% increase in the odds of death by suicide compared to those living in rural areas, after controlling for sex, age group, education, occupation, method, and season. Males have a 30% increase in the odds of death by suicide compared to female after controlling for other variables. Those in the age group of 35-49, 50-64, and 65 and above have ORs of 1.33, 1.59, and 2.13 for death due to suicide versus being alive compared to those in the age group 10-34, after controlling for other variables. People with primary and illiterate educational level have a 377% and 993% increase in the odds of death due to suicide compared to those with secondary and above educational level, after controlling for other variables. People who are farmer have a 35% increase in the odds of death due to suicide compared to those with vocations other than farming, after controlling for other variables. Those hanging and using other methods to commit suicide have a 2579% and 270% increase in the odds of death due to suicide compared to those using pesticides, after controlling for other variables. While those using other poison have a 91.0% decrease in the odds of death due to suicide compared to those using pesticides, after controlling for other variables. People committing suicide in autumn, winter and spring have a 11%, 65%, and 65% increase in the odds of death due to suicide compared to those committing suicide in summer, after controlling for the

other variables. The p values of urban residency, 10-34 age group, Farming, and suicide in autumn are not significant, while others are significant (p-value > 0.05).

Poisson Regression		
	OR(95% CI)	p Value*
Residency		
<i>Rural</i>	1.00	Reference
<i>Urban</i>	0.21 (0.17, 0.26)	<0.001
Sex		
<i>Female</i>	1.00	Reference
<i>Male</i>	1.09 (0.97, 1.22)	NS
Age group		
<i>10-34</i>	1.00	Reference
<i>35-49</i>	1.19 (0.94, 1.50)	NS
<i>50-64</i>	1.68 (1.36, 2.08)	<0.001
<i>65 and above</i>	2.84 (2.33, 3.50)	<0.001
Education		
<i>Secondary and above</i>	1.00	Reference
<i>Primary</i>	1.38 (1.20, 1.60)	<0.001
<i>Illiterate</i>	1.48 (1.28, 1.72)	<0.001
Occupation		
<i>All Others</i>	1.00	Reference
<i>Farming</i>	3.73 (3.15, 4.45)	<0.001
Method		
<i>Pesticide</i>	1.00	Reference
<i>Other poison</i>	0.04 (0.02, 0.07)	<0.001
<i>Hanging</i>	0.75 (0.66, 0.85)	<0.001
<i>All others</i>	0.20 (0.15, 0.27)	<0.001
Season		
<i>Summer</i>	1.00	Reference
<i>Autumn</i>	0.96 (0.82, 1.13)	NS
<i>Winter</i>	0.97 (0.82, 1.14)	NS
<i>Spring</i>	1.31 (1.12, 1.52)	<0.001

*Test for regression coefficients in the Poisson regression model.
NS, not significant (p>0.05); SSA, serious suicide attempts.

Table 2: Results of Poisson regression analysis (ORs and 95% CIs)

The result of a Poisson regression is summarized in Table 2. Compared to people living in rural areas, those in urban have a 79% decrease in fatality rate after controlling for sex, age group, education, occupation, method, and season. Compared to females, males have a 9% increase in fatality rate after controlling for other variables. Compared to those in the age group of 10-34, people in the age group of 35-49, 50-64, and 65 and more age groups have 19%, 68%, and 184% increase respectively in fatality rate, after controlling for other variables. Compared to those with secondary and above education, people with primary and illiterate education have 38% and 48% increase respectively in fatality rate after controlling for other variables. Compared to people with occupations other than farmers, those farmers have a 273% increase in fatality rate after controlling for other variables. Compared to people using pesticides to commit suicide, those using other poison, hanging, and using other methods have 96%, 25%, and 80% decrease respectively in fatality rate, after controlling other variables. Compared to cases in summer, those who suicide in autumn, winter, and spring have 4%, 3% decrease and, 31% increase respectively in fatality rate, after controlling for other variables. The p-values of male, age group of 35-49, suicide in autumn and winter are not significant(p-value >0.05), while others are statistically significant.

