

Original Article

Selection of Appropriate Model for Quality of Life Data of Strabismus Patients Despite Censorship

Zahra Ghassemi, Mehdi Yaseri* , Mostafa Hosseini*

School of Public Health, Department of Epidemiology and Biostatistics, Tehran University of Medical Sciences, Tehran, Iran.

ARTICLE INFO

ABSTRACT

Received 14.01.2020
 Revised 11.03.2020
 Accepted 05.03.2020
 Published 20.09.2020

Key words:

Multiple linear regression;
 The classic tobit model;
 CLAD model;
 Ceiling effect;
 Floor effect;
 The quality of life

Introduction: Previous studies on the quality of life of strabismus patients have not examined the existence of censoring to express the relation between the response variable and its predictors.

Methods & Materials: The information used in this study is a conducted cross-sectional study in 2012. The sample size is 90 children in the age range (4-18) years and with congenital strabismus. We used the RAND Health Insurance Study questionnaire with ten subscales to evaluate the quality of life, which was increased to 11 dimensions by adding some items related to eye alignment concerns introduced by Archer et al. The demographic profile is also recorded by 13 other questions. We have expressed the relationship between the independent and response variables in each of the 11 dimensions of the questionnaire and the overall quality of life score by fitting the multiple linear regression model. Then we fitted the two models of classic Tobit and CLAD, which are for censoring, to all dimensions of the questionnaire.

Results: We showed that in fitting the models to the overall quality of life scale variable, the best model is the multiple linear regression. Because the response variable was normal, and there was no censoring (ceiling and floor effect). However, in the depression subscale, due to the high censoring (28.89% of the ceiling effect) and the almost normal distribution of the response variable (p -value of skewness < 0.05), the appropriate model according to the criteria is the classic Tobit (AIC = 546.33). That is, the classic Tobit model is the best alternative to the multiple linear regression model in the presence of censoring. But these conditions did not exist in all variables. In the subscale, there was a severe censoring performance constraint (67.78% of the ceiling effect). When censoring is high, the distribution of the response variable becomes very skewed, and the distribution of response variables deviates drastically from normal. The distribution of the performance constraint variable was very skewed (p -value < 0.001). Here the RMSE standard scale for the classic Tobit model was 28.74, which is much higher than the standard scale for the multiple linear regression model (14.23). The best model for the high censoring was CLAD.

Conclusion: To use the appropriate statistical method in the analysis, one must look at how the response variable is distributed. The multiple linear regression model is very widely used, but in the presence of censoring, the use of this model gives skewed results. In this case, the classic Tobit model and its derived model, CLAD, are replaced. The nonparametric CLAD model calculates accurate estimates with minimum defaults and censoring.

Introduction

The quality of life is a broad concept of health that was defined by the World Health Organization (WHO). According to it, health is complete physical, mental, and social well-

being and not just not being sick or disabled (1).

Since the quality of life strongly associates with physical limitations due to the disease (2), strabismus, as a common disease in

*. Corresponding Authors' Email: mhosseini110@yahoo.com & myaseri@tums.ac.ir

children, through impaired vision and abnormal appearance, causes the lack of self-confidence and general disability in the patient, and it affects the quality of life of the children and families and their lifestyles negatively (3, 4, 5, 6, 7). So, treating strabismus can have a positive effect on the quality of the children and families' life. One way to treat strabismus is operation. Operation can have a significant positive effect on quality of life indicators by affecting the psychosocial parameters (8). Although the evaluation of field view is considered as an indicator of success in the treatment of strabismus, because it does not reflect patients' feelings and perceptions of their disease, these evaluations alone cannot be sufficient (9). These evaluations are not enough to decide on a person's health, and it is necessary to examine the patient's perception of his or her health level through disease-related questionnaires (10).

In the data from questionnaires, the highest score (100) indicates the high quality of life, and the lowest score is zero. There are many situations in experimental research in which it is not possible to view continuous data of the dependent and independent variables for the whole study population. So, it leads to the distribution of the observations of such data at a denser point. Consequently, we see a lot of data accumulation around the high or low limits, or both. It happens when the research instruments are unable to display values more or less than the defined limits. In this case, the data distribution is out of the normal state and is skewed (11). This large accumulation around the high and low limits is called the ceiling and floor effect (12).

Despite these effects, the response variable does not have a normal distribution and sometimes has significant skewness. In this case, using ordinary models such as the linear

model and ordinary least squares method leads to the poorly estimation of variables, underestimating the parameters and coefficients of the real value, and generally the biased and inconsistent estimates of the parameters of the model (11, 13).

In statistical methods for model selection, one should consider the characteristics of the dependent variable and the factors affecting it, and accordingly select the appropriate model for the research. One of these characteristics is the limited variance of the dependent variable (11).

