

Adjusting the Oral Health Related Quality of Life Measure (Using Ohip-14) for Floor and Ceiling Effects

Andiappan M¹, Hughes FJ², Dunne S³, Gao W⁴, Donaldson ANA⁵

ABSTRACT

Introduction

The influence of floor (lowest) and ceiling (highest) effects on the outcome measure is of serious concern particularly when the outcome is measured using Likert scales. Conventional regression methods yield biased results and hence tobit regression is to be used to adjust for these effects. This paper is an attempt to use tobit regression in finding the predictors of oral health related quality of life after adjusting for floor and ceiling effects.

Methods

A sample of 360 participants were asked to self assess their oral health related quality of life (OHRQoL) using Oral Health impact profile with 14 items which forms the data for this study. Apart from descriptive statistics, Ordinary Least squares regression and tobit regression were used to find the significant predictors of OHRQoL and the results of both methods were compared.

Results

The sample comprised of 41.1% men and 58.9% women. Majority of the participants (68.3%) were whites. The average item difficulty was 0.4 and the average item easiness was 0.03. The floor and ceiling values for the composite scores were 14 and 56 respectively. Age and gender were not statistically significant both in Ordinary Least Squares (OLS) regression and Tobit regression. Full time employment, student and retired have significantly lower scores in OLS but only retired had significantly lower scores in the tobit model.

Conclusion

Tobit model, after adjusting for floor and ceiling effect, gives higher values for the predictors and the OLS model underestimates the effects of predictors on OHIP scores.

KEY WORDS: Likert, OHIP, Regression, Prediction

Contact Author

Manoharan Andiappan

manoharan.1.andiappan@kcl.ac.uk

J Oral Health Comm Dent 2015;9(3)99-104

¹ M.Sc
Biostatistician and PhD Student
King's College London.

² DS, PhD
Professor of Periodontology
King's College London.

³ BDS (U Lond), LDS RCS (Eng), FDS RCS (Eng),
PhD (U Lond).
Professor of Primary Dental Care
King's College London.

⁴ PhD, M.Med, B.Med, GCAP
Senior Lecturer in Statistics and Epidemiology
King's College London.

⁵ BSc (Maths) MSc (Applied Stats) PHD Statistics
Professor of Biostatistics
King's College London.

INTRODUCTION

The majority of outcomes measuring the social and psychological impact of treatments have been defined as Likert scales, where items are scored within a finite range of values, often with the six possibilities 0 to 5, where 0 indicates no impact and 5 indicates the worst impact. This type of measurement is used in dental research to measure the patient's oral health related quality of life (OHRQoL). In particular, the Oral Health Impact profile (OHIP-14), a shortened version using only 14 of the 49 items of the original OHIP-49, has been widely used due to its good psychometric properties and reduced number of items.

This type of scale, and any derived composite (the total score of the items forming a given dimension, for example), yields measures that are censored on the right, on the left, or in both sides. In statistics, censoring occurs when the value of an outcome is only partially known. If a patient experiences severe day-to-day problems due to their dental illness, they are forced to give a score of 5 (the maximum). If the range of the score had been from 0-100 instead of 0-5, the individuals who were forced to give 5 for a particular question would have given different scores in the higher range. By adopting a shorter range (0-5), the items are right-censored, meaning that the true score of these individuals have not been measured accurately but have been curtailed at the highest value recorded.

When the scores of the participants cannot naturally exceed a certain value, an asymptote occurs at the largest value that the participants can score (1). The scores of the individuals with poor quality of life would be close to the maximum and this value would be, naturally, large enough to allow for the difference between individuals with severe dental problems to be measured. A ceiling occurs when the participants are forced to give a maximum, so called

the ceiling, which is much less than the asymptotic value. In OHIP-14 the ceiling takes place at 5. All those individuals with severe dental problems will score 5, the maximum allowed. This curtailment may be drastic and any bias introduced in this manner is termed ceiling effect. This potential bias comes about because the variability between those individuals with severe dental problems cannot be measured. Likewise, when the items are left-censored, we say that a floor has been fixed by the score and the bias introduced in this manner is termed floor effect.

