BIOS 667 Group 5

Lily Bai Nathalie Blasco Yihao Peng Lauren Rackley Sirui Wu

2025-11-19

week	site	id	treat	age	sex	twstrs
0	1	1	2	65	1	32
2	1	1	2	65	1	30
4	1	1	2	65	1	24
8	1	1	2	65	1	37
12	1	1	2	65	1	39
16	1	1	2	65	1	36

Introduction

The purpose of this project is to examine the effects of botulism toxin type B (BotB) to treat cervical dystonia over time. Cervical dystonia (CD) is a chronic neurological disorder, in which patients have painful involuntary contractions in neck muscles. CD is more prevalent in women (Jankovic et al., 2023). The prevalence of CD is estimated to be 28-183 cases per million. The data comes from a multicenter randomized clinical trial for cervical dystonia patients with 9 U.S. sites. Botulism toxin types A and B are first-line treatments for CD (Wetmore et al., 2025). The treatment groups included in the study were placebo, 5000 U BotB, and 10000 U BotB. The response variable is Toronto Western Spasmodic Torticollis Rating Scale (TWSTRS) total score, which ranges from 0 to 85 and is comprised of disability (0-30), pain (0-20), and severity (0-35) subscores. Only total score is included in this data.The TWSTRS score was measured at week 0 (baseline), and 2,4,6,8,12, and 16 weeks after treatment start. Site is included in the dataset but no further details about site were included in the available dataset documentation.

Methods

Study Population

Inclusion and exclusion criteria for the trial were not available. Duration of CD and age at onset were not known. The study included 109 patients (67 (61%)) females.The mean age was 56 (12). Median age was 56 years. The mean TWSTRS score at baseline was 46 (10). It was not known if the patients received prior BotB treatments. Information about the randomization schedule was not provided.

Statistical Analyses

Number of observations, mean, median, standard deviation (SD), minimum (min) and maximum (max) were provided for age. Mean and SD were calculated for TWSTRS score at baseline. Frequencies and percentages were reported for categorical variables. GLM, GLMM, and GEE models were fit using TWSTRS total score as the response variable and blank, blank, blank as covariates.

Results

Clear Figures/tables describing and summarizing key outcomes/variables Clear Figures/table illustrating results that directly addresses question(s) Proper and clear descriptions of results based on figures and tables, along with diagnostics Code chunks should be hidden in the document, with only text, figures, and tables showing

Table 1: Summary of Demographic and Baseline Characteristics

Characteristic	**Overall** N = 109	**Placebo** N = 36	**BotB**		**p-value**
			5000 U N = 36	**10000 U** N = 37	
sex					0.0706
Male	42 (39%)	15 (42%)	18 (50%)	9 (24%)	
Female	67 (61%)	21 (58%)	18 (50%)	28 (76%)	
Age (years)					0.6198
No. obs.	109	36	36	37	
Mean (SD)	56 (12)	54 (12)	57 (12)	56 (12)	
Median	56	56	57	54	
Min, Max	26, 83	26, 79	35, 83	34, 76	
TWSTRS total score at baseline					0.3307
Mean (SD)	46 (10)	44 (9)	46 (10)	47 (10)	

[1] BotB = botulinum toxin type B; TWSTRS = Toronto Western Spasmodic Torticollis Rating Scale.

[2] Pearson's Chi-squared test; Kruskal-Wallis rank sum test

Generalized Linear Model (GLM)

A generalized linear model (GLM model) using a Gaussian link was fit including week, site, treatment, age, and sex as predictors (Table 2).

Table 2: GLM Model Summary (Gaussian Link)

term	estimate	std.error	statistic	p.value
(Intercept)	30.227	2.850	10.605	0.000
week	0.234	0.080	2.923	0.004
treat5000 U	-0.030	1.120	-0.027	0.978
treat10000 U	0.287	1.127	0.254	0.799
age	0.072	0.039	1.854	0.064
sexFemale	1.938	1.006	1.927	0.054
site2	13.216	1.860	7.107	0.000
site3	-2.183	1.933	-1.130	0.259
site4	4.635	2.132	2.174	0.030
site5	7.403	2.413	3.068	0.002
site6	9.617	1.849	5.201	0.000
site7	1.862	1.912	0.974	0.331
site8	-2.493	1.773	-1.406	0.160
site9	10.198	2.011	5.071	0.000

The coefficient for week was statistically significant, indicating an overall linear trend in TWSTRS score over time. There was notable variation across study sites, with several sites showing significantly higher TWSTRS scores compared to the reference site, highlighting site-level differences. However, because the GLM assumes independence of observations, the repeated measures within individuals violate this assumption. As a result,

the standard errors and p-values may be underestimated, and inference should be interpreted with caution. This motivates the subsequent use of correlation-aware models such as GEE and GLMM.

GLM Model Diagnostics

Diagnostics were assessed for the GLM model.

