

## Research Article

# Quantile Regression Analysis of Depression and Clinical Symptom Degree in Chinese Patients with Spinocerebellar Ataxia Type 3

Xiaoping Chen, Lihui Zheng , and Jianqi Yao

*Department of Statistics, College of Mathematics and Informatics & FJKLMAA, Fujian Normal University, Fuzhou 350000, China*

Correspondence should be addressed to Lihui Zheng; [fbx20180024@yjs.fjnu.edu.cn](mailto:fbx20180024@yjs.fjnu.edu.cn)

Received 26 February 2020; Accepted 24 April 2020; Published 10 June 2020

Academic Editor: Chenxi Huang

Copyright © 2020 Xiaoping Chen et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Spinal cerebellar ataxia type 3 is a common SCA subtype in the world. It is a neurodegenerative disease characterized by ataxia. Patients exhibit common neuropsychological symptoms such as depression and anxiety. Some patients have suicidal tendencies when they are severely depressed. So, it is very important to study the severity of depression and clinical symptoms (SARA), to find out the patient's psychological state in time and to help patients actively respond to treatment. A total of 97 Chinese SCA3 patients were enrolled in the study. The Beck Depression Scale was used to investigate the prevalence of depression in the confirmed patients. The distribution of depression data in these patients was investigated. Then, the quantifier was used to model the depression status of Chinese SCA3 patients. An analysis was conducted to identify the key factors affecting depression under different quantiles. Studies have shown that SARA and gender are important factors affecting depression; the effect of initial SARA is small, then the degree of influence increases, and the degree of influence decreases in the later period, but it is always positively correlated with depression; the development of women's SARA is gentler than that of men, and the degree of depression is lower than that of men.

## 1. Introduction

Machado–Joseph disease or spinocerebellar ataxia 3 (MJD/SCA3) is a clinically heterogeneous, neurodegenerative disorder characterized by varying degrees of ataxia, ophthalmoplegia, peripheral neuropathy, pyramidal dysfunction, and movement disorder. MJD/SCA3 is caused by a CAG repeat expansion mutation in the protein coding region of the ATXN3 gene located at chromosome 14q32.1 [1]. So far, the pathogenesis and pathological effects of such diseases have not been clarified, and there was no effective treatment. Conventional treatment can only improve the clinical symptoms.

Previous studies on SCA3 focused the pathogenesis of SCA3, CAG mutation amplification, and ethnic differences. In the study of the role of the proteasome in the pathogenesis of SCA3/MJD, it was found that the proteasome plays a direct role in suppressing polyglutamine aggregation in disease. Also, the result suggested that modulating

proteasome activity is a potential approach to altering the progression of this and other polyglutamine diseases [2]. The CAG repeat amplified in the SCA3 gene caused the patient's restless legs syndrome (RLS) and sleep impairment, and impaired sleep in SCA3 was associated with older age, long-standing disease, and brainstem involvement [3]. There were significant differences in predominant spinopontine atrophy, lack of dystonic features, and larger CAG repeat expansion between families with spinal cord ataxia 3 in two different ethnic origins in the United States [4]. The frequency of trinucleotide repeats in different ataxic categories in Chinese patients was studied to find that the frequency of SCA3/MJD is substantially higher than that of SCA1 and SCA2 in patients with autosomal dominant SCA from Chinese kindreds, who are non-Portuguese. Dementia and hyporeflexia were more frequent in patients with SCA2, while spasticity, hyperreflexia, and Babinski signs were more frequent in patients with SCA3/MJD, and those might be helpful in clinical work to primarily distinguish patients with

SCA3/MJD and SCA2 from others with different types of SCA [5].