In the presence of censoring, estimators of effects based on the conventional Ordinary Least Squares (OLS) models are inconsistent (in the sense that the coefficients will not necessarily approach the "true" population parameters as the sample size increases) and yield biased estimations of both the gradient (under-estimating it) and the intercept (over-estimating it) of the model. Many researchers have studied the impact of the ceiling effect. Some studies have identified artificial non linearity and under-estimation of the regression parameters (2-3). At the same time, researchers have developed new analytical techniques that can handle the ceiling effect. The use of growth-curve models have been suggested for handling ceiling effects in longitudinal data (4-5). In cross sectional data, Tobin (6) developed the Tobit regression model, which has been widely used in cross sectional studies where the measure has ceiling effects. These approaches apply to both ceiling and floor effects in a similar way.

The initial development of the Tobit model, based on econometrics data, was proposed by Tobin in 1958. According to Tobin, the linear regression model with censored data can be represented as $y^* = a + bx + \epsilon$, where y^* the latent (non observable) variable of is interest, which is to be modelled based on the observed data, and ϵ is the normally distributed error

term. The observed score is y and, $y = y^*$ if $y^* < c$ and $y = c$ if $y^* > c$. The observed score y is measured with right-censoring and y cannot exceed the censoring point c . Many researchers have attempted to handle the floor and ceiling effects in various measurements. Brown and Muthen have separately documented the benefit of using Tobit regression to address the shortcomings of Pearson correlation when analyzing censored data (7-8). Nevertheless, floor and ceiling effects in the measure of Oral Health Related Quality of Life (OHRQoL) using OHIP-14 has not been explored. In this paper we attempt to examine these effects in the measures provided by the total OHIP-14 or its domains and to evaluate their impact. As for any outcome measured in the Likert scale, it is important to assess the floor and ceiling effects in the analysis of the OHIP-14.

MATERIALS AND METHODS

Patients attending King's College Dental Hospital for various dental treatments were approached to take part in this study, with approval from the London-Bromley ethical committee (13/LO/0366). Patients who could not give consent individually or did not have sufficient command of English or were under the age of 15 were excluded. After getting a written consent to take part in this research, participants were asked to fill in a questionnaire that contained the OHIP-14 instrument and other socio-demographic and oral health related behavior questions. Although the study was longitudinal (baseline and 2 and 4 months after treatment), this paper is based only on the cross-sectional analysis of the baseline data collected for the 360 individuals enrolled.

STATISTICAL ANALYSIS

In addition to the usual descriptive summaries, multiple linear regression was used to assess the relationship between OHRQoL and demographic variables. In a conventional way, the ordinary least squares (OLS), without

considering floor or ceiling effects, was applied. In addition, the Tobit regression was used, considering any potential floor and ceiling effects. The results of these two regression models were compared using the difference in the effect of predictors and the standard errors.

CENSORING IN OHIP-14

As OHIP-14 uses Likert scales, the actual score is expected to be biased due to censoring. We considered the possibility of both censoring at the lower and the upper end yielding floor and ceiling effects, respectively. We analysed the seven dimensions of the OHIP-14 score as well as the aggregate total score of all the 14 items (a composite score).

