Investigative Depression Potential Influencing Factors of Studies

Peiwen Yan *

School of Mathematics, The University of Manchester, Manchester, M13 9PL, England

* Corresponding Author Email: peiwen.yan@student.manchester.ac.uk

Abstract. Depression, a mental health disorder, is increasingly becoming an important public health issue worldwide. With the increasing social pressure, interpersonal problems and life challenges, the incidence rate of depression is also rising. This situation has sparked widespread research interest to gain a deeper understanding of the underlying causes, pathophysiological mechanisms, and effective treatment methods of depression. This article will delve into the research progress of depression, including the latest scientific discoveries, diagnostic methods, and treatment strategies. We will use a combination of charts and logistic regression models to conduct research. We will focus on expert opinions from different fields and aim to provide a comprehensive overview to better understand depression and provide useful insights for future research and treatment directions. Depression is no longer an isolated problem, but a challenge that global society needs to face together. This paper looks forward to making more contributions through in-depth research to alleviate the pain of depression patients and improve the overall mental health level of society.

Keywords: Depression; logistic regression model; public health issue.

1. Introduction

Depression is a group of mental disorders characterized by persistent emotional depression, often accompanied by slow thinking and sluggish behavior, as well as various somatic symptoms [1]. Depression is common especially if it is brief, transient unhappiness mood can be said to be an experience that everyone has in life. Curt depression is a syndrome or syndrome involving mental health, health, and health a series of symptoms [1]. Moreover, symptoms such as experiencing guilt or feelings of low self-worth, diminished energy, difficulties in concentration, as well as the presence of suicidal and psychotic symptoms could also be observed.

The extent of a depressive episode is assessed based on the number of symptoms present, their severity, the duration of the episode, and the impact on one's social and occupational functioning. At the same time of depression, patients with depression may also be accompanied by some pain, such as physical discomfort, trunk pain, common cervical pain, back pain, muscle pain and other symptoms, so it is called hidden depression. Pain is a complex physiological and psychological response, and it is always accompanied by an emotional response. Pain can deepen depression. In terms of treatment, most patients with pain choose powerful analgesics to effectively control it [2]. This is very helpful for human headaches, neuralgia.

Pregnant women are also a very significant group of illness, which involves postpartum depression, clinical manifestations involve emotions, cognition, behavior, life and other aspects. The main characteristics are mainly related to babies and husbands, and it is generally believed that the cause is multi-factor [3]. However, postpartum depression is influenced by crucial factors such as alterations in endocrine functions, the mother's personality and psychological traits pre- and post-delivery, and the occurrence of certain stressful life events and prenatal complications, which play a significant role in its development. At present, the diagnosis is still lack of objective indicators, mostly based on a variety of symptoms self-assessment scale, which is determined by the corresponding score results after the maternal self-filling. The emergence of signs of depression is also connected to physical well-being concerns, adverse effects of medical treatment, life events, and societal influences, in addition to alcohol and substance misuse. Moreover, these factors have the potential to induce symptoms of depression in individuals across the board.

The global lifetime prevalence of depression is roughly 15%, but the incidences of episodes with a severity level not meeting the requirements for a depressive diagnosis are far more prevalent. Actigraph recordings of motor activity are considered an objective method for observing depression, although this topic is far from exhaustive studied within psychiatric research [4].

This study will utilize a logical model, which is an abstract tool designed to unveil the relationship between disease symptoms and potential underlying causes. It achieves this by integrating clinical data, biological knowledge, and medical expertise to establish a theoretical framework for understanding pathological physiology, molecular mechanisms, as well as physiological and psychological processes related to diseases. Within the logical model, disease symptoms are segmented into distinct components, and the interconnections among these components are represented as logical associations [5].

