

Barinder Singh¹, Shubhram Pandey³, Supreet Kaur², Akanksha Sharma³
¹Pharmacoevidence Ltd., London, UK; ²Heorlytics Pvt. Ltd., SAS Nagar, Mohali, India; ³Pharmacoevidence Pvt. Ltd., SAS Nagar, Mohali, India

Objective

Background

- To compare the different regression models where the heavy ceiling effect was observed in the EQ-5D scores

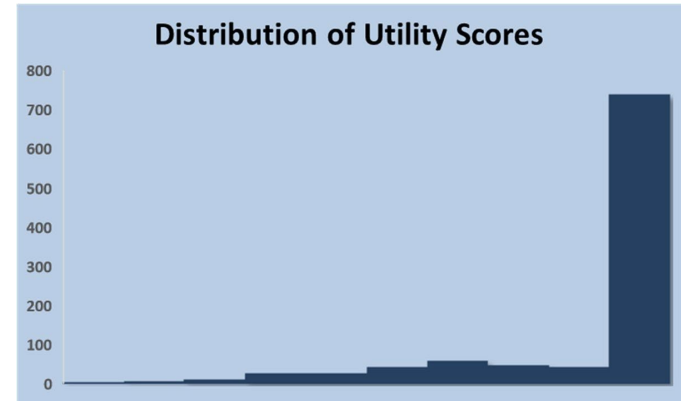


Figure 1: Histogram showing the distribution of utility scores

- The EQ-5D index is the preferable and widely used instrument that serves as the basis for summarizing and comparing health outcomes to assess preference-based health status
- The preference-based approach used to measure Health-related quality of life (HRQoL) quantifies the extent to which an individual prefers a specific health state out of provided health state options that lies between 0 to 1, where 0 indicates death and 1 indicates full health
- A small negative value indicates that life is worse than death
- OLS regression is the most preferred way of generating marginal utility estimates from per-patient utility scores
- However, the OLS regression provides unreliable and biased estimates when a heavy ceiling effect is observed in the utility scores

Methods

- The data for EQ-5D scores and baseline covariates (including age, sex, and weight) were simulated with an assumption that ceiling effect were observed in 75% of patients
- Conventional and non-conventional regression models such as Ordinary Least Square (OLS), Tobit/Censored (around mean/median), Beta regression and its extensions (bias correction and bias reduction), Two-part model (TPM) with OLS, TPM with Tobit, Log-transformed TPM with OLS and Log-transformed TPM with Tobit were fitted on simulated data
- EQ-5D scores were considered as a dependent variable and baseline characteristics were independent variables
- OLS attempts to estimate the coefficient of EQ-5D scores by minimizing the sum of the squared differences between observed and fitted values
- Tobit regression with median assumes that the median is more robust than the mean, to the ceiling effects and minimizes the sum of absolute differences between observed and fitted values
- TPM splits the whole data into two groups, one, at the ceiling and another, below the ceiling and then it fits the regression as usual
- Log transformed TPM transforms the data using a log link function, that are below the ceiling to explicitly recognize the skewed distribution of the data
- The predictive accuracy/goodness-of-fit of all the models was compared using absolute error (R1) and squared error (R2)

Results

- Log-transformed TPM with OLS was best fitted model on EQ-5D data with a heavy ceiling effect having predictive accuracy metric R^1 (-24.39) and R^2 (-444.36) followed by Log-transformed TPM with Tobit R^1 (-22.98) and R^2 (-395.88)
- The least preferable model was the OLS regression model ($R^1 = 0.09$) because of its linearity assumption

Conclusion

- The log-transformed TPM model with OLS was superior followed by log-transformed TPM with Tobit to all the other approaches in handling data with heavy ceiling effects

Table 1: Regression models depicting accuracy measures

Model	Predictive Accuracy	
	R^1	R^2
OLS Regression	0.09	0.18
Tobit Model	-1.76	-5.71
Tobit Model around Median	-1.13	-3.13
Beta Regression	-14.92	23.50
Beta Regression (Bias Correction)	-14.90	23.49
Beta Regression (Bias Reduction)	-14.90	23.50
Two-part model with OLS	-3.98	-73.36
Two-part model with Tobit	-3.98	-73.36
Log-transformed two-part model with OLS	-24.39	-444.36
Log-transformed two-part model with Tobit	-22.98	-395.88