

Introduction

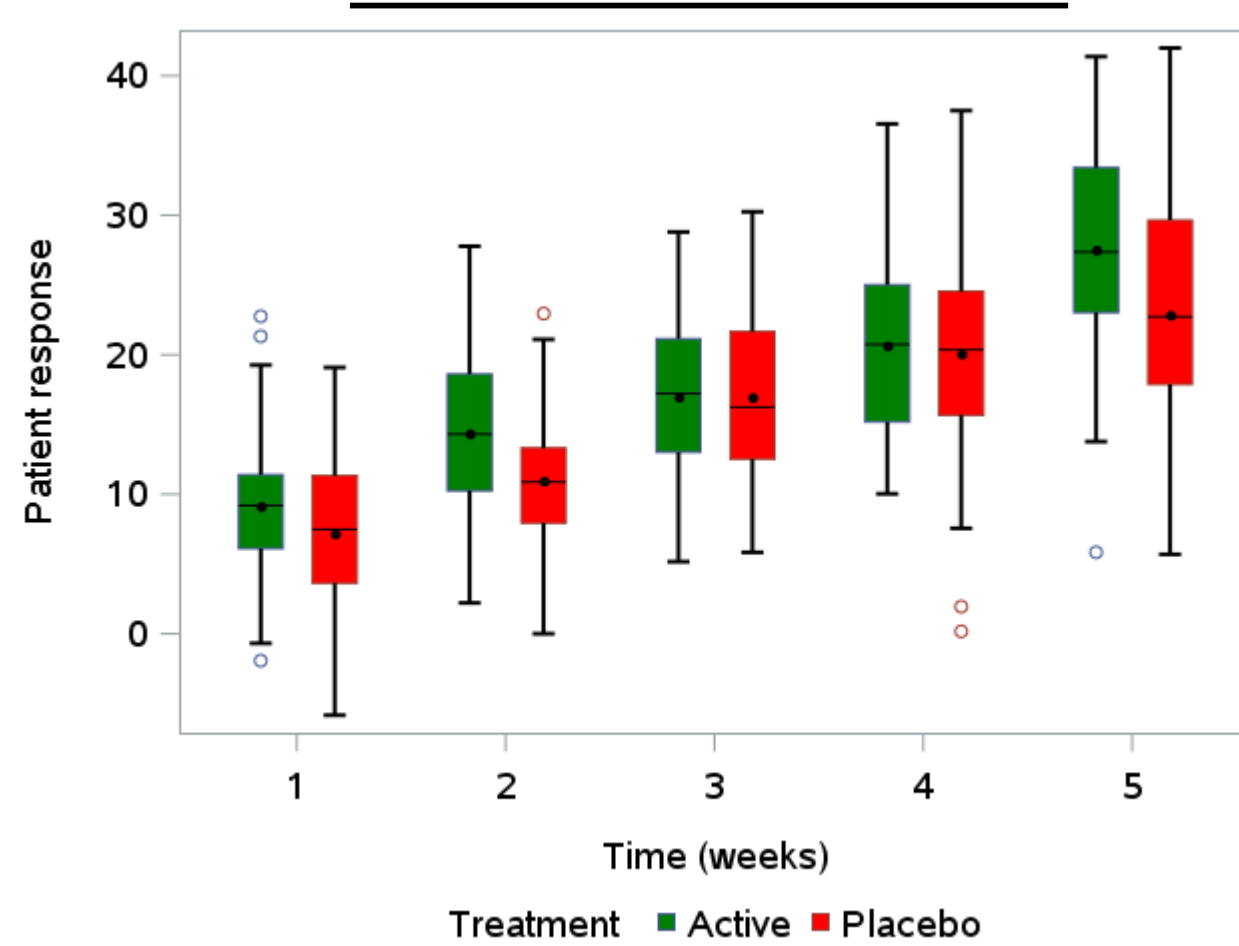
- Longitudinal studies data involve repeated measurements on subjects followed over time.
- For a continuous response, the most commonly used methodology is the linear mixed model (LMM) which can be performed via the *Procedure Mixed* in SAS or via the function *lmer* in R.
- LMM depends on fulfillment of primary data assumptions; data are normally distributed and that the values of the outcome are linearly related to independent variables.
- LMMs are sensitive to outliers occurring either at the level of the within-subject or the between-subject errors or both.
- In the case of skewed data, outliers or small sample sizes, parametric methods may result in unreliable or false conclusions.
- An alternative way would be use of nonparametric (rank-based) methods that are robust to outliers and are flexible to use on different forms of data including heavily skewed data.
- We illustrate a rank-based method and show how the results differ from the use of standard LMMs.