It is usual to assume the lowest score to be the left-censoring point. However, in the case of composite scores like OHIP-14 score, there are two options. One is to take the total of the lowest scale scores of all the items (in this case the lowest score is 0 for all the 14 items and hence the total is 0), which is an extreme point. The other is to consider the sum of all the lower end scores of all the items (if the scale range is 0 to 5, maybe 0 and 1 are considered as the lower ends). As there is no clear defini-

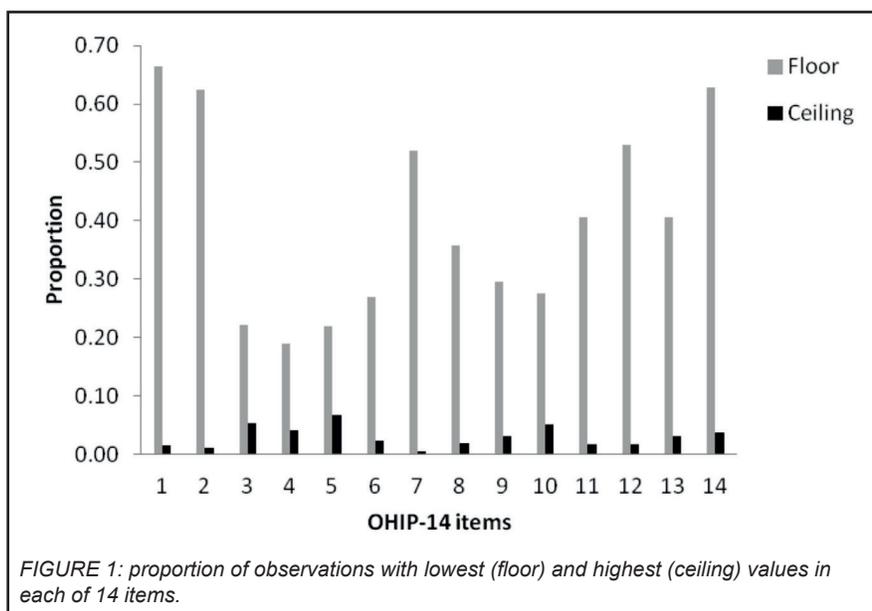
tion of lower end scores, there would be many possible options according to the range of scales. In this paper, we adopted the method suggested by Ven den Oord and Van den ark 1997 to decide on the floor and ceiling points in the case of composite data. We calculated the average item difficulty, representing the proportion of subjects whose rating is consistently biased towards the lower end of the scale (in reference to the case when a question is too difficult and most of the respondents get a low score on it). Similarly, for the ceiling point, we calculated the average item easiness, representing the proportion of subjects whose rating is consistently biased towards the upper end (in reference to the case when a question is too easy and most of the respondents get a high score on it). Tobit regression is fitted based on the calculated floor and ceiling points for the composite data.

RESULTS

Of the sample of 360 patients, 148 (41.1%) were men. The distribution across the ethnic groups was: White (68.3%), Black (16.6%), Mixed (8.4%), Asian (4.7%) and Other (1.1%). Ethnicity was missing for only 4 participants. All the participants filled in the OHIP-14 questionnaire and the oral health

knowledge and behavior questionnaire, along with their demographic details. Each item of the OHIP-14 was scaled from 0 to 5, with higher scores indicating poorer quality of life. The missing data in the OHIP-14 items were filled using EM (Expectation Maximization) algorithm. The proportion of patients who have given the lowest score (floor) and the highest score (ceiling) for each of the 14 items are depicted in Fig-1.

It is clear from the figure 1 that the floor score is seen considerably more often than the ceiling score. The amount of left censoring (floor) ranged from 18.9% to 66.4% for the 14 items, while the amount of right censoring (ceiling) ranged from 0.6 to 6.7%. The average item difficulty was 0.4 indicating that on average 40% of the patients consistently gave the lower end of the score. On the right censoring (ceiling), the average item easiness was 0.03, meaning that 3% of the patients tend to give higher side of the score. Therefore, the censoring points at lower side (floor) and higher side (ceiling) for the composite score of OHIP-14 were 14 and 56 respectively. The results of the ordinary least squares (OLS) and Tobit regressions (with a floor value of 14 and a ceiling value of 56) are given in table-1 for the total OHIP-14 and for each of the seven dimensions.