Since the accurate estimation of the relationship between the quality of life and its predictors is a big issue in health research, we should not ignore the analysis and the existence of these effects. One approach to solve this problem is to consider the effects, censoring. Because censoring occurs when it is not possible to view continuous data from the dependent variable for the entire sample under study, and all values within a defined range of the dependent variable become a fixed value (e.g., zero) (11). After censoring these effects, we use the methods that consider these censoring and is the main topic of this study to analyze the data.

Various techniques have been proposed to address this limitation in the data (limited dependent variable). One of these techniques is the Tobit model, also called the censored regression model (11). This model is an efficient method to estimate coefficients and the relationship between variables in the presence of censoring (14).

In various studies where the data collection instrument was a questionnaire, and there was censoring (ceiling or floor effect) in the data, a comparison between the performance of the Tobit model and other linear models has made. Their results summarize that this

model works very strongly in the presence of censoring (15, 16, 17).

In this study, we compare the Tobit model and one of its derivative models, namely CLAD, with the multiple linear regression model in the presence of censoring and without it.

Methods

The information used in this study is from a cross-sectional study that was conducted in 2012 in three referral university hospitals in Tehran, named Labafi Nejad Medical Center, Imam Hossein Ophthalmology Hospital, and Farabi Hospital. Patients went to these centers for operation by a pediatric ophthalmologist. The sample size is 90 children in the age range (4-18) years and with congenital strabismus.

All patients underwent ophthalmologic examinations, including preoperative visual acuity assessment and ocular alignment before and after the operation. We used the RAND Health Insurance Study questionnaire to evaluate the quality of life and added some items related to eye alignment concerns introduced by Archer et al. The demographic profile is also recorded by 13 other questions. The minimum and maximum scores of each scale of the questionnaire are 0 and 100. 0 and 100 indicate the worst and highest quality of life, respectively. The questionnaire includes 11 subscales of Function limitation, Anxiety, Depression, Positive well-being, Resistance/susceptibility, Prior health, Satisfaction with development, Eye alignment concerns, and Parent-child closeness and the overall quality of life score is the sum of these 11 subscales. Other details on how to collect data are in the Article (18). Among the demographic variables, we removed most or all of those who answered similarly. We considered the preoperative

scores in all subscales as the independent variables, and the dependent variables were the postoperative scores.

Tobit Model:

The Tobit model was introduced in 1958 by James Tobin. He called this model the model of limited dependent variables. Economists know this model and its variants as the Tobit model, a term coined by Goldberger (19). The Tobit model originates from the probit (probability unit) analysis and multiple linear regression. The advantage of this model is the use of all the information that probit (logit) or OLS models use separately.

The general form of Tobit is:

$$\begin{aligned} Y_i &= \beta'X_i + \varepsilon_i & \text{if } \beta'X_i + \varepsilon_i > 0 \\ Y_i &= 0 & \text{if } \beta'X_i + \varepsilon_i = 0 \end{aligned}$$

$i = 1, 2, \dots, N$

In which N is the number of observations, Y_i is the dependent variable and X_i is the vector of the independent variables, β is the vector of the estimated parameters and ε_i is the error which is independent and normal with zero mean and constant variance. The maximum likelihood estimation method is used to estimate the parameters. The maximum likelihood estimation function for this model has only one maximum, and its estimates are consistent and asymptotically (14).

When the instrument is a questionnaire, there is the possibility of censoring at the upper and lower limits. The model for the case which the dependent variable has upper and lower limits is:

$$y_i^* = \beta'x_i + u_i$$

In which y_i^* is an unobservable random variable if the observed values are indicated by y_i ;

$$\begin{cases} Y_i = L_1 & \text{if } y_i^* \leq L_1 \\ Y_i = y_i^* & \text{if } L_1 < y_i^* < L_2 \\ Y_i = L_2 & \text{if } y_i^* \geq L_2 \end{cases}$$

In which L_1 and L_2 are the upper and lower limits of the dependent variable, respectively. The maximum likelihood estimation function is:

$$L(\beta, \sigma_u^2 | y_i, x_i, L_1, L_2) = \prod_{y_i=L_1}^{\dots} \Phi(L_{1i} - \beta'x_i/\sigma_u) \times \prod_{y_i=y_i^*} \sigma_u^{-1} \Phi(y_i - \beta'x_i/\sigma_u) \times \prod_{y_i=L_2} [1 - \Phi(L_{2i} - \beta'x_i/\sigma_u)]$$

In the above models L_1 and L_2 are constant values that indicate the censoring points. If there is no censoring from the left, then we can set $L_1 = \infty$, and if there is no censoring from the right, then $L_2 = \infty$. The most common Tobit model is the case in which $L_1 = 0$ and $L_2 = \infty$ (20).