SCA3 not only affected patients' ataxia and cognitive ability but also led to depression or anxiety with the development of the disease, seriously affecting patients' physical and mental health. Foreign studies, namely, the study of stuneuropsychological and neurobehavioral characteristics of SCA3 patients and the degree of emotional dysfunction of patients, found that, in addition to anxiety and depression, SCA3 patients also showed the development of increased apathy [6]. The research on the influence between SCAs depression and ataxia found that depression was very common in SCA and would not develop in 2 years. The suicidal ideation of SCA3 patients was significantly higher than that of other types, and there were differences in the suicidal ideation of SCA patients in different regions (North America and Europe). The effect of depression on the progression of ataxia varies by the SCA type [7]. Depression in patients with SCA3 has also been studied in China. Through the establishment of multiple linear equations, it was found that depression is caused by the movement disorder caused by ataxia. In addition, the two had interaction, and gender and ICARS were important factors affecting depression [8].

In summary, domestic and foreign scholars have conducted a large number of in-depth studies on the factors affecting the incidence of SCA3 and other related content. However, there are a few studies on depression in Chinese SCA3 patients, and the main method is linear regression. There are a few studies combining the quantile regression method with the influencing factors of depression in Chinese SCA3 patients. Considering the quantile, regression can effectively avoid the heteroscedasticity and nonnormal distribution of the data, and with the change of the quantile used, it can more accurately describe the influence of independent variables on dependent variables and characterize the conditional distribution. Therefore, it is necessary to study the factors affecting the quantile regression in SCA3 patients with depression.

## 2. Objects and Methods

**2.1. Study Subjects.** Participants in this study included 97 patients with confirmed SCA3, all of whom over 18 years of age. This study was approved by the ethics committee of the first affiliated hospital of Fujian Medical University. Written informed consent forms were signed by all subjects.

**2.2. Observation Method.** Beck depression inventory (BDI) was used to evaluate the severity of depression in SCA3 patients. There were 21 questions in the whole scale, and each question was divided into different grades. According to the total score of the scale, the severity of depressive symptoms can be evaluated. In this study, a BDI score greater than or equal to 19 was defined as clinically related depression [8].

The severity of the patient was assessed using the ataxia rating scale (SARA). SARA is a semiquantitative neurologic

functional assessment scale that describes and quantitatively evaluates symptoms of cerebellar ataxia.

**2.3. Statistical Analyses.** Firstly, the data set was descriptively analyzed. The K-S discriminant method was used to analyze the normal distribution of data variables, and the mean and median values of the variables in the nondepressive group and the depression group were calculated.

Secondly, using univariate and multivariate quantile regression, different vacancies were taken, and the significance and correlation of each variable were judged according to the test  $P$  value. The density function maps of depression and SARA were plotted, as well as the density function maps by gender. The density function maps of depression BDI scores at 10% and 90% of the SARA were plotted. The dependent variable was the depression BDI score (Beck), and the independent variables were the gender, SARA score, disease duration, age of onset, small CAG, and large CAG.

Quantile regression [9] was a modeling method to estimate the relationship between regression variable  $x$  and dependent variable  $y$  at any probability level. Suppose the distribution function of the random variable was  $F(y) = \text{Prob}(Y \leq y)$ , and define the  $\tau$  quantile of  $y$  to be

$$Q(\tau) = \inf\{y: F(y) \geq \tau\}, \quad 0 < \tau < 1. \quad (1)$$

When  $\tau$  was equal to 1/2, that was the median. In the quantile regression model, the loss function was defined as a piecewise linear function:

$$\rho_\tau(u) = (\tau - I(u < 0))u, \quad u = y_i - \hat{\xi}, \quad (2)$$

where  $\hat{\xi}$  was the expectation  $E(y)$  of the dependent variable and  $I$  was the indicator function. The basic idea of regression analysis was to minimize the distance between the sample value and the fitting value, so the optimization problem from the loss function expectation can be converted to another form:

$$\min \sum_{i=1}^n \rho_\tau(y_i - \hat{\xi}). \quad (3)$$

Given the information set  $x$  and condition  $Q_\tau(\varepsilon | x) = 0$ , the conditional quantile function of  $y$  can be expressed as  $Q_y(\tau | x) = x' \beta_\tau$ .

The software tool was the R 3.5.1 version. Model results were significant at  $P < 0.05$  and were statistically significant.