In the analysis of the Total OHIP-14 (Table 1), under both the OLS and Tobit models, age ($P=0.27$) and gender ($P=0.83$) were not found to be significant predictors of oral health related quality of life. The multiple linear OLS regression analysis showed that full-time employed, students and retired subjects had significantly lower scores (e.g. better quality of life) in relation to the unemployed group while, in contrast, under the Tobit model, only 'retired' was found to be significantly different from 'unemployed'. The higher magnitude of the regression coefficients obtained under the Tobit model indicates that the estimated effect of predictor variables on the total OHIP score is higher under the Tobit

Table 1: Analysis By Domains				
Predictors	Ordinary Linear Regression (OLs)		Tobit Regression	
	Effect (95% Ci)	P Value	Effect (95% Ci)	P Value
TOTAL OHIP-14				
Age	-0.02 (-0.15, 0.12)	0.83	-0.02 (-0.20, 0.16)	0.83
Gender	1.90 (-1.48, 5.27)	0.27	2.79 (-1.76, 7.33)	0.23
Profession:		0.06		0.06
Fulltime	-5.7 (-11.09, -0.32)	0.27	-6.2 (-13.27, 0.86)	0.09
Part Time	-5.1 (-11.64, 1.50)	0.13	-6 (-14.61, 2.71)	0.18
Student	-10.9 (-20.42, -1.27)	0.03*	-12.9 (-26.10, 0.31)	0.06
Retired	-10.2 (-17.37, -3.01)	0.01*	-14.7 (-24.6, -4.9)	0.003**
Others	-6.7 (-13.96, 0.54)	0.07	-7 (-16.60, 2.64)	0.16
FUNCTIONAL LIMITATION				
Age	0.03 (0.01, 0.04)	0.009*	0.05 (0.01, 0.08)	0.016*
Gender	-0.08 (-0.56, 0.40)	0.74	-0.84 (-1.02, 0.85)	0.86
Profession:		0.43		0.57
Fulltime	-0.40 (-1.17, 0.37)	0.31	-0.27 (-1.76, 1.22)	0.73
Part Time	-0.84 (-1.78, 0.10)	0.08	-1.49 (-3.38, 0.40)	0.12
Student	-0.26 (-1.62, 1.11)	0.71	-0.11 (-2.83, 2.62)	0.94
Retired	-0.85 (-1.87, 0.17)	0.10	-0.98 (-2.93, 0.97)	0.33
Others	-0.77 (-1.80, 0.26)	0.14	-0.96 (-3.00, 1.09)	0.36
PHYSICAL PAIN				
Age	0.01 (-0.01, 0.03)	0.30	0.01 (-0.01, 0.04)	0.30
Gender	0.22 (-0.35, 0.79)	0.45	0.21 (-0.42, 0.85)	0.51
Profession:		0.35		0.27
Fulltime	-0.23 (-1.15, 0.68)	0.62	-0.32 (-1.33, 0.69)	0.53
Part Time	-0.02 (-1.14, 1.10)	0.97	-0.003 (-1.22, 1.23)	0.99
Student	-0.61 (-2.23, 1.02)	0.46	-0.82 (-2.63, 0.99)	0.37
Retired	-1.23 (-2.45, -0.01)	0.049	-1.41 (-2.76, -0.08)	0.038*
Others	-0.39 (1.63, 0.84)	0.53	-0.46 (-1.81, 0.90)	0.51
PSYCHOLOGICAL DISCOMFORT				
Age	-0.01 (-0.04, 0.01)	0.39	-0.01 (-0.04, 0.01)	0.30