Data

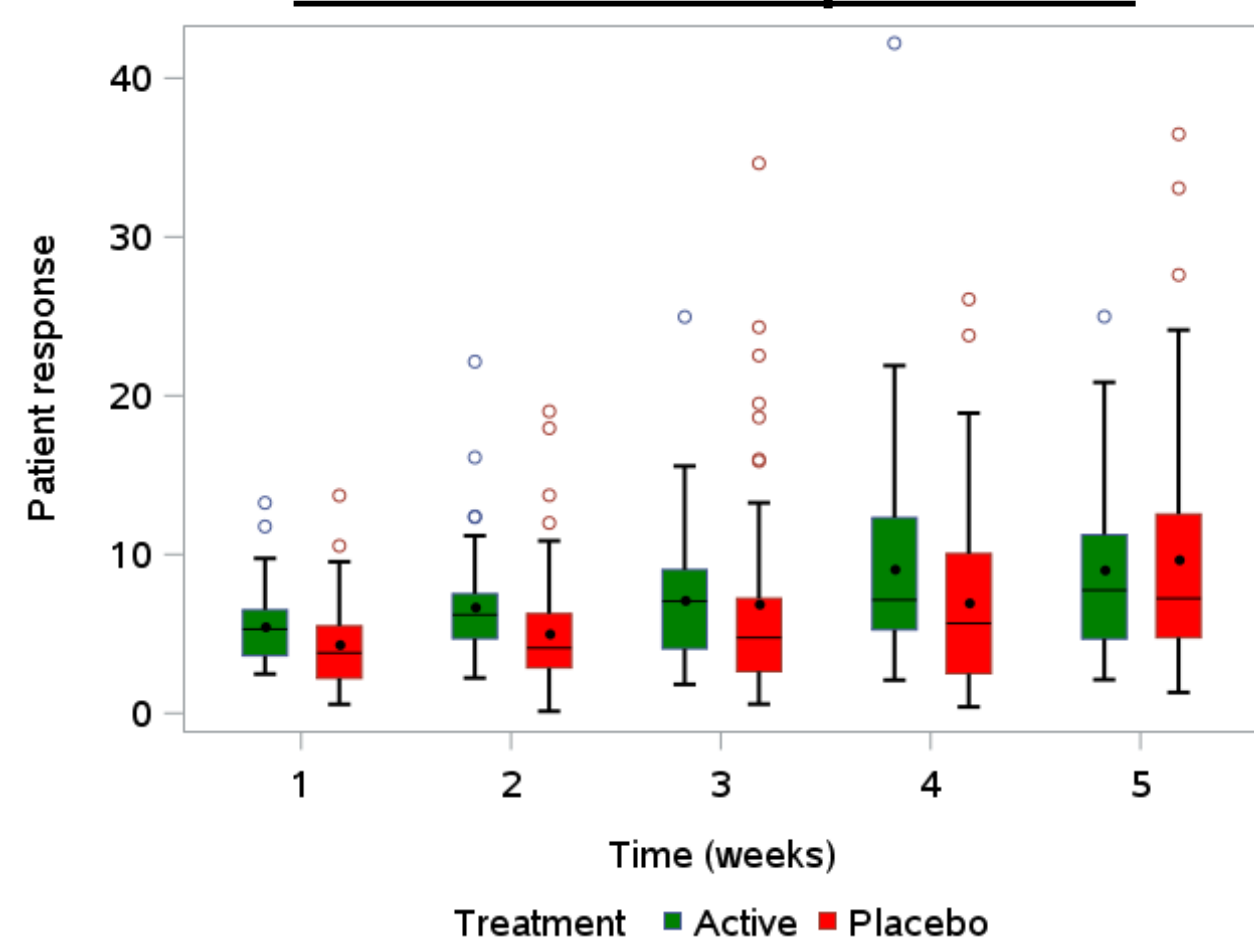
- Data used involve a simulation of a clinical trial setting to evaluate treatment effect on response for 1) normal distribution and 2) chi-square distribution as follows:

- 111 subjects with response over 5 time points
- Intercept and slope random effects following the response distribution
- Treatment effect of factor 1.5