## 3. Result

**3.1. Variable Descriptive Statistics.** Of the 97 patients with SCA3, 41 were nondepressed and 56 were depressed. The depression rate was 57.73%, and SCA3 patients had a higher depression rate. The mean course of the patients with depression ( $9.259 \pm 8.0$ ) was significantly larger than that of the nondepressed group ( $6.329 \pm 6$ ); SARA ( $13.22 \pm 13$ ) was higher than that of nondepressed patients ( $8.085 \pm 7$ ). The depression degree ( $33.21 \pm 32$ ) in the depression group was 4 times higher than that in the nondepression group

TABLE 1: Descriptive statistics of variables.

|                      | Distribution | Nondepressed group | Depression group |
|----------------------|--------------|--------------------|------------------|
| Number               | ---          | 41                 | 56               |
| Gender (M/F)         | ---          | 19/22              | 35/21            |
| CAG small            | Skewed       | 20.9 ± 20          | 21.57 ± 14       |
| CAG large            | Normal       | 74.34 ± 74.34      | 74.84 ± 75       |
| Age at onset (years) | Normal       | 35.24 ± 34         | 33.59 ± 34.5     |
| Disease duration     | Skewed       | 6.329 ± 6          | 9.259 ± 8.0      |
| SARA                 | Skewed       | 8.085 ± 7          | 13.22 ± 13       |
| BDI score            | Normal       | 8.878 ± 9          | 33.21 ± 32       |

TABLE 2: Univariate regression results.

|                      | OLS                     | $Q_{\tau}(0.1   X)$ | $Q_{\tau}(0.25   X)$ | $Q_{\tau}(0.5   X)$      | $Q_{\tau}(0.75   X)$ | $Q_{\tau}(0.9   X)$ |
|----------------------|-------------------------|---------------------|----------------------|--------------------------|----------------------|---------------------|
| CAG small            | -0.005                  | $1.202e-17$         | 0.233                | -0.116                   | 0.071                | -0.02               |
| CAG large            | 0.032                   | 0.333               | 0.75                 | 0.25                     | -0.833               | -1.33 ** (0.00092)  |
| Age at onset (years) | 0.059                   | 0                   | -0.154               | -0.2                     | 0.36                 | 0.429 ** (0.0005)   |
| Disease duration     | 0.924 ** (0.00101)      | 0.538               | 1.190 * (0.012)      | 1.333 ** (0.00036)       | 0.9375 * (0.012)     | 0.571 ** (0.005)    |
| SARA                 | 0.952 ** ( $3.25e-06$ ) | 0.727 ** (0.0043)   | 0.868 ** (0.0024)    | 1.273 ** ( $3.044e-05$ ) | 0.765 ** (0.004)     | 0.647 ** (0.00057)  |

Note. (1) If the P value is less than 0.01, it is considered that the effect of the independent variable is highly significant, which is represented by “\*\*.” (2) If the P value is less than 0.05, it is considered that the effect of the independent variable is significant, which is indicated by “\*.”