Gender	0.64 (0.02, 1.26)	0.043*	-0.79 (0.09, 1.49)	0.028*
Profession:		0.020*		0.014*
Fulltime	-0.96 (-1.95, 0.03)	0.06	-1.09 (-2.20, 0.01)	0.05
Part Time	-0.72 (-1.93, 0.49)	0.24	-0.90 (-2.25, 0.45)	0.19
Student	-2.31 (-4.07, -0.55)	0.010*	-2.79 (-4.83, -0.75)	0.007*
Retired	-1.89 (-3.21, -0.57)	0.005*	-2.25 (-3.74, -0.75)	0.003*
Others	-1.56 (-2.89, -0.23)	0.022*	-1.72 (-3.22, -0.22)	0.025*
PHYSICAL DISABILITY				
Age	0.01 (-0.02, 0.03)	0.66	0.01 (-0.02, 0.04)	0.63
Gender	0.36 (-0.22, 0.94)	0.22	0.42 (-0.31, 1.15)	0.26
Profession:		0.44		0.43
Fulltime	-0.69 (-1.62, 0.24)	0.14	-0.80 (-1.96, 0.36)	0.18
Part Time	-0.55 (-1.69, 0.58)	0.34	-0.56 (-1.97, 0.85)	0.44
Student	-1.09 (-2.74, 0.56)	0.19	-1.23 (-3.32, 0.85)	0.25
Retired	-1.23 (-2.46, 0.01)	0.05	-1.57 (-3.13, -0.02)	0.047*
Others	-0.48 (-1.73, 0.77)	0.45	-0.53 (-2.10, 1.03)	0.50
PSYCHOLOGICAL DISABILITY				
Age	-0.02 (-0.04, 0.01)	0.15	-0.02 (-0.05, 0.004)	0.09
Gender	0.36 (-0.24, 0.96)	0.24	0.04 (-0.30, -1.11)	0.26
Profession:		0.017*		0.019*
Fulltime	-1.39 (-2.34, -0.44)	0.004*	-1.57 (-2.68, -0.46)	0.006*
Part Time	-1.37 (-2.54, -0.21)	0.021*	-1.60 (-2.96, -0.25)	0.021*
Student	-2.35 (-4.04, -0.65)	0.007*	-1.76 (-4.79, -0.73)	0.008*
Retired	-1.98 (-3.25, -0.71)	0.002*	-2.30 (-3.81, -0.80)	0.003*
Others	-1.51 (-2.79, -0.22)	0.022*	-1.66 (-3.17, -0.16)	0.031*
SOCIAL DISABILITY				
Age	-0.02 (-0.04, 0.01)	0.12	-0.03 (-0.06, 0.00)	0.08
Gender	0.43 (-0.14, 1.01)	0.14	0.60 (-0.17, 1.37)	0.12
Profession:		0.027*		0.032*
Fulltime	-1.12 (-2.05, -0.20)	0.017*	-1.29 (-2.49, -0.08)	0.036*
Part Time	-0.89 (-2.02, 0.24)	0.12	-1.00 (-2.47, 0.46)	0.18
Student	-2.32 (-3.96, -0.67)	0.006*	-3.05 (-5.28, -0.82)	0.007*
Retired	-1.75 (-2.98, -0.52)	0.005*	-2.29 (-3.94, -0.64)	0.007*
Others	-1.03 (-2.28, 0.21)	0.10	-1.20 (-2.82, 0.43)	0.15
HANDICAP				
Age	-0.01 (-0.03, 0.01)	0.36	-0.02 (-0.06, 0.02)	0.27
Gender	-0.04 (-0.64, 0.55)	0.89	-0.01 (-0.94, 0.92)	0.98
Profession:		0.19		0.20
Fulltime	-0.90 (-1.85, 0.04)	0.06	-1.24 (-2.68, 0.20)	0.09
Part Time	-0.67 (1.83, 0.48)	0.25	-0.97 (-2.73, 0.80)	0.28
Student	-1.92 (-3.60, -0.23)	0.026*	-3.12 (-5.89, -0.35)	0.027*
Retired	-1.26 (-2.52, 0.00)	0.06	-1.92 (-3.90, 0.06)	0.06
Others	-0.97 (-2.24, 0.30)	0.14	-1.35 (-3.32, 0.62)	0.18

model than under the OLS regression model.