The logical model can serve several purposes in the analysis of disease symptoms, Firstly, Causal Analysis: the logical model aids in identifying possible causes by linking various biological, environmental, and lifestyle factors, thereby helping to reveal the fundamental origins of the disease. Secondly, Diagnostic Support: by analyzing symptoms and other relevant information, the logical model can assist medical professionals in making more accurate diagnoses and avoiding the omission of critical details. Then, Treatment Strategies: the logical model can help predict the effects of different treatment methods, guiding the development of personalized treatment plans. Next, Research Guidance: Within medical research, the logical model can help researchers comprehend the mechanisms underlying a disease, providing direction for experimental design and data analysis. Finally, Prevention and Management: the logical model can be applied to formulate strategies for disease prevention and management, aiding in the identification of risk factors and the implementation of corresponding measures [6].

Consequently, the logical model is a potent tool that facilitates a systematic analysis of disease symptoms, elucidates potential cause-and-effect relationships, and furnishes scientific support for medical decision-making. Nonetheless, in practical applications, it is crucial to consider data accuracy and the limitations of the model, while also integrating clinical experience and professional knowledge to make comprehensive judgment.

2. Methodology

2.1. Data Source

The depression data set used in this paper were collected from the Kaggle website. which was consists as a study about the life conditions of people who live in rurales zones. The data set is owned by Frankcc while adjusted by collaborator Diego Babativa.

2.2. Variable Selection

The dataset employed in this study encompasses a total of 299 individuals, encompassing both those who are affected by depression and those who are not. Among these, 27 individuals are male, while 272 individuals are female. The ages of the patients in the dataset span from 17 to 85 years. This dataset comprises 20 distinct variables, which include gender, age, marital status, number of children, educational attainment, total family members, acquired assets, durable assets, savings, living expenses, other expenditures, salary income, personal farm income, business income, non-business income, agricultural income, farm-related expenses, primary labor input, long-term investments, and non-long-term investments.

Table 1. The number of depressions with three outstanding characteristics

Elements	variables	Number of total	Number of depression
Married	X_1	299	39
Male	X_2	27	11
Female	X_3	272	44

The table 1 above illustrates the number of depression regard to few remarkable characteristics. It is easy to find that marriage may not be a direct factor to being depression. While there are nearly half of male and a quarter of females are depressed.

The Table 2 above illustrates few constraints which the number is determined by the median number.

Table 2. The number of depressions with other 12 characteristics

Elements	Classification	Number of total
Number children	=>3	168
	<3	131
education level	=>9	200
	<9	99
total member in the family	=>5	205
	<5	94
save asset	=>23399979	270
	<23399979	29
living expenses	=>26692283	206
	<26692283	93
other expenses	=>28203066	189
	<28203066	110
incoming salary	1	49
incoming own farm	1	84
incoming no business	1	73
incoming agricultural	=>30028818	213
	<30028818	86
farm expenses	=>31363432	200
	<31363432	99
labor primary	1	60

2.3. Research Protocol

This paper uses the logistic Regression model, whether or not to have depression is the dependent variable(Y), and the 20 factors are the independent variables(X), where 0 represents no and 1 represents yes.

$$P(y=1|x) = \frac{1}{1+e^{-(\beta_0+\beta_1x_1+\beta_2x_2+\beta_3x_3+\cdots+\beta_nx_n)}} \quad (1)$$

Where $P(y=1|x)$ is the probability that the dependent variable y is 1 given the independent variable x. e is the base of the natural logarithm. β_0 is the intercept. β_1 is the coefficient for the independent variable x.

3. Results and Discussion

3.1. Descriptive Analysis

In Figure 1, the largest proportion is occupied by families that have given birth to six to eight children. Based on the analysis, it is evident that having an excessive number of children places significant financial strain on a family. Each child requires expenses for food, nutrition, and clothing, and when the number is substantial, the overall expenditure rises significantly. Not all families can afford the expenses associated with having so many children. Raising numerous children under limited economic circumstances means that each child receives fewer resources, making their upbringing more challenging compared to families with fewer children. Consequently, this puts substantial financial pressure on the family and can lead to family depression.