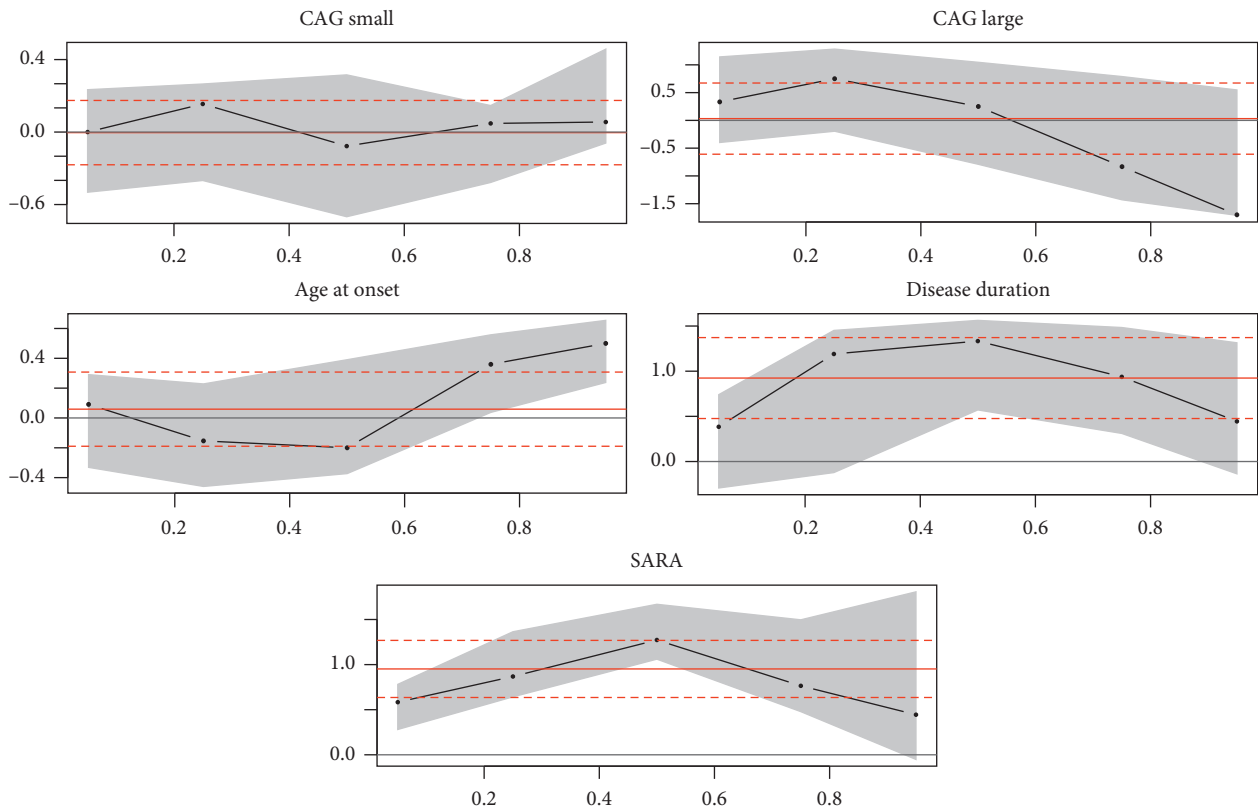


FIGURE 1: Variation of the quantile coefficient.

(8.878 ± 9), and the depression degree was relatively deep. In the data, the small CAG, the disease duration, and SARA were skewed. The large CAG, the age of onset, and Beck were

normally distributed. Therefore, it is more appropriate to use the quantile regression relative least squares regression (the data are shown in Table 1).

TABLE 3: Estimated values of each quantile parameter.

|             | OLS                 |            | $Q_{\tau}(0.05   X)$ |             | $Q_{\tau}(0.1   X)$  |            | $Q_{\tau}(0.25   X)$ |            |
|-------------|---------------------|------------|----------------------|-------------|----------------------|------------|----------------------|------------|
|             | Estimated value     | $P$        | Estimated value      | $P$         | Estimated value      | $P$        | Estimated value      | $P$        |
| (Intercept) | 14.223              | $2.15e-07$ | -3.26                | 0.189       | 0.2                  | 0.920      | 4.143                | 0.170      |
| Gender (F)  | -6.196*             | 0.0181     | -6.282*              | 0.038       | -3.4                 | 0.296      | -4.143               | 0.184      |
| SARA        | 1.036**             | $4.48e-07$ | 1.026**              | $1.098e-07$ | 0.8**                | $4.08e-07$ | 0.905**              | $7.7e-4$   |
|             | $Q_{\tau}(0.5   X)$ |            | $Q_{\tau}(0.6   X)$  |             | $Q_{\tau}(0.75   X)$ |            | $Q_{\tau}(0.9   X)$  |            |
|             | Estimated value     | $P$        | Estimated value      | $P$         | Estimated value      | $P$        | Estimated value      | $P$        |
| (Intercept) | 11.571              | $3.3e-03$  | 31.795               | 0.58844     | 24                   | $5.0e-09$  | 34                   | $1.25e-10$ |
| Gender (F)  | -6.0                | 0.102      | -10.506**            | 0.00049     | -9*                  | 0.019      | -5                   | 0.18       |
| SARA        | 1.238**             | $3.1e-06$  | 1.010**              | 0.00000     | 1*                   | $1.48e-05$ | 0.67*                | 0.02       |

Note. (1) If the  $P$  value is less than 0.01, it is considered that the effect of the independent variable is highly significant, which is represented by “\*\*.” (2) If the  $P$  value is less than 0.05, it is considered that the effect of the independent variable is significant, which is indicated by “\*.”