In the analysis by domains (Table 1), in both the OLS and Tobit models, age was found to be a significant predictor of the OHIP-14 dimension Functional Limitation ($P=0.01$) and gender was found to be significantly associated with the OHIP-14 dimension Psychological Discomfort ($P=0.04$). Under both models, 'profession' was not found to be significantly associated with the OHIP-14 domains Physical Pain and Physical Disability, but it was found to be significantly associated with all the other domains. The effect of profession on these other domains was as found for the total OHIP-14: 'retired' is significantly different from 'unemployed'. In all the cases negative coefficients are obtained, showing 'unemployed' to have the worst outlook. The estimates of the regression coefficients tend to be larger under Tobit regression, confirming that the OLS, without considering floor and ceiling effects, under-estimates the effects. However, as observed, the dif-

ferences were on the marginal side and the results of the significance tests for the regression coefficients were very consistent between the two models, OLS and Tobit. The results obtained under OLS and Tobit's regressions for the total OHIP-14 and the dimensions were compared using log-likelihood and model fit statistics (Table-2). The log-likelihood values for the Tobit model for all the domains and for the total score were less than the likelihood values for the OLS model, indicating that the fit under the Tobit model tends to be better than under the OLS model.

DISCUSSION

The presence of Floor and Ceiling effects in Likert scales has long been established (Van Den Oord and Van Der Ark 1997). The presence of the floor effect in the OHIP-14 was larger than the ceiling effect, indicating a tendency towards a good OHRQOL for this particular patient population.

Many studies have analysed the relationship between OHIP-14 scores and demographic variables although none

seem to have taken floor and ceiling effects into account. Liu et.al (2012), using bivariate analysis, showed that neither age nor sex affect the total OHIP-14 score. Similarly, a weak association between OHIP-14 mean score and age was established by Slade et.al (2011). Slade and Sanders (2011) found this association to be weak for people with no clinical conditions but, in contrast, a strong three-fold inverse association between age and mean OHIP score was found for persons with two or more clinical conditions. Similarly, Macedo et.al (2011), using Fisher's exact test, found no significant association between OHIP scores and demographic characteristics. Our findings, with and without adjustment for floor and ceiling effects, were in agreement with these results, in the sense that neither age nor sex are significant predictors of the Total OHIP-14 score.

The influence of floor and ceiling effects on the Total OHIP composite score is demonstrated in our study by the difference observed between the regression coefficients of both models (OLS and Tobit). The difference in coefficients was maximum for the retired group, whereas, in other predictors it was rather marginal. Therefore, any model fitted without adjusting for floor and ceiling effects when a considerable proportion of cases suffer from these would be underestimating (in terms of magnitude) the effects. Based on simulated data, Matthew McBee (2012) also established similar results. Wang et.al (2009), using empirical and simulated data, showed that ceiling effects in longitudinal data lead to biased parameter estimation. In this paper we also show that ignorance of floor and ceiling effects can lead to misleading results.

In our study, the multiple OLS regression analysis (ignoring floor and ceiling effects) found full time, student and retired persons to be significantly associated with the Total OHIP-14 score. However, the Tobit model (a model that takes floor and ceiling effects into

Table-2: Comparison of dimensions using log-likelihood and model fit statistics

Outcome	Model	Log-likelihood
Total OHIP score	OLS	-1448.16
	Tobit	-996.03
Functional Limitation	OLS	-767.73
	Tobit	-485.15
Physical Pain	OLS	-829.85
	Tobit	-774.38
Psychological Discomfort	OLS	-856.85
	Tobit	-769.10
Physical Disability	OLS	-834.39
	Tobit	-703.00
Psychological Disability	OLS	-844.33
	Tobit	-741.88
Social Disability	OLS	-833.09
	Tobit	-686.13
Handicap	OLS	-841.33
	Tobit	-619.21

account) showed only Retired persons to be significantly associated. The Tobit model demonstrated a significant reduction in OHIP scores for Retired people indicating better quality of life when compared to the unemployed category. If Retired people are expected to have less tolerance and more severe dental problems resulting in poorer quality of life then the results of Tobit regression analysis is in opposition to this assumption. However, this could be viewed in relation to the day to day activities of individuals; as retired people are expected to be involved in less activities when compared to all other categories of people, the impact of oral health problems on their day to day life is expected to be less.