Figure 1: Residuals vs Fitted Plot

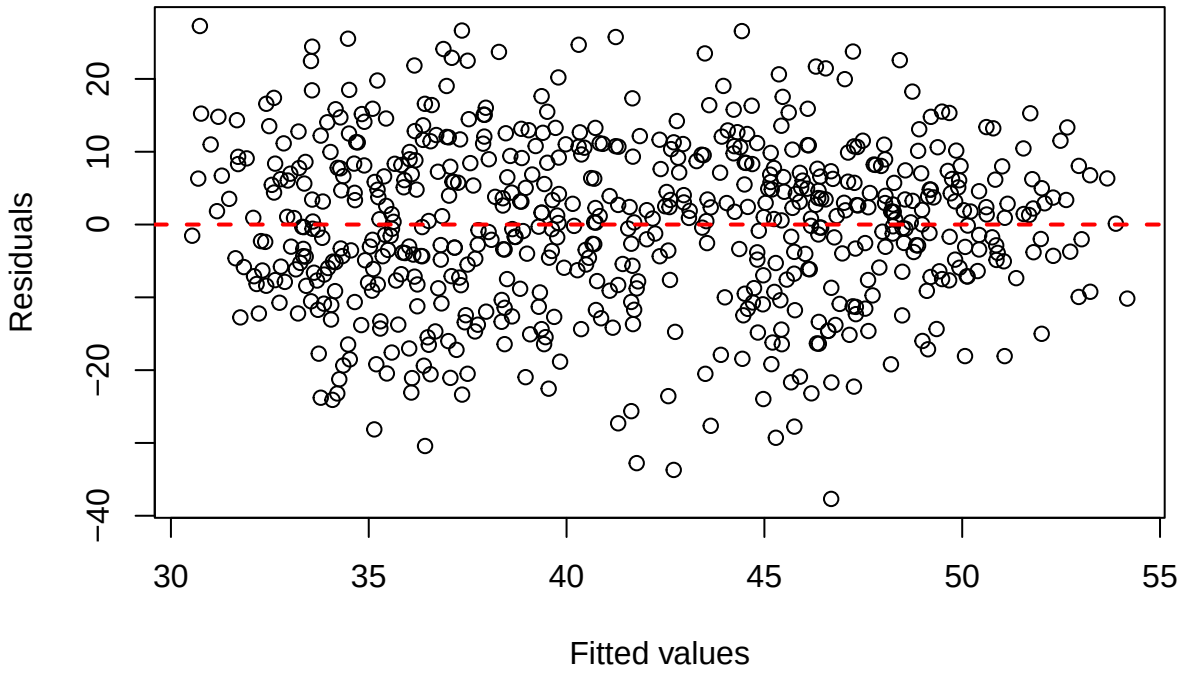


Figure 1 shows the residuals versus the fitted values. The absence of strong patterns or fanning suggests that the linearity and homoscedasticity assumptions are reasonably met.