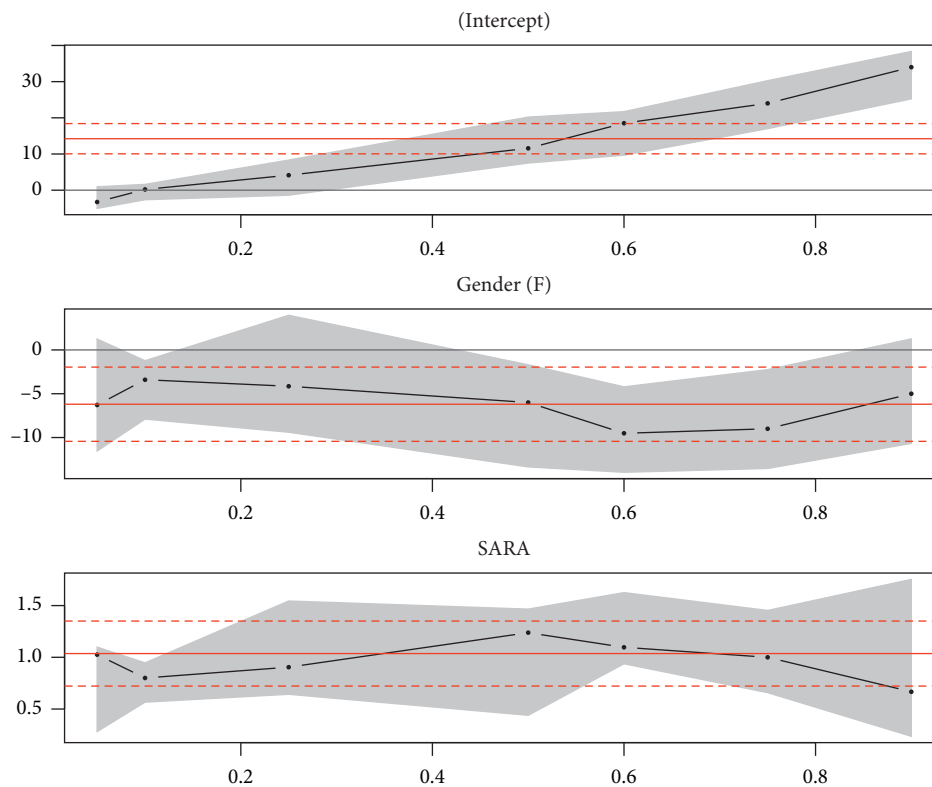


FIGURE 2: Variation of the quantile coefficient.

3.2. *Univariate Regression Analysis.* The Beck depression score was used as the dependent variable, and the CAG small, CAG large, age of onset, course of disease, and SARA were used as independent variables. Quantile regression and OLS regression were used to model the variables. The results were as shown in Table 2. From the regression results of the OLS model, it can be seen that in the linear regression model, the duration of disease, and SARA have a significant positive effect on depression. The longer the duration of disease is, the greater will be the SARA score and the more severe the depression will be. However, the CAG small, CAG large, and age of onset had no effect on depression.

The OLS model required that all variables obey normal distribution, but some of the data were biased. The validity of

the OLS model fitting was limited, while the quantile regression did not require normal distribution. Columns 3–7 showed the quantile regression coefficients, and  $P$  values in parentheses. Compared with the regression coefficients of the OLS model, the quantile regression coefficients changed. The variation of the coefficients is shown in Figure 1, where the red horizontal line represents the OLS regression coefficient and the curve represents variation of the quantile regression coefficient.