The analyses we present here do not include the clinical conditions of the patients who were involved. The analysis, including the clinical conditions, will no doubt help to explore the reasons for better quality of life for retired people in relation to unemployed. However, the objectives of the analysis we present in this paper show that oral health related quality of life using OHIP-14, as an outcome measure in a Likert scale, suffer from floor and ceiling effects and that, even after adjusting for the floor and ceiling effects, gender and age are not predicting the oral health related quality of life

as measured by the OHIP-14. Though the impact of floor and ceiling values were not striking, we verified that the Tobit model fits the oral health related quality of life data better than the OLS model. Hence the Tobit model for OHQoL data improves the model predictions and fits the data better than the OLS model. Our main purpose was to demonstrate the problems of floor and ceiling effects and our findings suggest that any future studies on OHRQoL using OHIP-14 should take these effects into account when exploring the relationship between OHIP-14 composite scores and demographic and clinical variables.

Acknowledgement

The authors would like to thank Dr. Helena Lewis Greene (King's College London) for her help with the data collection. Thanks and thoughts of healing are also due to all the patients who so kindly consented to take part in this study and provided data.

REFERENCES

1. Liu L, Xiao W, He Q, Jiang W. Generic and oral quality of life is affected by oral mucosal diseases. *BMC Oral Health* 2012;**12**:2.
2. Slade GD, Sanders AE. The Paradox of Better Subjective Oral Health in Older Age. *J Dent Res* 2011;**90**(11):1279-85.
3. Macedo CG, Queluz DP. Quality of life and self-perceived oral health among workers

from a furniture industry. *Braz J Oral Sci* 2011;**10**(4):226-32.

4. Wang L, Zhang Z, McArdle JJ, Salthouse TA. Investigating Ceiling Effects In Longitudinal Data Analysis. *Multivariate Behav Res* 2009;**43**(3):476-96.
5. Van Den Oord EJCG, Van Der Ark LA. A note on the use of the Tobit approach for tests scores with floor and ceiling effects . *British Journal of Mathematical and Statistical Psychology* 1997; **50**:3510364.
6. McBee M. Modeling Outcomes with Floor or Ceiling Effects: An introduction to the Tobit Model. *Gifted Child Quarterly* 2010; **54**(4): 314-20.
7. Miller GA. The magical number of seven, plus or minus two: some limits on our capacity for processing information. *Psychological Review* 1956;**63**:81-97.
8. Genia V. Evaluation of the spiritual well-being scale in a sample of college students. *The international journal for the psychology of Religion* 2001;**11**(1): 25-33.
9. Ledbetter MF, Smith LA, Vosler-Hunter WL, Fisher JD . An evaluation of the research and usefulness of the spiritual we-being scale. *Journal of Psychology and Theology* 1991;**19**:49-55.
10. Fitzmaurice GM, Laird NM, Ware JH. *Applied Longitudinal analysis*. Wiley; New York: 2004.
11. Laird N and Ware J. A Random-effects model for longitudinal data. *Biometrics* 1982;**38**:963-74.
12. Tobin J. Estimation of relationships for limited dependent variables. *Econometrica* 1958;**26**(1):24-36.
13. Muthen B. TOBIT factor analysis. *British Journal of Mathematical and Statistical Psychology* 1989;**42**:241-50.
14. Brown RL. Congeneric modeling of reliability using censored variables. *Applied Psychological Measurement* 1989;**13**:151-159.