Figure 2: Q−Q Plot

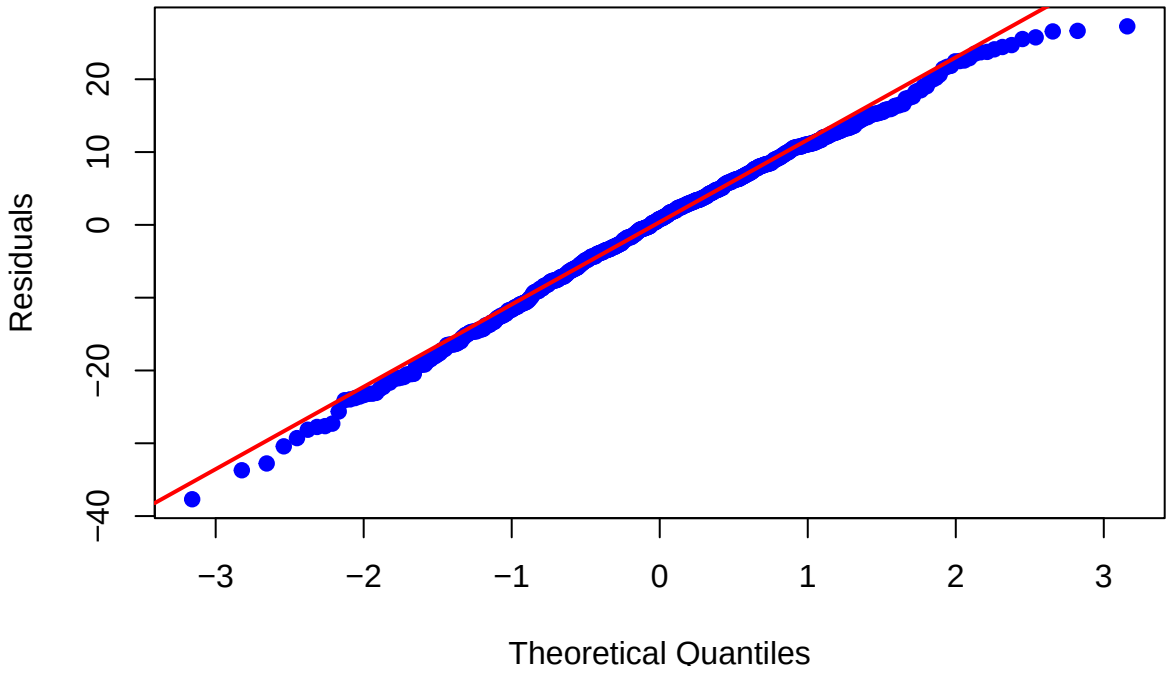
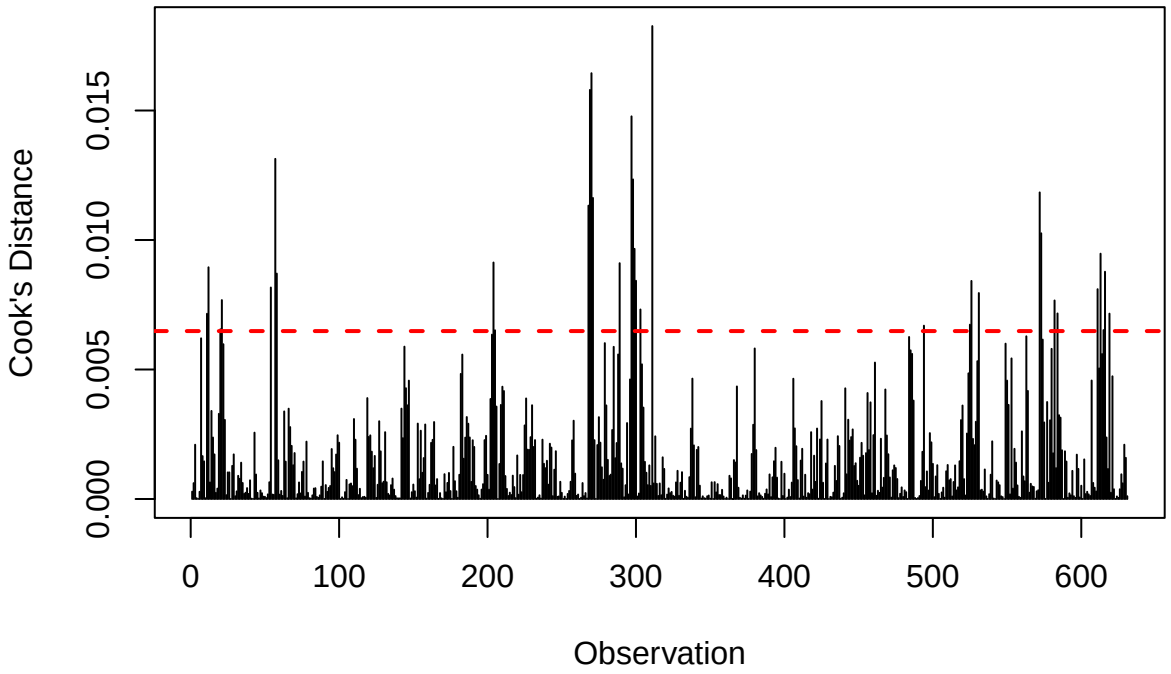


Figure 2 presents a QQ plot of the residuals, which largely follow the 45 degree reference line with mild deviations in the tails. This indicates that the normality assumption is approximately satisfied.

Figure 3: Cook's Distance Plot



Finally, figure 3 displays Cook's distance for all observations. The conventional threshold, 4/(n-k-1), where n is the number of observations and k is the number of predictors, was used to assess the influence of the observations. Several observations exceeded the threshold and were considered potentially influential and

warranted further investigation. Therefore, a second GLM was fit, excluding those observations.

Table 3: GLM Model Summary (Excluding Influential Observations)

term	estimate	std.error	statistic	p.value
(Intercept)	31.836	2.571	12.384	0.000
week	0.215	0.072	3.002	0.003
treat5000 U	-0.185	1.004	-0.184	0.854
treat10000 U	0.800	1.001	0.799	0.424
age	0.050	0.035	1.449	0.148
sexFemale	1.046	0.903	1.158	0.247
site2	13.405	1.667	8.043	0.000
site3	-1.417	1.741	-0.814	0.416
site4	7.636	1.947	3.921	0.000
site5	10.761	2.283	4.714	0.000
site6	10.152	1.661	6.111	0.000
site7	2.380	1.707	1.394	0.164
site8	-3.384	1.605	-2.108	0.035
site9	10.049	1.861	5.400	0.000

Excluding influential observations resulted in modest shifts in the parameter estimates and slightly improved precision. Notably, the effect of week remained statistically significant and variation across study sites continued to be pronounced. This underscores that site-level differences remain important even after removing influential observation.s

Overall, excluding influential points slightly adjusted the estimates and improved precision, but the main patterns of association were consistent with the original GLM. As with the previous model, the GLM assumptions are not appropriate for longitudinal data with correlated repeated measures. The following GEE and GLMM analyses address this limitation by explicitly modeling within-subject correlation and random effects.

Discussion

Summarization of main points, conclusions based in results Summarization of the various models applied, which ones you prefer and base your interpretations off of and why Discussion of limitations if any

References

Add references here later