This model and the models derived from it have been used in econometrics more than other sciences for the cases which have limitations in measurement. In medical sciences, the models eliminate the shortcomings of measurement instruments. The results of previous studies (17) show that even if one independent variable is exposed to the ceiling (floor) effect, using linear regression methods result in a significant error in estimating regression coefficients. The solution to this problem is the use of the Tobit model and its derived models (21). Because selecting the appropriate statistical method for data analysis depends not only on the distributive assumptions but also on the characteristics such as the sequential or continuous response variable and whether this variable is limited to a bounded distance. As a result, the Tobit regression is a suitable model assuming that it is normal when there is a ceiling or floor effect (22).

However, usually, the data from the questionnaires in health assessments (due to the ceiling and floor effects) are not normally distributed, so the conditions of using the Tobit method are not met and the CLAD model, derived from Tobit, is a better alternative (23). Indeed, the heterogeneity or abnormality of the response variable leads to a drastic error in the Tobit estimator, so we need a more robust method than the Tobit and linear regression to analyze the data (25, 24). The CLAD model obtains robust estimators in the presence of censoring if it is abnormal and heterogeneous (26).

CLAD Model:

Many scientists have developed the Tobit model. One of them was Powell. He introduced the CLAD model as an alternative to the Tobit model if the error distribution was abnormal (27).

In the classic Tobit model, the MLE method is used to estimate the coefficients β in the case that ϵ_i is independent of X_i and has a normal distribution ($(\sigma^2$ and $o) N$). These are two important hypotheses when using the classic Tobit model. Goldberger (1980) and Arabmazar and Schmidt (1982) indicated that if the dependent variable is limited and the two hypotheses are not valid, even the estimators based on the standard linear regression model are more suitable for an extensive class of residual distributions. Therefore, the estimators of the classic Tobit model are very sensitive to violating these two hypotheses.

In 1984, Powell (27) showed that if the ϵ_i is nonparametric, the β estimator is not compatible. He introduced $\beta' X_i$ as a conditional median Y_i , that is:

$$Med(y_i^* | X_i) = X_i' \beta$$

He also demonstrated:

$$Med(y_i|X_i) = \max (Med (y_i^*|X_i), 0) = \max (X_i'\beta , 0)$$

Thus, the conditional median is a nonlinear function of $X_i'\beta$. This relationship indicates that the censored observations follow the nonlinear median regression model.

$$y_i = \max(X_i'\beta, 0) + \varepsilon_i$$

$$Med(\varepsilon_i|X_i) = 0$$

The appropriate method to estimate the conditional median is the least absolute deviations (LAD). Powell proposed the following relation:

$$S_n(\beta) = \sum_{i=1}^n |y_i - \max(X_i'\beta, 0)|$$

And presented the following relation as its equivalent:

$$S_n(\beta) = \sum_{i=1}^n 1(X_i'\beta > 0) |y_i - X_i'\beta|$$

The $\hat{\beta}$ estimator that minimizes the $S_n(\beta)$ function is called the CLAD estimator.

Powell also displayed its asymptotic distribution as follows:

$$\sqrt{n} (\hat{\beta} - \beta) \rightarrow d \rightarrow N(0, V)$$

Pullenayegum (28) introduces this model based on the assumption that the median is more resistant than the average in contrast with the ceiling and floor effects. The coefficients in this model are estimated so that they minimize the sum of absolute deviations from the regression line. That is: For the floor effect:

$$\sum_{i=1}^n |y_i - \max(X_i'\beta, 0)|$$

And for the ceiling effect:

$$\sum_{i=1}^n |y_i - \min(1, X_i'\beta)|$$

Previous studies emphasize that the validity of any statistical method depends on the error distribution so that the CLAD method is considered a suitable alternative to the classic Tobit model with the least hypothesis when a censoring occurs drastically (30, 29). Because this model is based on the median regression, the error distribution is very drastic against changes and finally provides us compatible estimators (31).

In this paper, we have calculated the relationship between the independent variables and the response variables in each subscale using three models of the multiple linear regression, the classic Tobit, and the CLAD, and compared these models with the considered criteria and with the presence and absence of censoring. We carried out the data analysis with the Stata software version 14 and excel 2013.

Results

According to Table 1, the subscales of function limitation, anxiety, depression, social interaction, general health perceptions, resistance/susceptibility, prior health, and eye alignment concerns did not have a normal distribution (p<0.05).