In the quantile regression, the CAG was not significant at each quantile. The CAG and the onset age were not significant at other quantile, and only had an effect on depression at 90%. At this time, CAG was negatively correlated (-1.33). The age

of onset was positively correlated (0.429). From the low quantile to the high quantile, the CAG large coefficient symbol changed from positive to negative, and the onset age coefficient symbol changed from negative to positive.

SARA was significant at each quantile. The course of the disease was 10%, and the others were significant. The two variables were positively correlated with depression and the same as the univariate linear regression coefficient. When considering the univariate influence on depression, SARA and disease duration were important variables.

**3.3. Multivariate Quantile Regression.** In the establishment of multiple quantile regression, variable modeling with higher significance should be selected as far as possible. So, the course of disease and SARA were selected as independent variables. In the study of Lin et al. [8], gender was found to be an important factor affecting depression. So, gender was also selected as an independent variable. Through preliminary modeling, it was found that the disease course coefficient is not significant. Therefore, the disease course was excluded, and only the quantile model of SARA and gender was established.

The R software was used to conduct quantile regression on SCA3 patient data. Table 3 lists the parameter estimation, P value, and significance results of different quantile points  $\tau = (0.05, 0.1, 0.25, 0.5, 0.6, 0.75, 0.9)$  in quantile regression.

Comparing the results of parameter estimation under 7 quantiles, it can be seen that SARA has always played a significant role, and the gender was significant when the quantile was 0.05, 0.6, and 0.75. Figure 2 is the parameter estimation chart of each variable under each quantile. In the figure, it can be seen that although the coefficients of the SARA points that have a significant effect on depression are changing, they were always positive, indicating a positive correlation with depression. When the quantile was small, it can be seen that the SARA coefficient has a decreasing trend. It showed that the effect of SARA on depression was small in the initial period of time. With the increase of SARA, the degree of depression will be aggravated, and the decrease in the coefficient value indicated that the development of SARA to the later stage has a reduced impact on depression.

Gender (female) was negatively correlated with depression, and only became significant when the coefficient value was less than about  $-6$ . When the depression score was used as the dependent variable, gender and SARA were used as independent variables to establish linear multiple regression. The coefficient of gender (female) was  $-6.196$ , the quantile at this time was between about 0.5 and 0.8, and the degree of depression in women was lower than that in men.

**3.4. Density Function Diagram.** Furthermore, the density function diagram of BDI scores and SARA, as well as the density function diagram of male and female, was drawn. From Figure 3, it can be more intuitively seen that the BDI score distribution was close to the normal distribution, and SARA presented a skewed distribution. The peak BDI scores of women and men were different, about 10 for women and