The result of multiple logistic regression and Poisson regression is slightly different. Most of

the CIs of variables in multiple logistic regression are wider than those in Poisson regression. The ORs obtained in the Poisson regression may be more precise than those obtained in multiple regression. In addition, people living in urban are more likely to die due to suicide estimated by a logistic regression model, while this is not the case in Poisson regression model. Hanging and all other methods are more likely to lead to death after suicide compared to pesticide, while they are not in Poisson regression. People committed suicide in autumn and winter are more likely to die by suicide than summer, while they are not in Poisson regression. Therefore, multiple logistic regression and Poisson regression both shows that those people who are male, who are older, who are with low education level, or as farmers, who commit suicide in spring, are more likely to die as a result of suicide.

In the diagnostic testing of the multiple logistic regression model, the variance inflation factor(VIF) is calculated and summarized in Table 3. Nearly all VIF scores are near to 1. This indicates the absence of collinearity, as the VIF scores are far less than 5. In addition, leverage values are calculated to evaluate the presence of outliers in the independent variables, we obtain 153 high-leverage points (Figure 1a). All the hat values of outliers are less than 0.1. In the estimation of Cook's distance, there are 148 observations with high Cook's distance values (Figure 1b). Furthermore, in Hosmer-Lemeshow chi-squared test, p-value is 0.442. The p-value is large enough that we fail to reject the null hypothesis that the model is not well fitted, at the 0.05 level of significance. We conclude that this multiple logistic regression model is poorly fitted. The Nagelkerke's R^2 is 0.523, meaning that about 52.3% of the variability in the dependent variable is explained by the independent variables.

In the diagnostic testing of the Poisson regression model, the VIF is calculated and summarized in Table 4. Nearly all VIF scores are less than 5, indicating the absence of collinearity. In addition, there are 55 high-leverage points (Figure 2a). Two of them have hat values of more than 0.1. In the estimation of Cook's distance, there is 63 observation with high Cook's distance values (Figure 2b). Furthermore, in the goodness of fit testing, residual deviance (731.91) is not equal to the degrees of freedom (465). Therefore, the data are overdispersed and the Poisson model may not be appropriate.

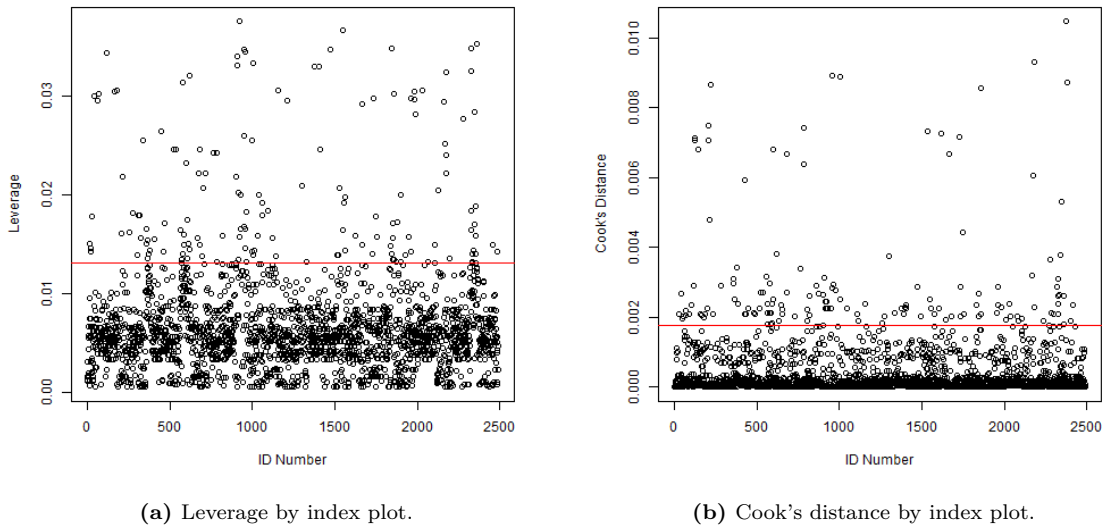


Figure 1: These plots depict the diagnostic testing results of the multiple logistic regression model.