Table 1. Table for Skewness/Kurtosis tests for Normality of different dimensions of the questionnaire

QOL dimensions	Number of observations	Skewness	Kurtosis	P
quality of life	84	0.23	0.92	0.47
Function limitation	90	P<0.001	P<0.001	P<0.001
Anxiety	90	0.002	0.429	0.013
Depression	90	0.015	0.214	0.033

Positive well-being	90	0.059	0.754	0.152
Social relations	90	0.014	0.523	0.048
General health perceptions	89	P<0.001	P<0.001	P<0.001
Resistance/susceptibility prior health	90	P<0.001	0.047	P<0.001
Satisfaction with development	84	P<0.001	0.006	P<0.001
Eye alignment concerns	89	0.196	0.168	0.156
Parent-child closeness	90	P<0.001	0.094	0.003
Parent-child closeness	90	0.193	0.221	0.193

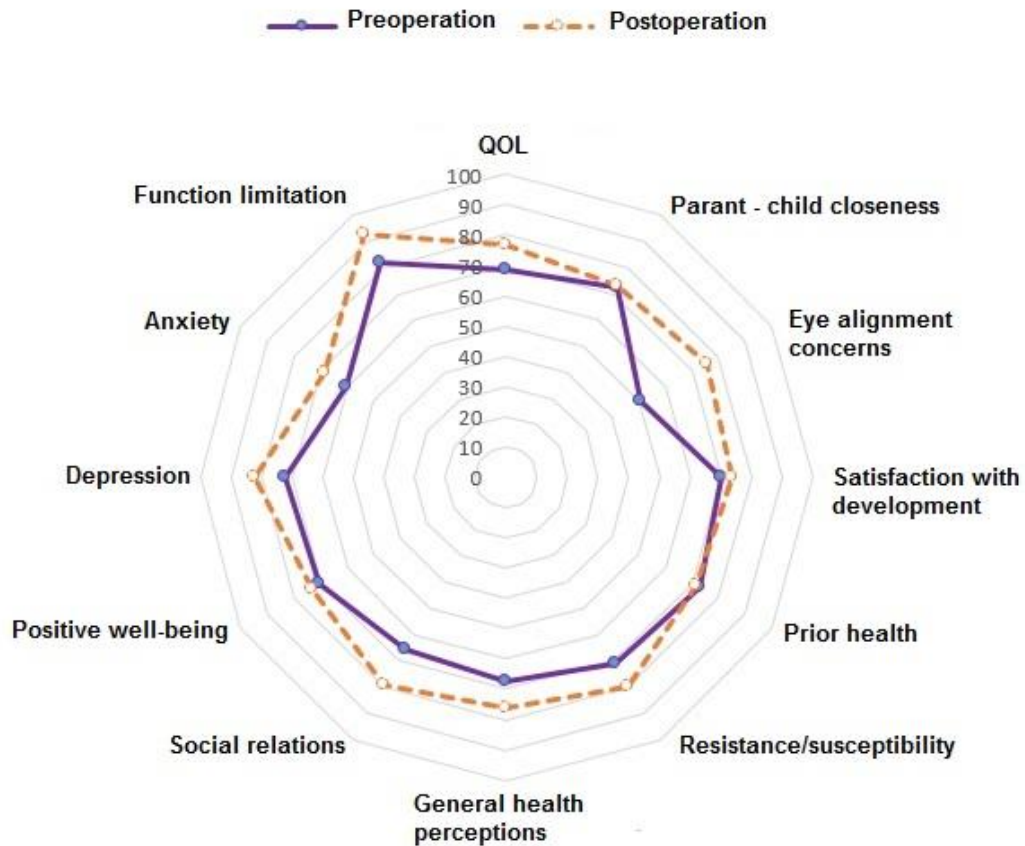


Figure 1. Differences in pre- and postoperative dim

Therefore, we used the Wilcoxon test to study the differences between the before and after scores of all variables. The results of this test in Table 2 show a significant difference between the preoperative and postoperative overall quality of life score ($p < 0.05$). This difference was observed in all subscales as

well except for prior health and parent-child closeness ($p < 0.05$).

Figure 1 confirms the Wilcoxon test results in the image. In this diagram, we have displayed the comparison between pre- and post-scores of the overall quality of life and its dimensions, in which the most differences are

related to eye alignment concerns, while the least ones are for parent-child closeness.