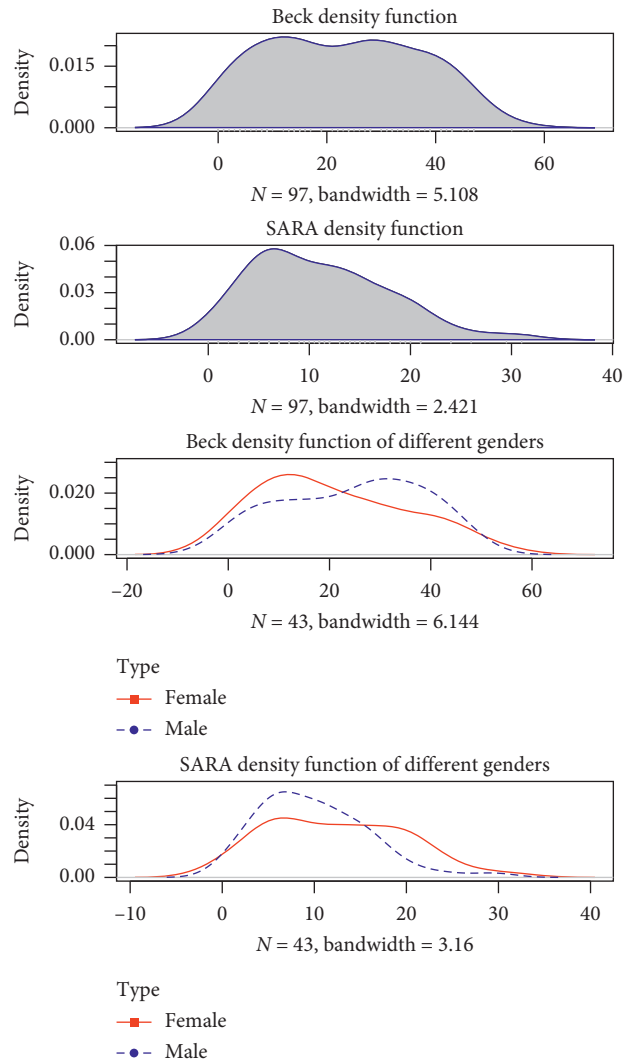


FIGURE 3: BDI score and the SARA density function.

35 for men. The distribution of the SARA density function was relatively flat for women and steep for men.

In order to study the effect of SARA on depression, the effects of clinical symptom severity on low and high scores were compared, and the comparison of BDI scores at 10% and 90% of SARA was drawn (90% is the blue dotted line). From Figure 4, it can be found that the BDI scores of SARA at different points were significantly different. The higher the value of SARA, the higher the BDI score. The right figure reflected the density image of the BDI score in the case of a large SARA value and small SARA value. The larger SARA value was, the larger the BDI score was. The BDI score with a small SARA value was concentrated around 5, while the BDI score with a large SARA value was concentrated around 40.

#### 4. Discussion

In this study, SCA3 patients had a higher rate of depression, and SARA had a significant effect on depression. In the early stage of development, SARA had less effect on depression. At this time, the patient's depressive symptoms were not

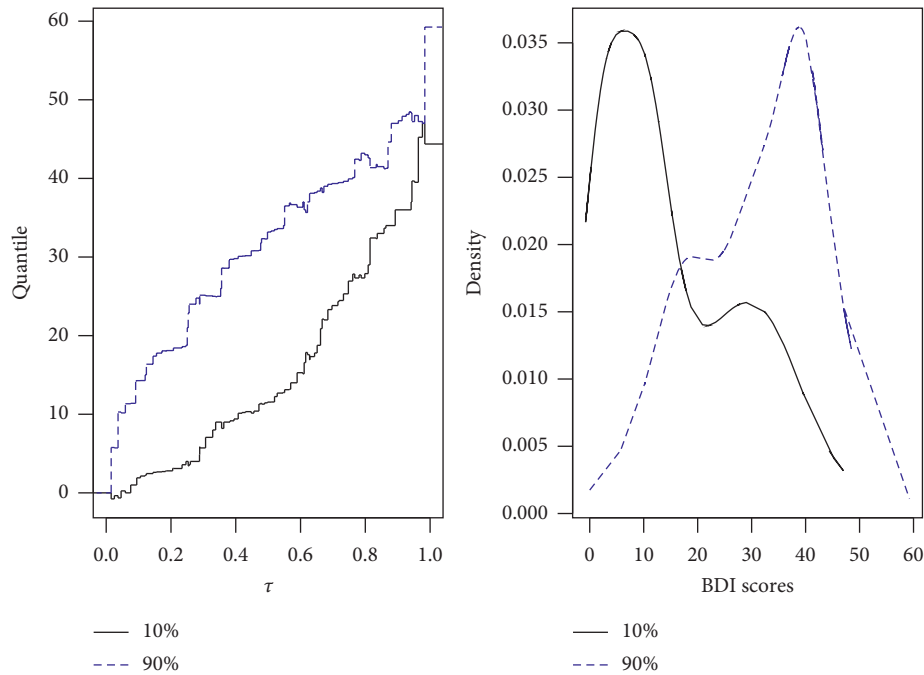


FIGURE 4: BDI score density map for SARA at 10% and 90%.

obvious. With the aggravation of clinical symptoms, depression of patients will be more affected and may be accompanied by anxiety, apathy, and other manifestations. By the end of the disease, SARA still had an effect on depression, but the degree of influence was reduced. Lo et al. [7] mentioned in their study that depression was very common in SCA and will not develop in 2 years. Therefore, it was speculated that the initial period when SARA has a small impact on depression may be 2 years. After 2 years, we need to pay attention to the impact of SARA on patients' depression. Some SCA3 patients in this article were also accompanied by suicidal tendencies, lethargy, fatigue, etc, which seriously affected the quality of life and physical health of patients. From the early stage of the disease, it was necessary to pay attention to the depressive symptoms and psychological state of SCA3 patients. The treatment of depressive symptoms may be beneficial to improve the clinical symptoms of patients and reduce the risk of suicide.