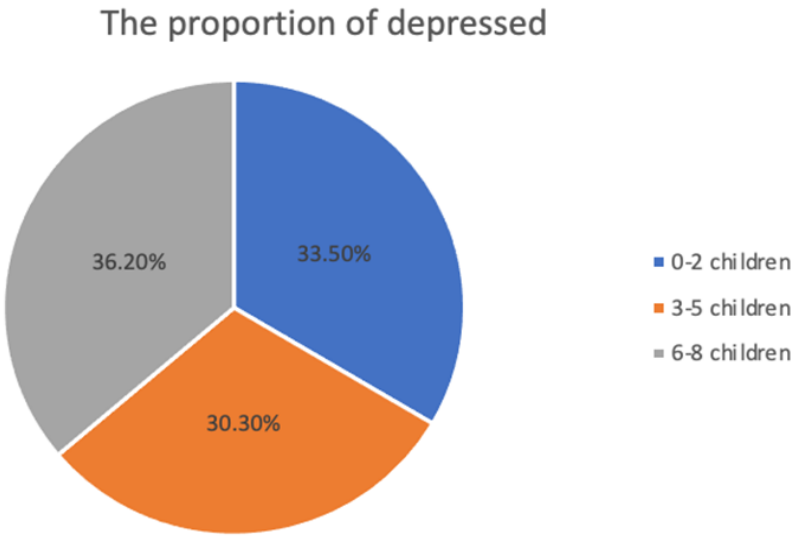

Fig. 1 Relationship between number of births and depression

From the perspective of the group with the lowest prevalence of depression, having three to five children doesn't impose as much pressure on parents, and having multiple children can bring a richer and more colorful life, reducing the probability of the family experiencing depression.

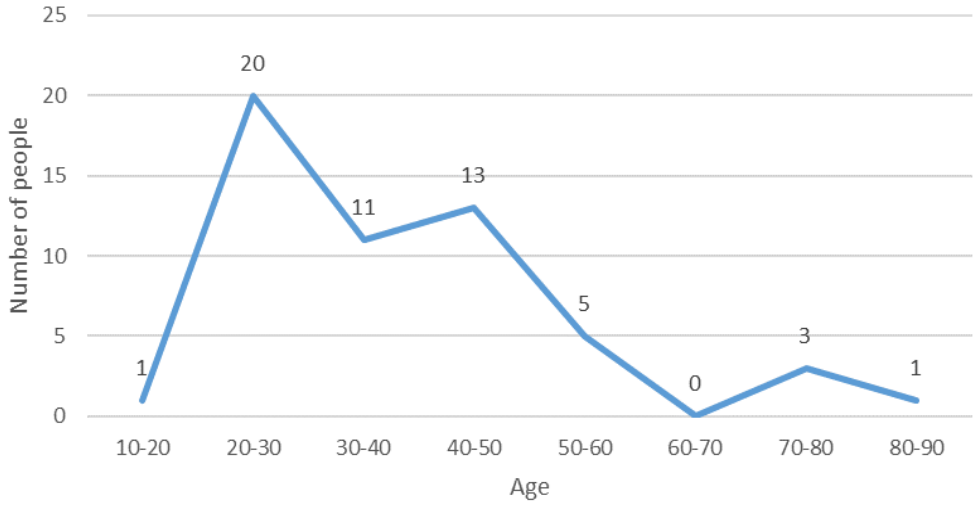

Fig. 2 Relationship between age and depression

In Figure 2, it is evident that there are significant differences in the prevalence of depression among different age groups. For instance, the age group between twenty and thirty years old shows a notably higher rate of depression compared to other age groups, with the figure approaching nearly twenty individuals. According to the survey analysis, this age group predominantly comprises young individuals, including students who are still in school and young adults who are either in the process of job hunting or at the early stages of their careers. These individuals are navigating significant life transitions, such as graduating and job searching, which can introduce substantial stressors. Given the importance of securing employment after graduation, the added pressure they face can lead to depression.