Table 2. Wilcoxon test on pre- and post-operative scores for questionnaire dimensions

QOL	Pre-operation		Post-operation		P
	Mean		Mean		
QOL dimensions	Mean		Mean		
quality of life	68.66	9.12	76.70	8.59	0.00
Function limitation	82.15	20.92	92.36	16.78	0.00
Anxiety	60.28	19.19	68.61	18.15	0.00
Depression	72.36	17.72	82.32	16.42	0.00
Positive well-being	70.56	15.96	73.33	14.68	0.02
Social relations	65.39	13.99	79.43	11.52	0.00
General health perceptions	67.36	18.89	76.40	16.47	0.00
Resistance/susceptibility	71.02	14.58	79.72	13.39	0.00
prior health	72.78	15.29	71.73	15.89	0.29
Satisfaction with development	70.07	14.97	73.81	16.07	0.00
Eye alignment concerns	50.84	19.01	75.44	15.89	0.00
Parent-child closeness	72.50	17.99	72.92	15.82	0.81

After studying the difference between the pre- and post-scores, we have written the amount of accumulation in the upper and lower limits as a percentage in Table 3 to check the presence or absence of the ceiling and floor effects. The percentage of the accumulation in the lower limits of all subscales is zero but general health

perception, while the maximum amount in the upper limit of the function limitation subscale is 67.78%, followed by the subscales of depression (28.89%), general health perceptions (16.85%), parent-child closeness (13.33%), and satisfaction with development (12.36%), respectively.

Table 3. The amount of accumulation in the upper and lower limits as a percentage of the subscale scores Post-operation

QOL dimensions	Lower limit(%)	Upper limit(%)
quality of life	0	67.78
Function limitation	0	1.11
Anxiety	0	28.89
Depression	0	5.56
Positive well-being	0	4.44
Social relations	1.12	16.85
General health perceptions	0	7.78
Resistance/susceptibility	0	12.36
prior health	0	1.11
Satisfaction with development	0	13.33
Eye alignment concerns	0	9.52
Parent-child closeness	0	0

The accumulation at the upper limit for the other subscales is less than 10%. We did not find an accurate measure of the ceiling and floor effect. Here we have defined the accumulation of at least 20% of the data at the upper and lower limits as the reason for the ceiling and floor effect (32).

To study the advantages or disadvantages of the models than each other, we fitted all three multiple linear regression models, the classic Tobit, and the CLAD to the overall quality of life score and its 11 subscales. The response

variables of the postoperative scores, the independent variables of the preoperative scores, demographic variables, and other variables affected each dimension of the questionnaire. After fitting the models, we have listed the Pseudo-R², RMSE, MAE, and AIC criteria in Table 4. As we know, the model with the highest Pseudo-R² value and the lowest value for RMSE, MAE, and AIC criteria is the superior model.

Table 4. Demonstrate four criteria for all three models in questionnaire dimensions

	QOL dimensions	QOL	Function limitation	Anxiety	Depression	Positive well-being	Social relations	General health perceptions	Resistance/susceptibility	prior health	Satisfaction with development	Eye alignment concerns	Parent-child closeness
Multiple linear	MAE	5.31	8.56	12.02	9.39	7.87	8.64	10.07	9.04	9.43	8.48	11.82	7.86
	AIC	484.63	717.05	735.38	653.58	542.68	679.69	556.86	522.47	558.39	687.55	572.25	501.10
	RMSE	6.54	14.23	15.17	12.55	10.20	10.64	14.25	11.82	13.97	11.51	15.27	9.74
	Pseudo-R²	0.41	0.29	0.29	0.42	0.58	0.15	0.25	0.33	0.29	0.483	0.06	0.66
TOBIT	MAE	5.31	25.39	12.00	10.26	7.89	8.66	10.67	9.08	9.74	8.69	11.82	8.07
	AIC	486.63	329.42	732.97	546.33	527.23	666.86	506.55	513.53	530.47	646.13	570.10	461.86
	RMSE	6.54	28.74	15.18	13.53	10.27	10.69	14.72	11.93	14.20	11.67	15.28	10.09
	Pseudo-R²	0.41	0.29	0.29	0.42	0.58	0.15	0.26	0.33	0.29	0.481	0.06	0.66
CLAD	MAE	5.89	-	12.33	-	9.29	9.43	-	-	-	-	13.07	-
	AIC	-	-	-	-	-	-	-	-	-	-	-	-
	RMSE	7.29	-	14.42	-	12.56	11.78	-	-	-	-	17.24	-
	Pseudo-R²	0.36	0.07	0.29	0.41	0.42	0.07	0.03	0.21	-	0.38	0.07	-

Table 4 indicates that the software has not calculated the MAE criterion value for the CLAD model in the subscales of function limitation, depression, general health perceptions, resistance/ susceptibility, prior health, satisfaction with development, and parent-child closeness. If we want to decide according to the MAE criterion, the comparison is between two models of multiple linear regression and classic Tobit. In these mentioned subscales, the lowest MAE criterion score is for the multiple linear regression model.

In the other subscales that the software has calculated the MAE score for its all three models, its score in the subscales of eye alignment concern and overall quality of life for both the multiple linear regression models and the classic Tobit has been the same and less than the CLAD model. For the anxiety subscale, the lowest score is for the classic Tobit model.