Gender was also an important factor affecting patients' depression. In this study, the depression degree of women was lower than that of men. Considering that the graph of SARA density function of women presents a relatively flat state and that of men was rather steep, the difference in SARA development between men and women may be the reason for the difference in depression. Therefore, SARA development of SCA3 patients of different genders can be considered in subsequent studies and targeted treatment programs for depression.

### Data Availability

The data used to support the findings of this study are included within the supplementary information files.

### Conflicts of Interest

The authors declare no conflicts of interest.

### Acknowledgments

The authors sincerely thank the participants for their help and willingness to participate in this study. This study was supported by the National Natural Science Foundation of China (11601083 and U1805263), the Program for Probability and Statistics: Theory and Application (IRTL1704), and Innovative Research Team in Science and Technology in Fujian Province University (IRTSTFJ).

### Supplementary Materials

The authors provide the data in the supplementary information files that they submit alongside their manuscript. (*Supplementary Materials*)

### References

- [1] A. D'Abreu, M. C. França, H. L. Paulson et al., "Caring for Machado-Joseph disease: Current understanding and how to help patients," *Parkinsonism and Related Disorders*, vol. 16, no. 1, pp. 2–7, 2010.
- [2] Y. Chai, S. L. Koppenhafer, S. J. Shoesmith, M. K. Perez, and H. L. Paulson, "Evidence for proteasome involvement in polyglutamine disease: localization to nuclear inclusions in SCA3/MJD and suppression of polyglutamine aggregation in vitro," *Human Molecular Genetics*, vol. 8, no. 4, pp. 673–682, 1999.
- [3] L. Schols, J. Haan, O. Riess, G. Amoiridis, and H. Przuntek, "Sleep disturbance in spinocerebellar ataxias: is the SCA3

- mutation a cause of restless legs syndrome?" *Neurology*, vol. 51, no. 6, pp. 1603–1607, 1998.
- [4] J. J. Higgins, L. E. Nee, O. Vasconcelos et al., "Mutations in American families with spinocerebellar ataxia (SCA) type 3: SCA3 is allelic to Machado-Joseph disease," *Neurology*, vol. 46, no. 1, pp. 208–213, 1996.
- [5] B. Tang, C. Liu, L. Shen et al., "Frequency of SCA1, SCA2, SCA3/MJD, SCA6, SCA7, and DRPLA CAG trinucleotide repeat expansion in patients with hereditary spinocerebellar ataxia from Chinese kindreds," *Archives of Neurology*, vol. 57, no. 4, pp. 540–544, 2000.
- [6] T. M. Zawacki, J. Grace, J. H. Friedman, and L. Sudarsky, "Executive and emotional dysfunction in Machado-Joseph disease," *Movement Disorders*, vol. 17, no. 5, pp. 1004–1010, 2002.
- [7] R. Y. Lo, K. P. Figueroa, S. M. Pulst et al., "Depression and clinical progression in spinocerebellar ataxias," *Parkinsonism & Related Disorders*, vol. 22, pp. 87–92, 2016.
- [8] M.-T. Lin, J.-S. Yang, P.-P. Chen et al., "Bidirectional connections between depression and ataxia severity in spinocerebellar ataxia type 3 patients," *European Neurology*, vol. 79, no. 5-6, pp. 266–271, 2018.
- [9] L. Hao and D. Q. Naiman, *Quantile Regression*, pp. 40–52, SAGE Publications Inc, Thousand Oaks, CA, USA, 2007.