What is the best approach to analyze longitudinal bounded scores? Application to Quality of Life data

A.F. Donneau [1], C. Russo [2], N. Dardenne [1], M. Mauer [3], C. Coens [3], A. Bottomley [3], E. Lesaffre [4]

(1) B-stat, Department of Public Health, University of Liège, Belgium (2) Department of Applied Mathematics and Statistics, University of Sao Paulo, Brazil
(3) EORTC Headquarters, Departments of statistics and quality of life, Brussels, Belgium (4) L-Biostat, KU Leuven, Leuven, Belgium

Introduction

In cancer clinical trials, the patient's quality of life (QoL) is a major outcome measure, generally assessed at specified time intervals by means of by means of Likert-type items questionnaire that covers different domains of the QoL. Usually, the items are summated and linearly transformed to construct a bounded score ranging from 0 to 100. Most papers concerned with the statistical analysis of QoL scores treat them as continuous rather than as bounded variables. The aim of the present study was to compare the results derived from the analysis of longitudinal bounded QoL scores from an EORTC trials under different statistical approaches, namely the linear mixed-effects model and the beta regression model.

Material

The EORTC 26981 study is a randomized multicenter phase III trial that evaluated the addition of temozolomide (TMZ) to standard focal radiotherapy (RT) in 573 patients with newly diagnoses glioblastoma. QoL was assessed using the EORTC QLQC30 version 2 questionnaire [1] and the Brain Cancer Module (BN20): at baseline; during radiotherapy at week 4; 4 weeks after completion of radiotherapy; at the end of the third and sixth cycle of adjuvant temozolomide; and every 3 months thereafter until disease progression for patients allocated RT+TMZ, and at equivalent time points for those allocated RT.

Statistical Methods

Differences between the two treatment groups were tested using the beta regression model (for bounded values) [2] and the linear mixed-effects model (for continuous values) [3]. To fit to the condition of application of the beta distribution, QoL scores were divided by 100. In the following, consider a sample of N subjects and let Y be an bounded QoL outcome variable assessed on T occasions on each patient. Denote by $\mathbf{Y}_{ij}$ the assessments of Y on the ith patient at the jth occasion. Associated with each patient, there is a $p \times 1$ vector of covariates, say $\mathbf{x}_{ij}$ measured at time j.

1. Linear mixed-effects model (LMM) :

$$\mathbf{Y}_{ij} = \mathbf{x}_{ij}^T\boldsymbol{\beta} + \mathbf{z}_{ij}^T\mathbf{b}_i + \boldsymbol{\epsilon}_{ij}$$

where $\mathbf{b}_i \overset{ind}{\sim} N(0, G)$, $\boldsymbol{\epsilon}_{ij} \overset{ind}{\sim} N(0, \sigma^2)$, with $\mathbf{b}_i$ and $\boldsymbol{\epsilon}_{ij}$ independent from each other.

2. Beta regression model:

$$log\frac{\mu_{ij}}{1-\mu_{ij}} = \mathbf{x}_{ij}^T\boldsymbol{\beta} + \mathbf{z}_{ij}^T\mathbf{b}_i$$