If we want to determine the appropriate model according to the AIC criterion, since the software does not provide the output of this criterion for the CLAD model, the comparison of this criterion is only between the two models of multiple linear regression and classic Tobit. According to the inserted tables for each subscale, when our response variable is the overall quality of life score, the lowest value is for the multiple linear regression model, and in the other subscales, the lowest AIC score is for the classic Tobit model.

The next comparison criterion is the RMSE. Like MAE, the score of this criterion has not been calculated in some subscales for the CLAD model. Therefore, in the subscales of function limitation, depression, general health perceptions, resistance/susceptibility, prior health, satisfaction with development, and parent-child closeness are comparisons

between two models of multiple linear regression and classic Tobit. The score of this criterion is the lowest in all these subscales for the multiple linear regression model. Multiple linear regression also has the lowest score in the subscales such as anxiety, positive well-being, social relations, eye alignment concerns, and overall quality of life that RMSE has been calculated for all three models. Anyhow, the RMSE criterion score for the multiple linear regression model is the lowest.

The last criterion for comparing models is Pseudo- R^2 . This criterion is calculated if the software can fit the model. We could not fit the CLAD model when our response variables were prior health and parent-child closeness. So, in these two subscales, we can only compare the other two models. According to Table 3 in the subscale of prior health, the criterion score for both the multiple linear regression models and the classic Tobit is the same. And for the parent-child closeness subscale, the criterion score for both of these models are equal. Therefore, according to this criterion, the superior model is not recognizable for these two subscales. In the rest of the subscales, the Pseudo- R^2 value for the overall quality of life score, function limitation, positive well-being, social relations, resistance/susceptibility, and satisfaction with development has been the same in both multiple linear regression models and classic Tobit and more than the CLAD model. This case is also true for the general health perceptions subscale. The small difference between the scores of this criterion in the two models of multiple linear regression and classic Tobit can be considered as the result of rounding error, and its value the same as both models and more than the CLAD model. In the Anxiety subscale, the criterion score is the same for

all three models. For the depression subscale, its score is almost equal in all three models. Therefore, the criterion does not recognize the superior model in two subscales (anxiety and depression). The only subscale that had the highest criterion score for the CLAD model with very little difference was eye alignment concerns. That is, according to the Pseudo- R^2 criterion, when the response variable is eye alignment concerns, the CLAD model is the appropriate one.

Discussion

In this study, we have fitted the three models of classic Tobit, CLAD, and multiple linear regression to the overall quality of life score and its 11 subscales to see the appropriateness of the models in the different distributions of the response variables, and detect the superior model to indicate the relationship between the response variables and its predictors.

We have selected the independent variables among the demographic variables. Besides, we have considered three clinical evaluation criteria as an independent variable. These criteria are visual impairment, the amount of ocular realignment, and residual deviation. In case that it affects the subscales, we will add it to the model as an influential variable.

In this study, a large improvement was observed in the scores of subscales after the surgery. The greatest improvement is related to the eye alignment concern subscale. The results of this section of the study are similar to the Ziaee et al.'s one (18).

The difference between this study and the study of Ziaee et al. (18) is in the significance of clinical variables in the dimensions of the questionnaire. In this study, the residual deviation is a significant variable in the well-being subscales ($p = 0.002$) and resistance /susceptibility ($p = 0.056$). But in

the Ziaee et al.'s study (18), this variable was effective only in the well-being subscale. Ziaee et al. have reported the effect of the eye alignment variable in their study (18) on two subscales of Satisfaction with development and eye alignment concerns positively. In our study, we observed the effect of this variable in more dimensions of the questionnaire, including general health perceptions, resistance/susceptibility, prior health, eye alignment concerns, positive well-being, and parent-child closeness. This variable was very significant among these dimensions in the subscales of well-being ($p = 0.002$), prior health ($p = 0.002$), and parent-child closeness ($p = 0.007$).

The preoperative visual impairment variable in this study has been described as another preoperative treatment, affected the dimensions of depression, resistance/susceptibility, eye alignment concerns, parent-child closeness, and overall quality of life score. The greatest effect was on the dimensions of depression ($p = 0.050$) and parent-child closeness ($p = 0.004$). These findings contradict the results of the previous study on this data (18).

Our findings in fitting the models to data are as follows. In fitting the models to the overall quality of life score variable, we can see that the best model is the multiple linear regression. Due to Table 1, it is normal by default, and according to table 3, we do not see the ceiling and floor effect in this scale.

In the depression subscale, due to the censoring in the upper limit (28.89% of the ceiling effect) and the nearly normal distribution of the response variable (p -value of skewness < 0.05), the appropriate model according to the criteria is the classic Tobit. That is, the classic Tobit model is the best alternative to the multiple linear regression model in the presence of censoring (33).