Conversely, the age group with the least prevalence of depression in the chart is between sixty and seventy years old, where there are almost no reported cases of depression. According to the survey findings, individuals in this age range are typically retired seniors. They have transitioned away from the demands of busy work lives and are now enjoying their retirement years, leading carefree lives. This carefree lifestyle significantly reduces the risk of depression among this age group.

3.2. Model Results

In Table 3, an array of independent variables, including gender (specifically, female), age, marital status, number of children, total household members, educational attainment, accumulated assets, savings, living expenses, miscellaneous expenditures, salary earnings, personal agricultural income, non-business-related earnings, agricultural income, farm-related expenditures, and primary labor status, have been employed in a multi-classification Logit regression analysis. The dependent variable under scrutiny is the presence of "depression." As discernible from the tabular representation, a comprehensive analysis was conducted with the participation of a total of 299 samples. The model's outcomes underscore the observation that individuals afflicted with depression constitute a distinct minority within the dataset.

Table 3. Basic summary of multi-class logistic regression analysis

name	Options	Frequency	percentage
depressed	0.0	244	81.61%
	1.0	55	18.39%
	total	299	100.0

Table 4. Multi-classification logistic regression model likelihood ratio test

Model	log-likelihood	Chi-square value	df	p	AIC	BIC
intercept	285.441	-	-	-	-	-
Final model	264.319	21.121	16	0.174	298.319	361.227

Table 5. Summary of multi-class logistic regression analysis results

1.0	coefficients	S.E.	z	Wald $\chi 2$	p	OR-95% CI
sex（F）	-1.540	0.513	-3.003	9.017	0.003	0.078 ~ 0.586
Age	0.007	0.014	0.522	0.272	0.602	0.980 ~ 1.035
Married	0.101	0.440	0.229	0.052	0.819	0.467 ~ 2.620
Number_children	-0.084	0.137	-0.614	0.377	0.539	0.703 ~ 1.202
total_members	0.116	0.135	0.860	0.739	0.390	0.862 ~ 1.464
education_level	-0.076	0.067	-1.134	1.286	0.257	0.812 ~ 1.057
gained_asset	0.000	0.000	0.429	0.184	0.668	1.000 ~ 1.000
save_asset	0.000	0.000	2.083	4.337	0.037	1.000 ~ 1.000
living expenses	-0.000	0.000	-1.275	1.625	0.202	1.000 ~ 1.000
other_expenses	0.000	0.000	0.157	0.025	0.875	1.000 ~ 1.000
incoming_salary	-0.208	0.922	-0.225	0.051	0.822	0.133 ~ 4.950
incoming_own_farm	-0.524	0.409	-1.282	1.644	0.200	0.266 ~ 1.319
incoming_no_business	0.207	0.370	0.560	0.314	0.575	0.596 ~ 2.538
incoming_agricultural	-0.000	0.000	-0.761	0.579	0.447	1.000 ~ 1.000
farm expenses	0.000	0.000	1.297	1.683	0.194	1.000 ~ 1.000
labor_primary	0.156	0.862	0.180	0.033	0.857	0.216 ~ 6.328
intercept	-0.593	1.291	-0.459	0.211	0.646	0.044 ~ 6.945

Incorporating the data from both Table 4 and 5, a comprehensive set of 16 variables encompassing factors such as gender (F), age, marital status, number of children, total family members, education level, accumulated assets, savings, living expenses, miscellaneous expenditures, salary, individual farm income, absence of business income, agricultural income, farm expenses, and primary labor engagement are treated as independent variables in the context of a multi-class logistic regression analysis. This analytical approach aims to ascertain their influence on the dependent variable,

"Depression," which contains two distinctive items, with 0.0 serving as the reference point. Consequently, the outcome yields a singular formula for the final model, articulated as follows:

In relation to the reference value of 0.0, while controlling for all other variables set at 1.0, it is observed that the regression coefficient for gender (F) stands at -1.540, signifying statistical significance at the 0.01 level ($z=-3.003$, $p=0.003<0.01$). This compellingly suggests that gender (F) exerts a substantial and negative impact on depression. Furthermore, the odds ratio (OR value) is determined to be 0.214, underscoring the notion that a unitary increment in gender (F) corresponds to a mere 0.214-fold reduction in depression.