Simulated Normal Data



Simulated Chi-Square Data



Methodology

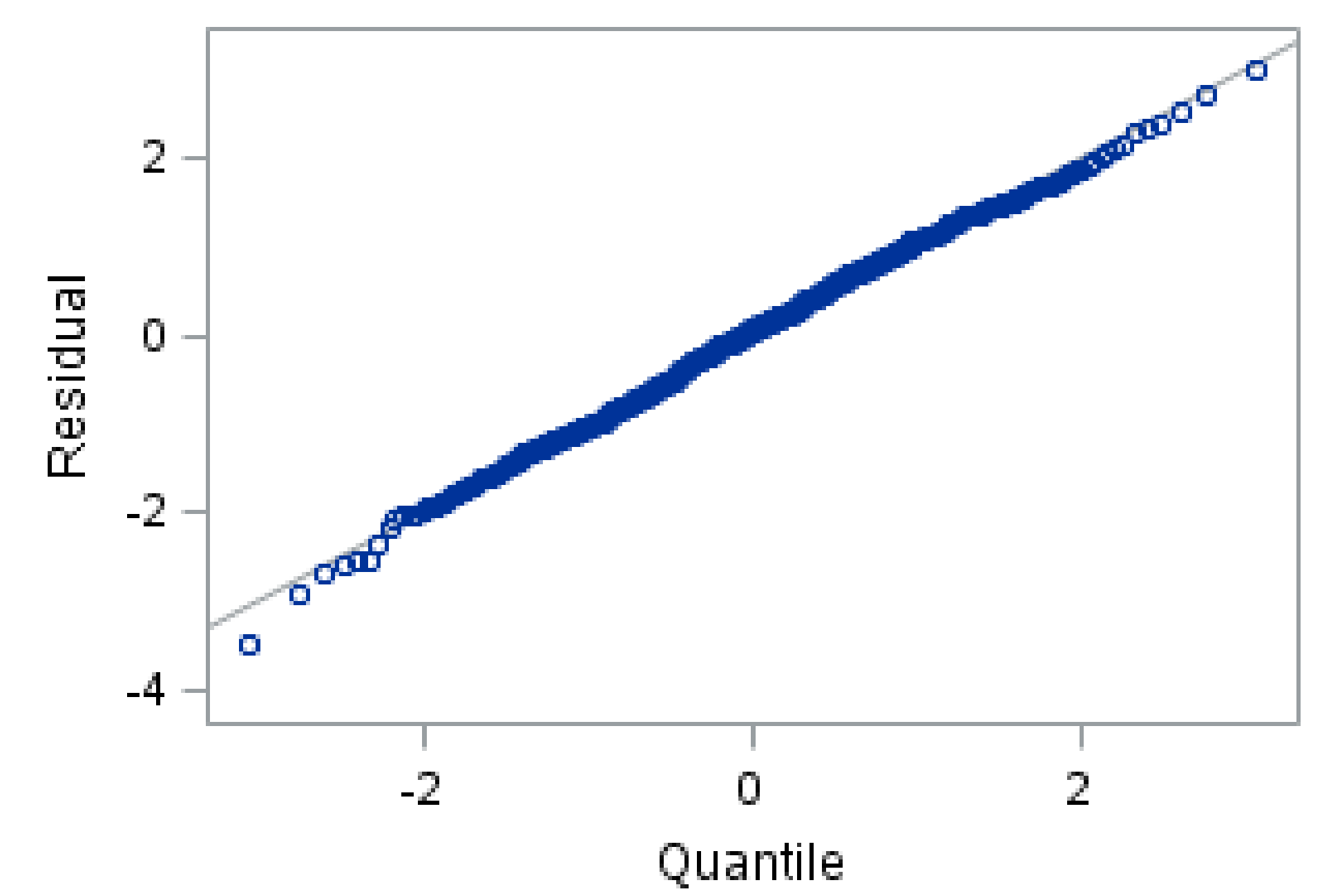