But it is usually very uncommon that the distribution of the response variable gets normal and there is censoring. When there is censoring, the distribution of the response variable is skewed, and the higher the censoring, the more skewed it is. Therefore, the distribution of the response variable is more deviant than the normal. In the function limitation subscale, this case is seen. In the case that censoring is severe, the classic Tobit model performs even worse than multiple linear regressions, and a robust method of both should be used (27, 25). This powerful method is the CLAD model (31,29,16,15). This model is not sensitive to deviations from normal (23, 24, 26). This method has estimated the value of the coefficients more than the other two models in the function subscale, which has big censoring. This result is completely consistent with the results of Austin's studies (15, 17).

The comparison results, according to RMSE and MAE criteria, were the same except for the anxiety subscale. Based on this research, it is not necessary to use both criteria simultaneously to compare the models.

Accordant with the Pseudo- R^2 criterion, the superior model cannot be identified either, because only in the subscale of eye alignment concerns, they had slightly different values. Other subscales had equal values for at least two models.

Another suggestion of this research is about the AIC criterion. AIC is a widely used criterion for comparing models in statistical methods. Since this criterion is not calculated for the CLAD model (34), when one of the models is CLAD, it is better not to use the AIC comparison criterion.

If we want to compare the models in terms of interpretability, the multiple linear regression model is the simplest one. Because in the

classic Tobit and CLAD models, we discuss the data which we do not observe (28, 35).

Another limitation to this study was the definition of the existence of the ceiling and floor effect criterion. Here we defined the acquisition of at least 20% of the lower and upper limit scores based on the existence of these effects (32). Accordingly, only two subscales of function limitation (68.78%) and depression (28.89%) have a ceiling effect. Although there is no exact criterion to confirm the existence of these effects, according to the number of samples, this criterion can be considered lower. In this case, more subscales with these effects are introduced.

Another case that we did not encounter in this study was the presence of both ceiling and floor effects simultaneously in the same data. In this respect, we cannot use the CLAD model in this software. Because we can only fit this model for one effect (ceiling or floor).

Conclusion

The multiple linear regression model is widely used to express the relationship between the response variable and its predictors, but in the presence of censoring, using this model gives skewed results. In this matter, the classic Tobit model is a good alternative. The classic Tobit model is also useful as long as the distribution of the response variable is normal, and the data is censored in upper and lower limits. But if the response variable distribution has a critical deviation from the normal, actually highly censoring, the classic Tobit model can no longer be useful, and gives us more deviant results from the multiple linear regression. In this regard, the nonparametric CLAD model calculates accurate estimates with minimum defaults and censoring.

Conflict of Interests

There are no conflicts of interest.

References

1. Park JE PK. Preventive and social medicine 17 ed. Jabalpure: M/s Banarsidas Bhanot; 2004.
2. Ghaffari Gh, Karimi A, Nozari H. Trend study of quality of life in Iran. Quarterly Journal of Social Studies and Research. 1391; 1(3):07-134.
3. T LT, N S, A M, SH N. Stress coping process in parents of children with seizure. Journal of Tehran Nursing and Midwifery Faculty (HAYAT). 2005;12(2):63-71.
4. Hatt SR, Leske DA, Kirgis PA, Bradley EA, Holmes JM. The effects of strabismus on quality of life in adults. American journal of ophthalmology. 2007;144(5):643-7.
5. Satterfield D, Keltner JL, Morrison TL. Psychosocial aspects of strabismus study. Archives of ophthalmology (Chicago, Ill : 1960). 1993;111(8):1100-5.
6. Archer SM, Musch DC, Wren PA, Guire KE, Del Monte MA. Social and emotional impact of strabismus surgery on quality of life in children. Journal of AAPOS : the official publication of the American Association for Pediatric Ophthalmology and Strabismus. 2005;9(2):148-51.
7. Lilakova D, Hejzmanova D, Novakova D. [Effect of strabismus on the quality of life in adults). Ceska a slovenska oftalmologie : casopis Ceske oftalmologicke spolecnosti a Slovenske oftalmologicke spolecnosti. 2003;59(3):184-7.
8. Nelson BA, Gunton KB, Lasker JN, Nelson LB, Drohan LA. The psychosocial aspects of strabismus in teenagers and adults and the impact of surgical correction. Journal of AAPOS : the official publication of the American Association for Pediatric

Ophthalmology and Strabismus. 2008;12(1):72-6.e1.

9. Clancy CM, Eisenberg JM. Outcomes research: measuring the end results of health care. Science (New York, NY). 1998;282(5387):245-6.