Conversely, when scrutinizing the regression coefficient of savings (save_asset), it becomes evident that this variable is assigned a value of 0.000, thus establishing statistical significance at the 0.05 level ($z=2.083$, $p=0.037<0.05$). This nuanced finding implies that savings (save_asset) wield a noteworthy and positive influence on depression. The odds ratio (OR value) is precisely 1.000, signifying that a unitary augmentation in savings is directly associated with a proportionate 1.000-fold escalation in depression. Notably, other variables under consideration do not appear to exert substantial influence on the outcome.

Table 6. Forecast accuracy summary.

		Predictive value		Prediction accuracy	prediction error rate
		0.0	1.0		
actual value	0.0(n=244)	240	4	98.36%	1.639%
	1.0(n=55)	52	3	5.45%	94.545%
Summary				81.27%	18.73%

Examining Table 6 reveals that the predictive accuracy attains an impressive 98.36% when the actual value is 0, but regrettably dwindles to a mere 5.45% when the correct value is 1. These outcomes can be attributed to several factors, foremost among them being the limited sample size when the true value is 1. Furthermore, a shortfall in the model's performance compared to the initial expectations curtails its utility in practical applications.

4. Conclusion

Data indicates that women have a higher likelihood of experiencing depression compared to men, particularly in households with multiple children. On a global scale, the prevalence of depression tends to be greater among women when compared to men. According to the World Health Organization (WHO), women are about twice as likely to suffer from depression as men. Depression is most likely to occur in 20-30 years old. At the same time, save_asset is also related to depression. At the same time, it was found that the level of family income also affected the occurrence of depression. This study was used to identify and evaluate various factors that may influence the risk of depression, including life events, family background, social support, genetic factors, etc. By analyzing these factors, researchers can better understand potential causes of depression and develop targeted prevention strategies. In the future, this research could allow doctors to better tailor treatment plans to each patient's specific situation and improve treatment effectiveness. This can lead to increased focus on early intervention and prevention of mental health issues, ultimately lowering the occurrence of such problems through the identification of potential risk factors and the provision of timely support.

References

[1] Shu L. Evaluation and treatment of depression. Chinese Medical Journal, 1999, 398-400.

[2] Chen Y J, Zhong Y B. Postpartum depression. Journal of Practical Obstetrics and Gynecology, 2000.

[3] Huang Ruxun, Lu Lin. Chronic Pain and Depression. Chinese Clinical Rehabilitation, 2002.

[4] Carney, et al. Major depressive disorder in coronary artery disease. American Journal of Cardiology, 1987, 60: 1273-1275.

[5] Damsker, et al. Canadian Network for Mood and Anxiety Treatments (CANMAT) clinical guidelines for the management of major depressive disorder in adults. III. Pharmacotherapy. Canadian Journal of Psychiatry Revue Canadienne De Psy, 2009, 117.1 S54-S64.

[6] Mckinnon M, et al. A meta-analysis examining clinical predictors of hippocampal volume in patients with major depressive disorder. Journal of Psychiatry & Neuroscience, 2009, 34-141.

[7] Greene W. Econometric Analysis. Prentice Hall, 2003.

[8] Peterson J. Regression Analysis of Count Data. Technometrics, 1999, 41(4): 1.

[9] Da H M. A Causal Model of Factors Influencing English Academic Performance of Middle School Students. Psychological Science, 2005, 28 (4): 5.

[10] Feng Lixia. Research on Factors Influencing Academic Performance of College Students. China Electric Power Education, 2013.