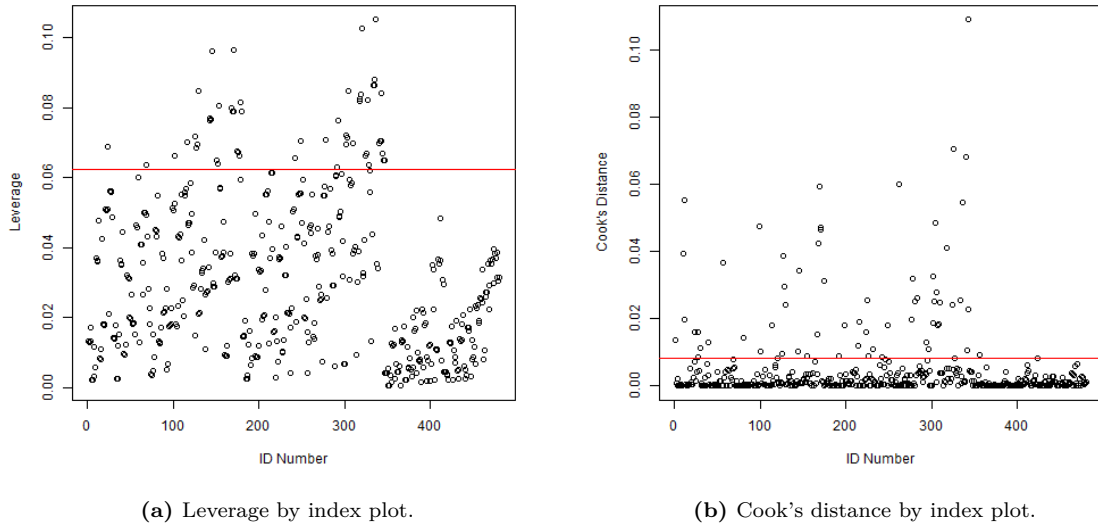


Figure 2: These plots depict the diagnostic testing results of the Poisson regression model.

	GVIF	DF	GVIF(1/(2*Df))
Urban	1.04	1	1.02
Sex	1.04	1	1.02
Age Group	1.50	3	1.07
Education	1.50	2	1.11
Occupation	1.07	1	1.04
Method	1.08	3	1.01
Season	1.03	3	1.01

Table 3: VIF for the detection of presence of collinearity in the diagnostic testing of multiple logistic regression.

	VIF
Residency: Urban	1.02
Sex: Male	1.02
Age Group: 35-49	2.22
Age Group: 50-64	2.91
Age Group: 65 and more	3.31
Education: Primary	1.60
Education: Illiterate	1.63
Occupation: Farming	1.02
Method: Other poison	1.01
Method: Hanging	1.06
Method: All others	1.03
Season: Autumn	1.47
Season: Winter	1.46
Season: Spring	1.52

Table 4: VIF for the detection of presence of collinearity in the diagnostic testing of Poisson regression.

Conclusion

In this study of fatality of suicide in a Chinese population, factors including age group, education, occupation, method, and season are tested to be influential on the fatality rate after suicide in both models. People who are living in urban areas, who are male, who are older, who are with low education, or with a farmer occupation and who hung to suicide have higher fatality rates as shown in the result of multiple logistic regression. Unlike the result in multiple logistic regression, in the Poisson regression, however, people living in rural areas have a higher fatality rate after suicide, and people who hung, who suicide in autumn or winter have a lower fatality rate due to suicide. And sex difference seems to be not significant in the fatality rate in Poisson regression model. This study gives suggestions about the targets which should be focused on and provided suicide prevention.

References

- [1] Ajdacic-Gross, V., Weiss, M., Ring, M., and et al (2008). Methods of suicide: international suicide patterns derived from the who mortality database. *Bull World Health Organ*, 86:726–32.
- [2] Chen, Z., Chen, Y., and Xie, Y. (2008). Report on the third national cause of death sampling survey. *Beijing:Peking Union Medical College*.
- [3] Phillips, M., Li, X., and Zhang, Y. (2002). Suicide rates in china,1995-99. *Lancet*, 359:835–40.
- [4] Sun, J., Guo, X., Zhang, J., Wang, M., Jia, C., and Xu, A. (2015). Incidence and fatality of serious suicide attempts in a predominantly rural population in shandong, china: a public health surveillance study. *BMJ Open*, 5(2):1405–1410.
- [5] Yang, G., Phillips, M., Zhou, M., and et al (2005). Understanding the unique characteristics of suicide in china: National psychological autopsy study. *Biomed Enciron Sci*, 18:379–89.