10. Sartorius N. The meanings of health and its promotion. Croatian. medical journal.2006;47 (4): 662-4.

11. Guo S, Fraser MW. PROPENSITY SCORE ANALYSIS(Statistical Methods and Applications). 2, editor: Thousand Oaks: Sage Publications; 2014.

12. Everitt, B. S. (2002). The Cambridge dictionary of statistics (2nd ed.). Cambridge, UK: CUP.

13. Berk KN, Lachenbruch PA. Repeated measures with zeros. Statistical Methods in Medical Research 2002; 11: 303-16

14. Tobin J. Estimation of Relationships for Limited Dependent Variables. Econometrica. 1958;26(1):24-36.

15. Austin PC. A comparison of methods for analyzing health-related quality-of-life measures. Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research. 2002;5(4):329-37.

16. GS M. Limited-Dependent and Qualitative Variables in Econometrics. New York : Cambridge University Press 1983.

17. Austin PC, Hoch JS. Estimating linear regression models in the presence of a censored independent variable. Statistics in medicine. 2004;23(3):411-29.

18. Ziaei H, Katibeh M, Mohammadi S, Mirzaei M, Moein HR, Kheiri B, et al. The Impact of congenital strabismus surgery on quality of life in children. Journal of ophthalmic & vision research. 2016;11(2):188-92.

19. AMEMIYA T. TOBIT MODELS : A SURVEY. *Journal of Econometrics* 24. 1984:61-3.
20. Baltagi B H. *Econometric Analysis of Panel Data*. Third edition. 2003; 652.
21. Rahmani E, Kaaouachi A. Detection of a Tobit Model. *Applied Mathematical Sciences*. 2015;9(39):1911-7.
22. Arostegui I, Nunez-Anton V, Quintana JM. Statistical approaches to analyse patient-reported outcomes as response variables: an application to health-related quality of life. *Statistical methods in medical research*. 2012;21(2):189-214.
23. Sullivan PW, Lawrence WF, Ghushchyan V. A national catalog of preference-based scores for chronic conditions in the United States. *Medical care*. 2005;43(7):736-49.
24. WILHELM MO. PRACTITIONERS' CORNER. Practical Considerations for Choosing Between Tobit and SCLS or CLAD Estimators for an Application Censored Regression Models with to Charitable Giving*. *OXFORD BULLETIN OF ECONOMICS AND STATISTICS*. 2008;70(4).
25. Stewart J. Tobit or Not Tobit? US Bureau of Labor Statistics. 2009.
26. Wocken M, Kneib T. Tobit regression to estimate impact of EU market intervention in dairy sector. Dublin-123rd EAAE seminar. 2012:23-4.
27. L.POWELL J. LEAST ABSOLUTE DEVIATIONS ESTIMATION FOR THE CENSORED REGRESSION MODEL. *Journal of Econometrics* 25. 1984:303-25.
28. Pullenayegum EM, Tarride JE, Xie F, Goeree R, Gerstein HC, O'Reilly D. Analysis of health utility data when some subjects attain the upper bound of 1: are Tobit and CLAD models appropriate? *Value in health : the journal of the International Society for Pharmacoeconomics and Outcomes Research*. 2010;13(4):487-94.
29. Chay KY, Powell JL. Semiparametric Censored Regression Models. *Journal of Economic Perspectives*. 2001;15(4):29-42.
30. Raidani, Pertiwi LS, Wulandari DY, Zuhri. Tobit and Interval Censored Regression Model. *Global Journal of Pure and Applied Mathematics*. 2016;12(1):981-94.
31. Clarke P, Gray A, Holman R. Estimating utility values for health states of type 2 diabetic patients using the EQ-5D (UKPDS 62). *Medical decision making : an international journal of the Society for Medical Decision Making*. 2002;22(4):340-9.
32. Wang Z, Ren H, Frey R, Liu Y, Raphael D, Bian W, et al. Comparison of the Adult Strabismus Quality of Life Questionnaire (AS-20) with the Amblyopia and Strabismus Questionnaire (ASQE) among adults with strabismus who seek medical care in China. *BMC ophthalmology*. 2014;14:139.
33. Arostegui I, Nunez-Anton V, Quintana JM. Statistical approaches to analyse patient-reported outcomes as response variables: an application to health-related quality of life. *Statistical methods in medical research*. 2012;21(2):189-214.
34. Crott R. Direct Mapping of the QLQ-C30 to EQ-5D Preferences: A Comparison of Regression Methods. *Pharmacoeconomics*. 2017;2:165-77.
35. Fernando R. Logit, Probit and Tobit : Models for Categorical and Limited Dependent Variables. *PLCS/RDC Statistics and Data Series at Western* 2011.