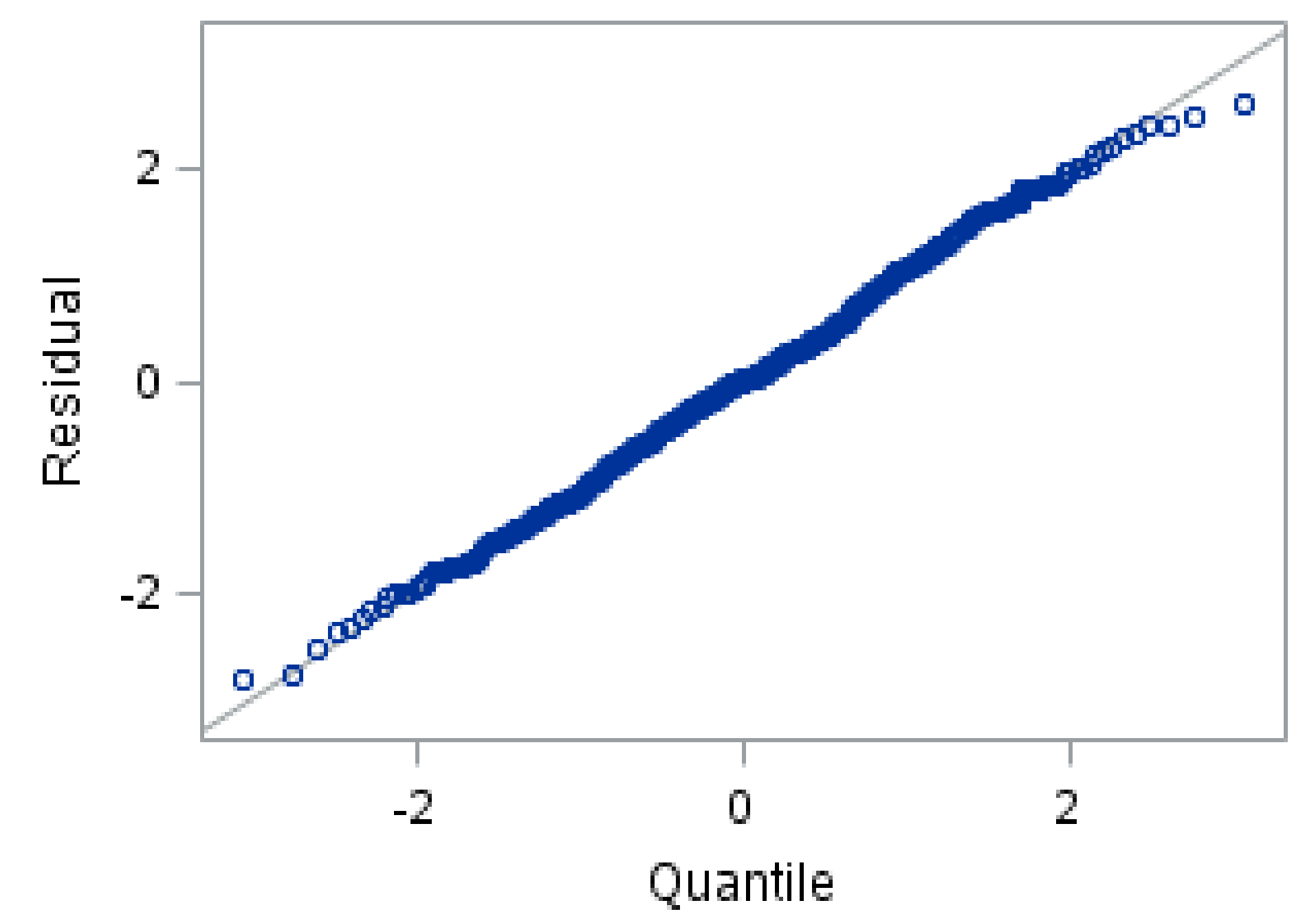
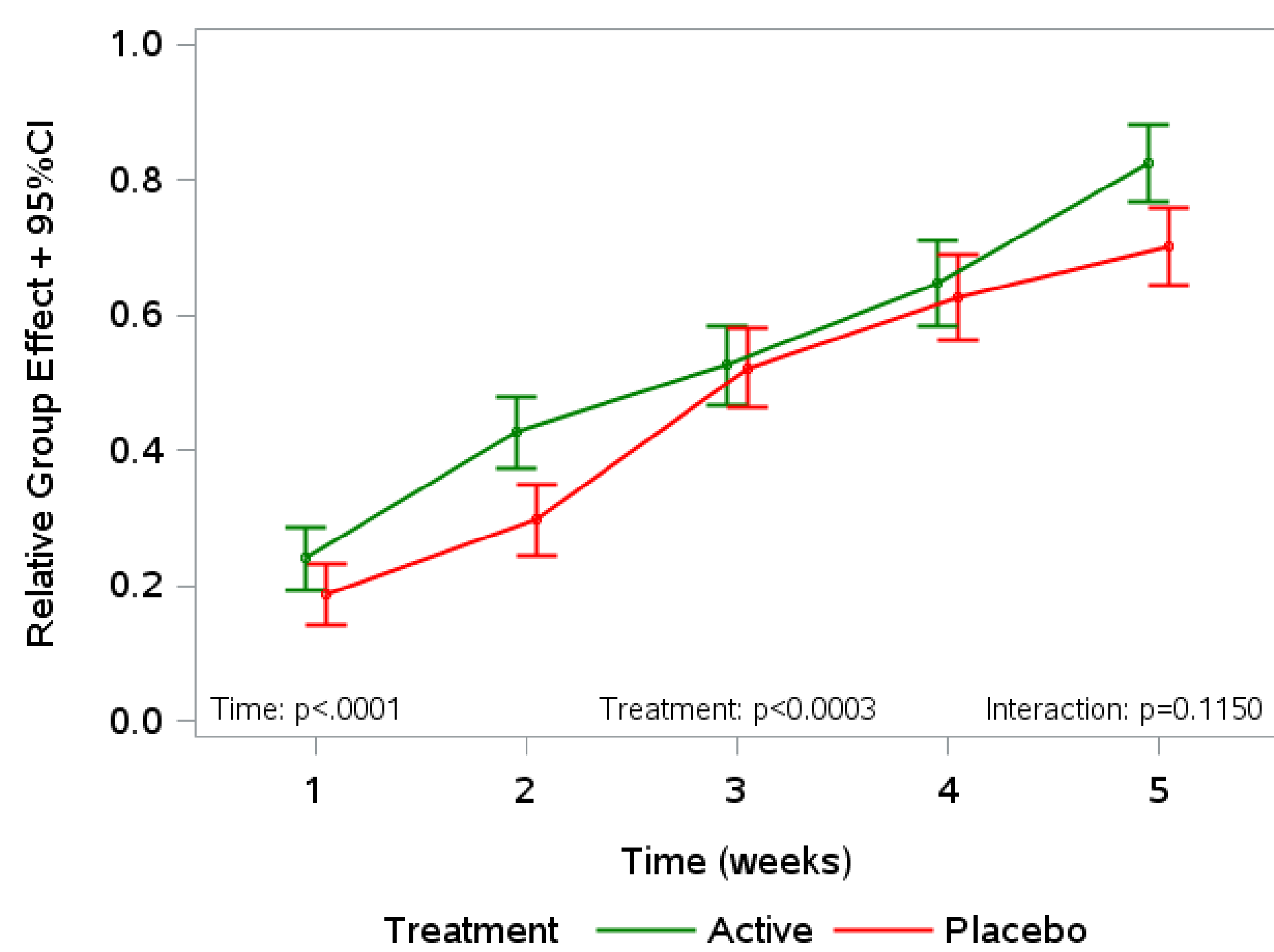
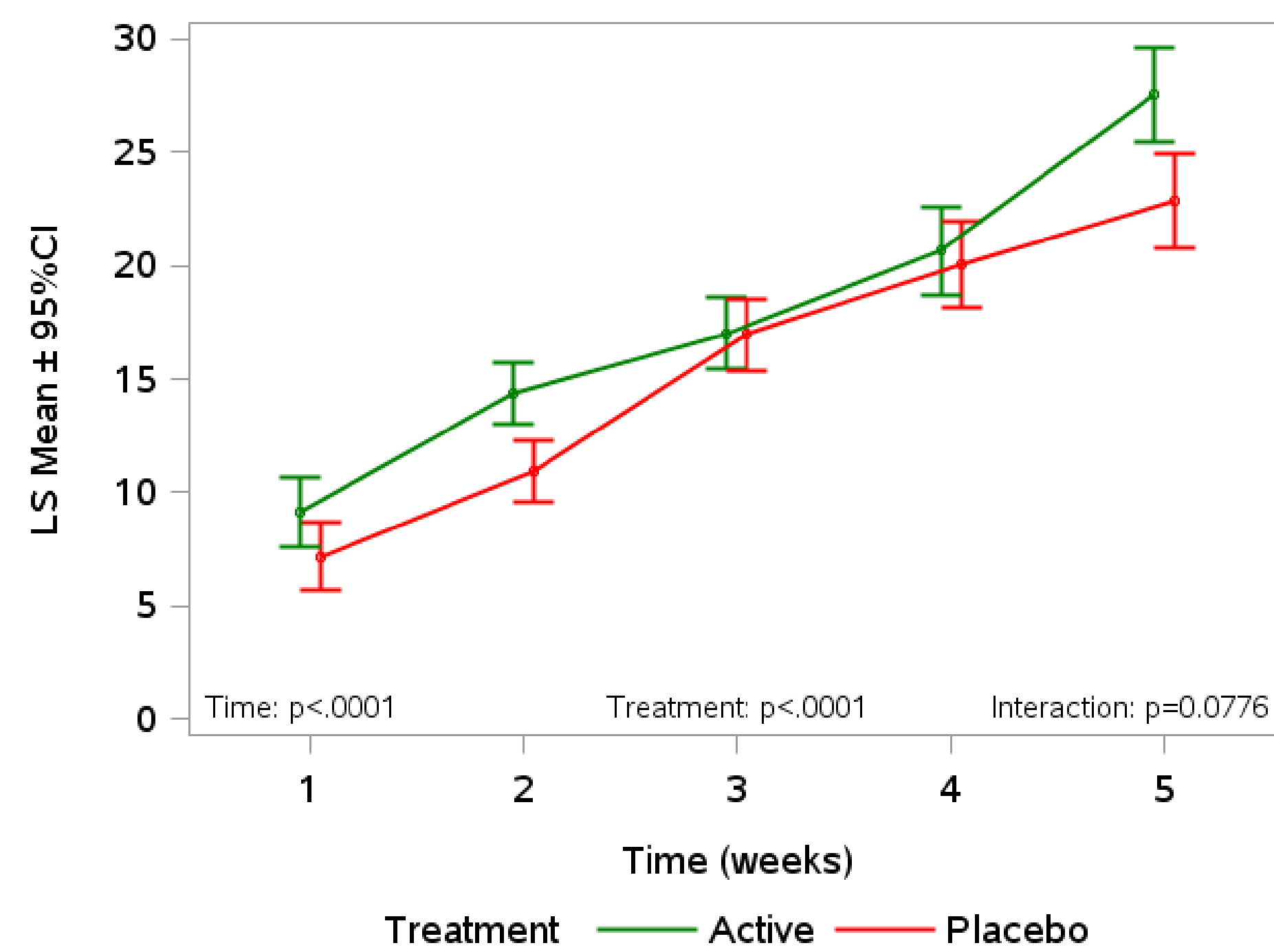
- Linear mixed models with time as categorical, treatment and their interaction were performed using SAS V9.4 on:

- Original response: LS means and their 95% CIs are presented
- Rank-transformed (mid-ranks) response: Relative group effects, \widehat{p}_{is} , based on rank mean estimates, \widehat{R}_{is} , and their 95% CIs are presented.

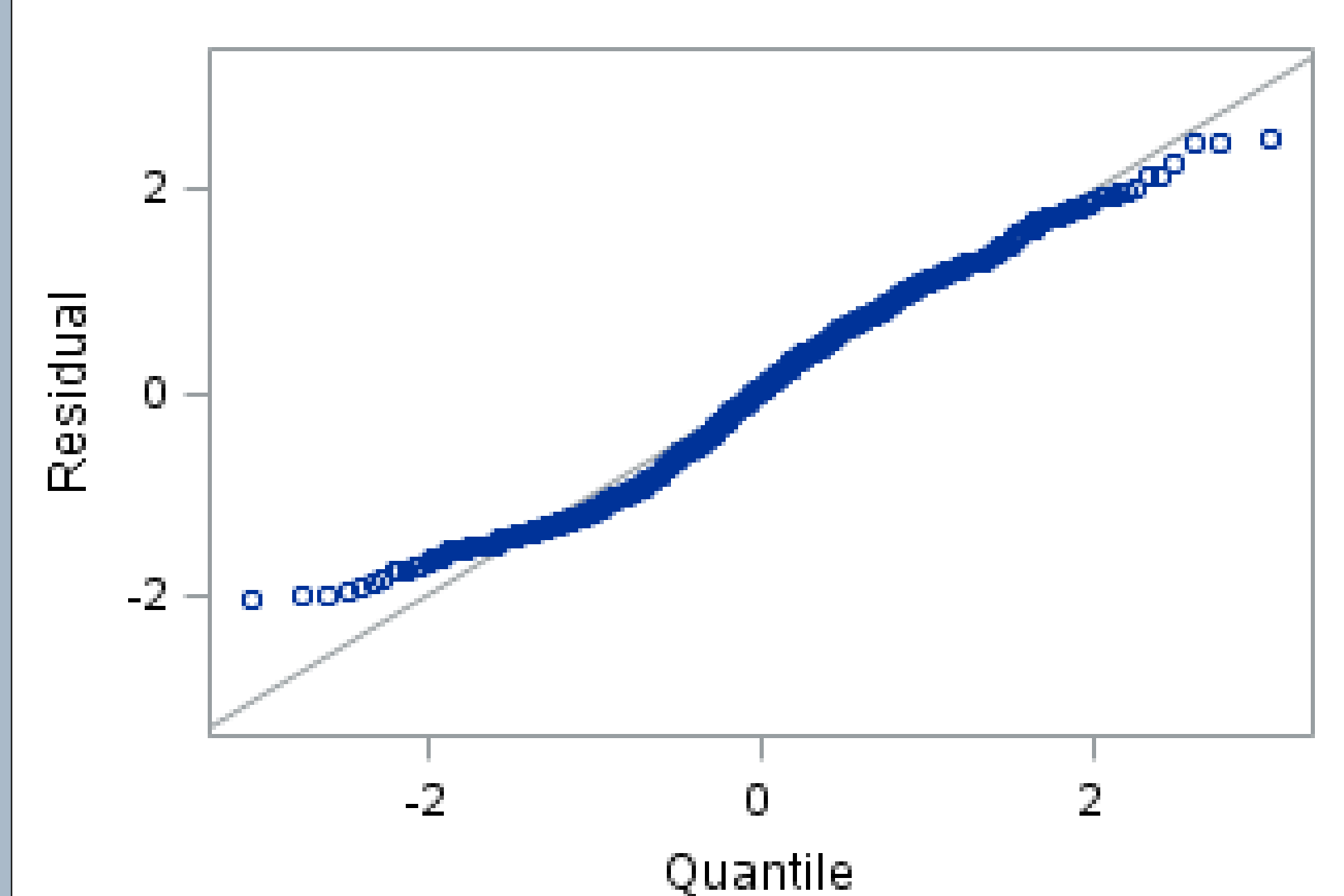
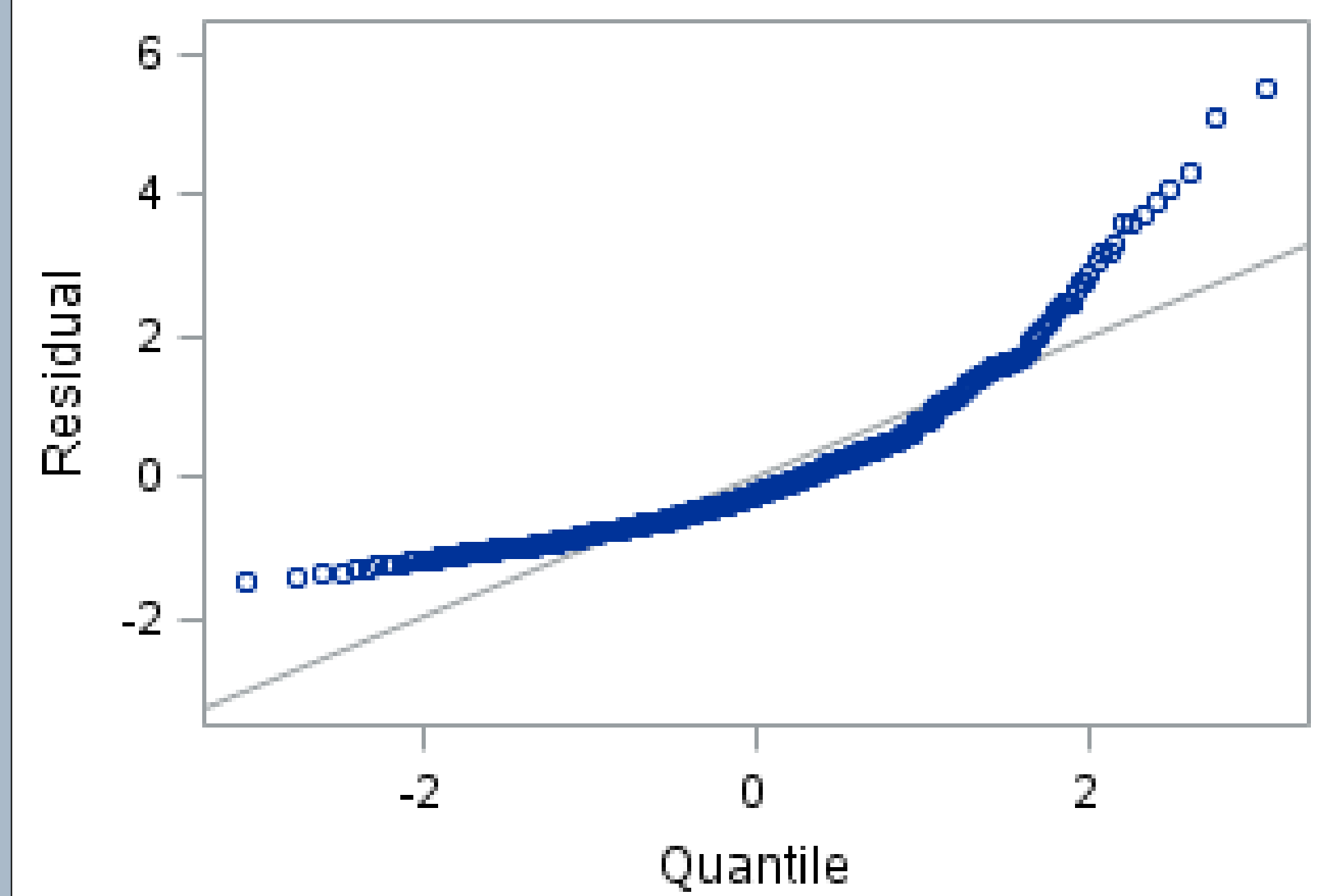
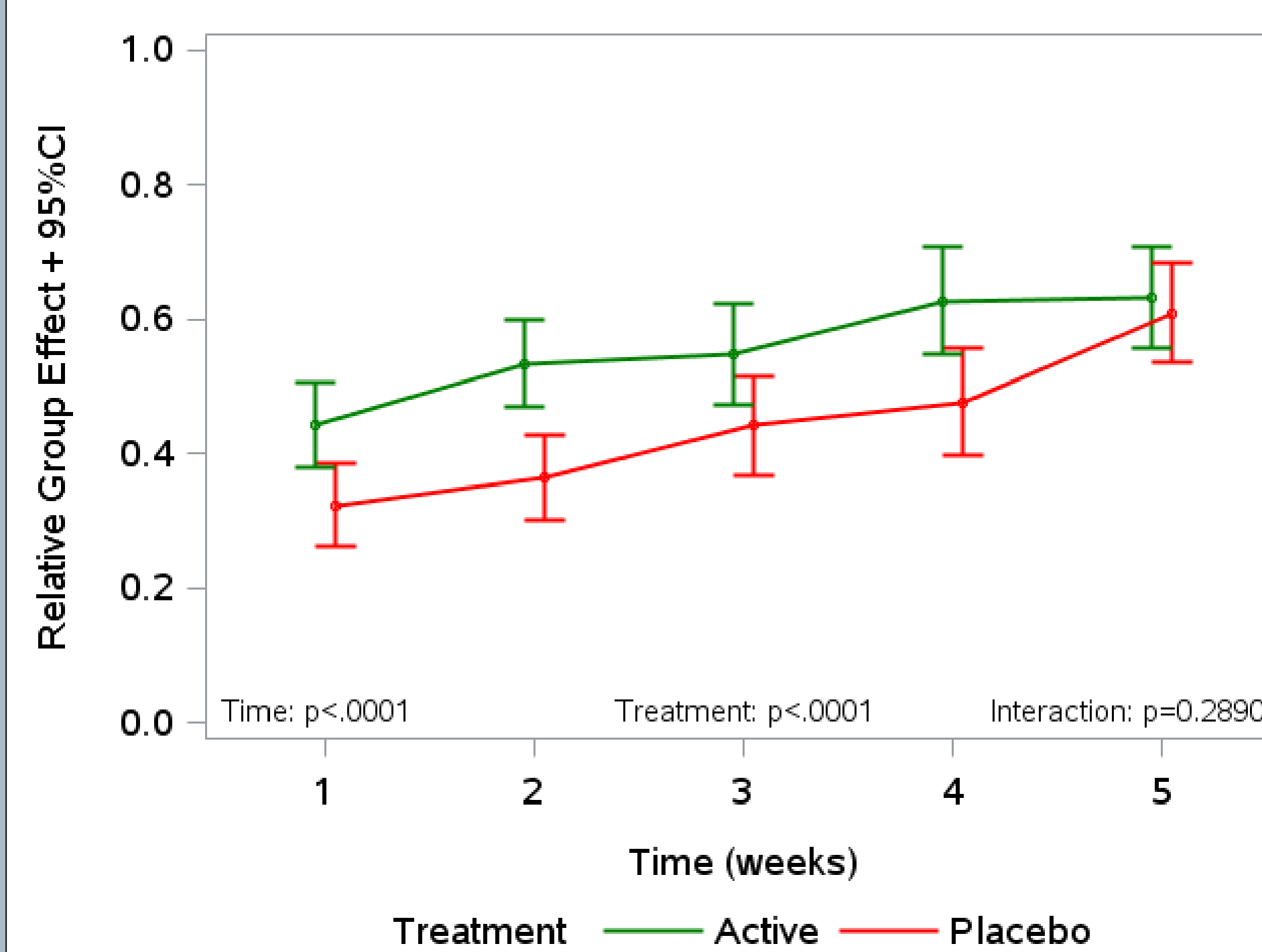
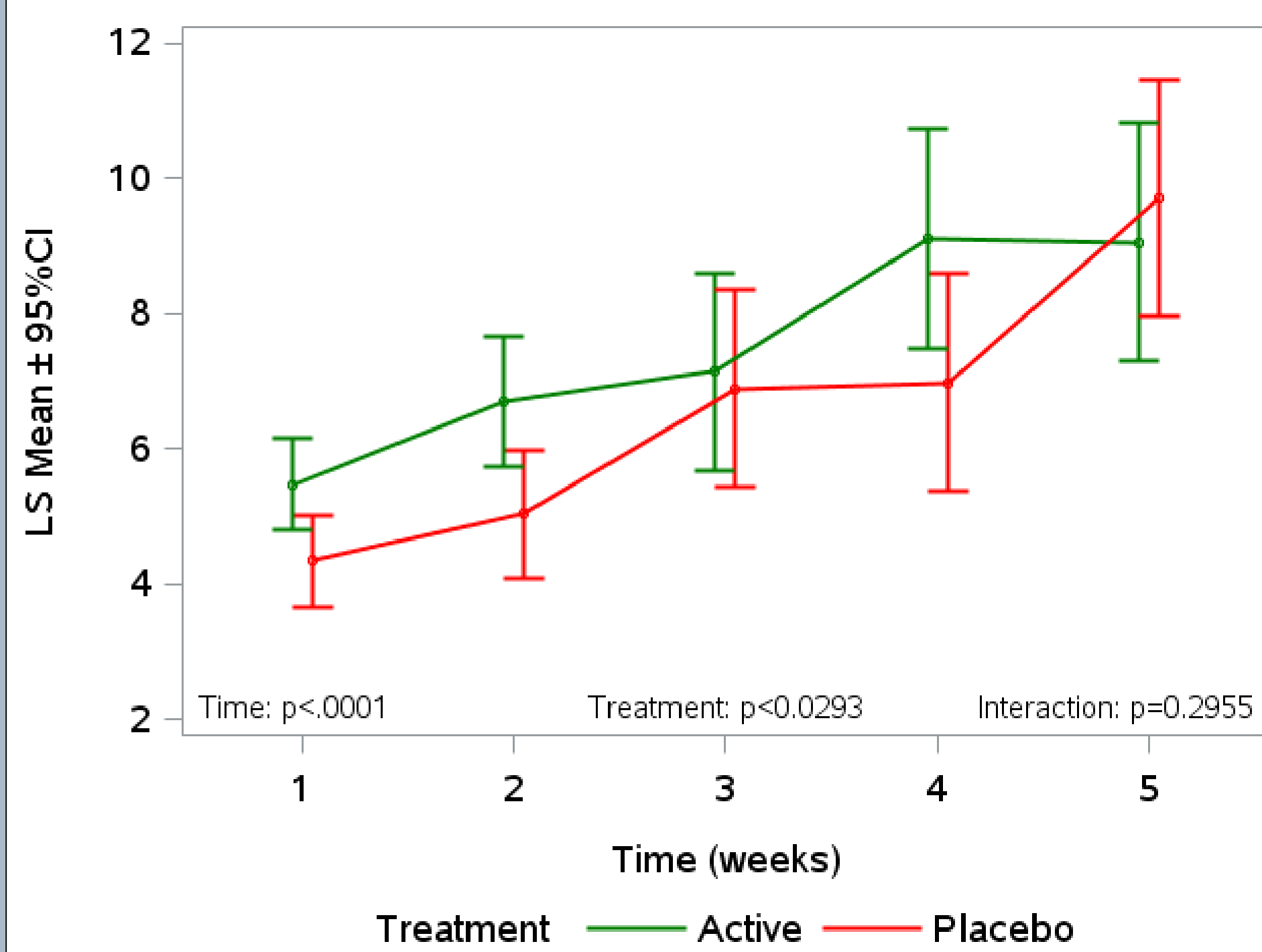
$$\widehat{p}_{is} = \frac{1}{N} (\widehat{R}_{is} - 0.5), i=1, \dots, g; s=1, \dots, t; 0 \leq \widehat{p}_{is} \leq 1$$

N is the number of observations used in analysis

Results for Normal data: LMM based on original response and ranks respectively



Results for Chi-Square data: LMM based on original response and ranks respectively



Discussion

- For normal data, CIs for LS Means and Relative Effects are comparable.
- For non-normal data, wider CIs are observed for LS Means compared to those for Relative Effect.
- For the non-normal data QQ-plot on the original response, data points form a curve that deviates from a straight line and markedly gets better on the rank-based QQ-plot.

Conclusion

- In case of non-normal data and/or outliers, rank-based methods offer a flexible and robust framework for data analysis.
- These methods are not restricted to only continuous outcomes but can also be applied to ordered categorical and count data or heavily skewed data.