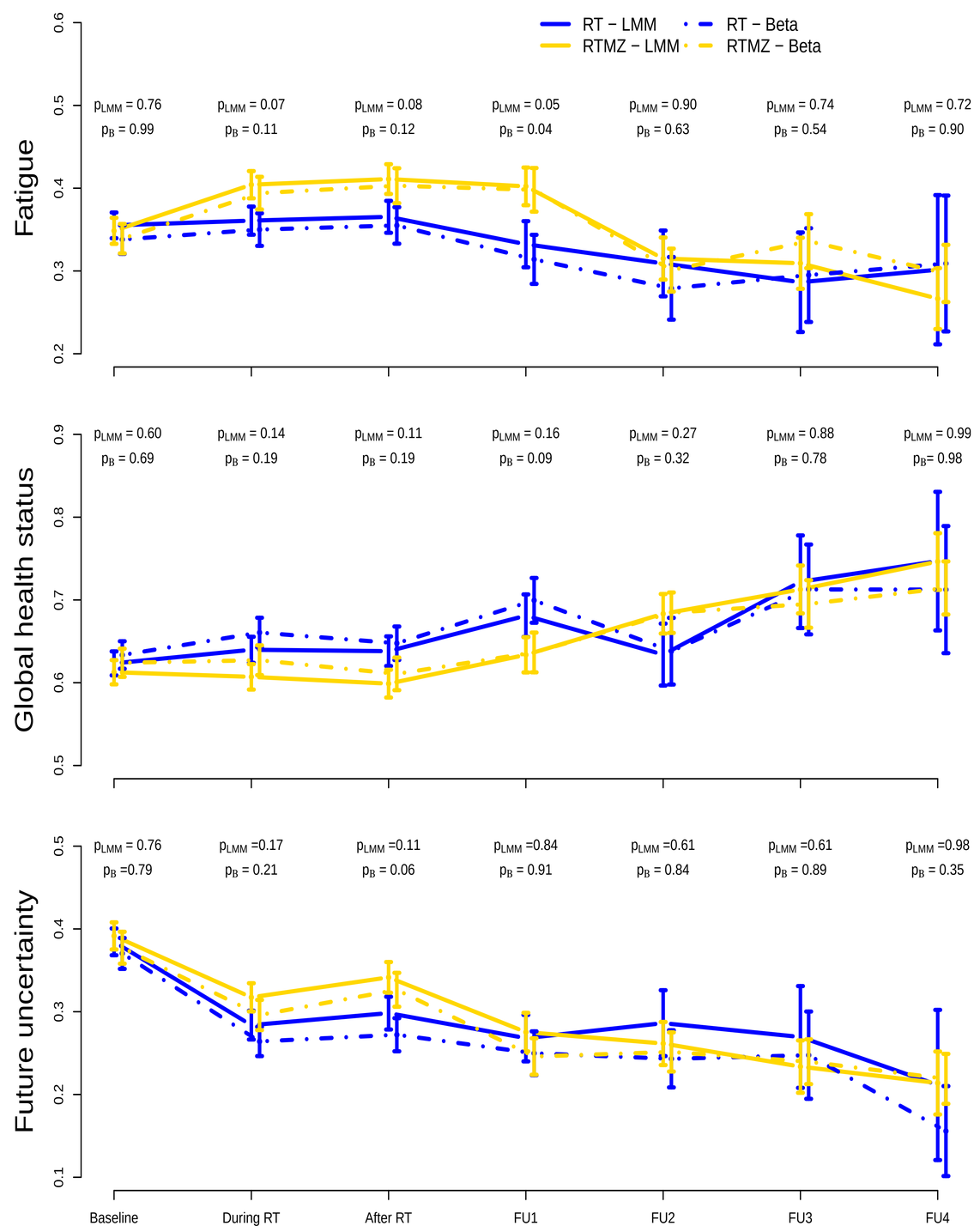
where $\mathbf{b}_i \sim N(0, G)$ and μ is the mean parameter of the beta distribution.

In both models, $\mathbf{z}_{ij}$ is vector related to random-effects, $\boldsymbol{\epsilon}_{ij}$ and $\mathbf{b}_i$ the error and the random effects vetor, respectively and G denotes the positive definite covariance matrix of the random effects.

Results

QoL scale	Information criteria	LMM	Beta
Fatigue	-2 Log Lik	-305.9	-785.6
	AIC	-273.9	-753.6
	BIC	-205.4	-685.0
Global Health	-2 Log Lik	-534.3	-874.3
	AIC	-502.3	-842.3
	BIC	-433.7	-773.7
Future uncertainty	-2 Log Lik	-299.1	-1109.7
	AIC	-267.1	-1077.7
	BIC	-198.6	-1009.2

Comparison of the information criteria revealed that lower values were found when using the beta regression approach. Estimated mean scores ± SE in both arms and P-value related to treatment effect at each assessment time for the three considered bounded QoL scales are depicted here below. Both statistical approaches presented comparable results.



Conclusions

The preliminary analysis of these QoL scales showed that both statistical approaches led to the same conclusion when considering the treatment effect, P-values and the mean scores. However, the beta regression model presented a better model fit for the QoL scales. This indicates that incorporating the bounded outcome assumption into the analysis methods can improve QoL hypothesis testing. Those results need to be confirmed by investigating application of other models for longitudinal bounded outcome scores, such as truncated regression model and coarsening approach. Simulation analysis would also be of interest to investigated the impact of various distribution (well-balanced, right or left skewed, U-shape) for the bounded QoL scores.

References

[1] Aaronson, N. et al. The EORTC QLQ-C30: A quality-of-life instrument for use in international clinical trials in oncology. *Journal of the National Cancer Institute*, 1993, **85**, 365 – 376.
[2] Ferrari, S. and Cribari-Neto, F. Beta Regression or Modeling Rates and Proportions. *Journal of Applied Statistics*, 2004, **31(7)**, 799 – 815.
[3] Verbeke, G., and Molenberghs, G. *Linear mixed models for longitudinal data*. 2000. New York: Springer.