Data analysis of cervical dystonia disease - repeated measures data research, Complex Data

Piotr Dzierza

June 2019

1 Introduction

This research is a data analysis of longitudinal data. This type of data is characterized by measuring an object of interest more than once, where every measurement is taken after some, usually specified, time. This type is commonly encountered in medicine, crime fighting fields or clinical psychology. Longitudinal data allow for the measurement of within-sample change over time, enable the measurement of the duration of events, and record the timing of various events.

This paper presents an analysis of cervical dystonia disease - repeated measures data research where the main point of interest lies in patient's response to treatment. To properly examine the effects of medication course there is a need to not only get the results after the treatment but to also gather the information during course. It is a natural way to obtain longitudinal data.

The research of this data will be based on profile analysis. We are going to investigate whether the covariance structure can be simplified from unstructured form. Then we will try to fit an appropriate reduced model for the effect of time based on the chosen covariance matrix. Another task afterwards will be simplifying the model by checking the hypothesis about parallelism. The analysis will end with testing the null hypothesis that the estimated response means at week 16 are equal for the three groups of treatment and compare each of the two botulinum neurotoxin treatment groups with different doses to the placebo group.

2 About cervical dystonia

Cervical dystonia, also known as spasmodic torticollis, is a rare neurological disorder that originates in the brain. It is the most common form of focal dystonia in an office setting. Cervical dystonia is characterized by involuntary muscle contractions in the neck that cause abnormal movements and postures of the neck and head. In some cases, these abnormal contractions may be sustained or continuous; in others, they may be present as spasms that can resemble tremor. The severity of cervical dystonia can vary, but the disorder can cause significant pain and discomfort as well as difficulty due to the abnormal postures. It can affect quality of life and activities of daily living including employment. Cervical dystonia typically begins in middle age, and rarely begins in adolescence and young adulthood. The cause of cervical dystonia is unknown, although a genetic susceptibility is thought to underlie some cases. Women are more

likely to develop cervical dystonia than are men. Overall the number of people with adult onset dystonia in the population is about 4.9 people per 100,000. The current treatment of cervical dystonia is local injections of botulinum neurotoxin. Botulinum neurotoxin type A (BotA) is the most frequently used; type B (BotB) is only proposed in selected cases.

3 Dataset overview

This set of data contains information about 109 patients with cervical dystonia. There were 3 forms of trial:

- placebo
- 5000 units of BotB
- 10000 units of BotB.

First group has 36 patients, second group has also 36 patients and third one has 37. One person had only one form of trial. The matter of interest was response of patients: total score on Toronto Western Spasmodic Torticollis Rating Scale (TWSTRS), measuring severity, pain, and disability of cervical dystonia, a numerical score ranging from 0 to 87; high scores indicate impairment. It was measured at baseline (week 0) and weeks 2, 4, 8, 12, 16 after treatment began. There is also information about sex, age and hospital id number of every patient.

3.1 Missingness of data

This dataset is not complete. The are 6 missing values at week 2, 3 at week 4, 5 at week 8, 5 at week 12 and 4 at week 16. It seems that missingness is completely at random (MCAR) because there are only two dropouts: one patient missed 3 last treatment sessions and one missed 2 last sessions. The rest of missingness does not show any pattern.

3.2 Descriptive statistics

There were 67 women and 42 men. The youngest person was 26 years old and the oldest was 83. The mean of age of patients equals 55.6 and the median 56. The first quantile is at 46 and the third one at 65.

The lowest value of total score on Toronto Western Spasmodic Torticollis Rating Scale was 6 and the highest at level 71. The mean was at level 41.48 and the median 43. The first quantile of TWSTRS is at level 32.5 and the third quantile is at level 51.

Boxplots of age and TWSTRS are presented at the beginning of next page (Figure 3.1). Information about hospital id number of patients is not going to be used, so it will be ignored.

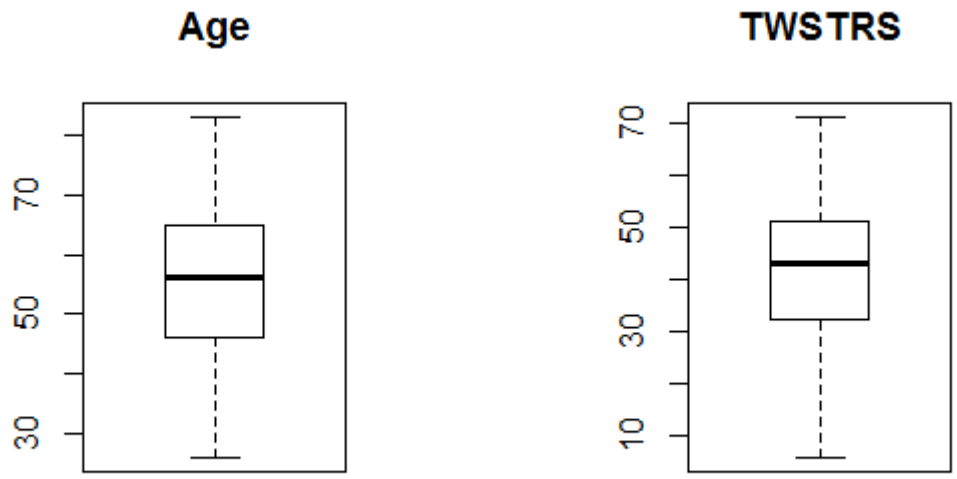

Figure 3.1: Boxplots for age and TWSTRS

4 Data analysis - all patients

In this section we are going to perform profile analysis. At the beginning we should take a look at plot of means of TWSTRS responses over time. The plot is presented below.

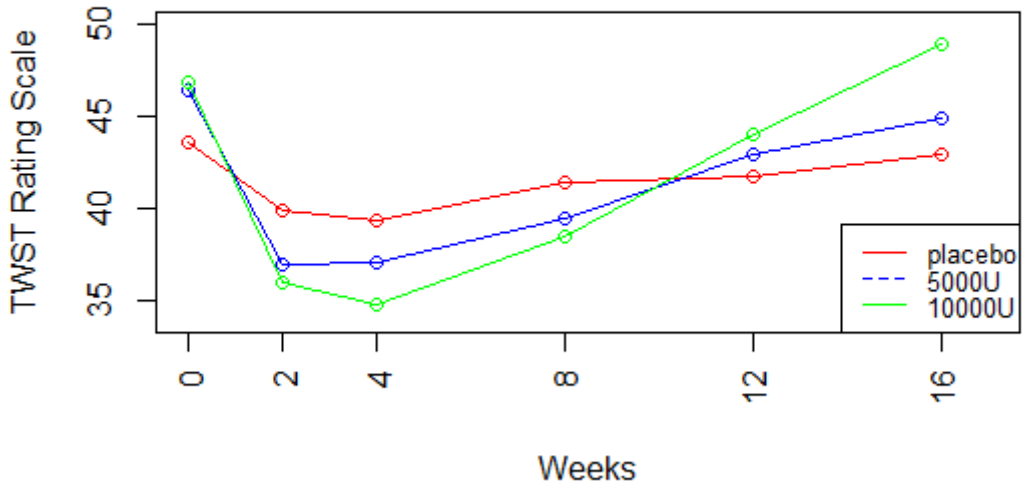

Figure 4.1: Means of TWSTRS over time grouped by treatment

Figure 4.1 shows that at beginning of the treatment the means for all groups were decreasing. Yet after 8th week the mean responses for all treatment started to increase. For group treated with the 10000 units of BtoB the mean response at the end of the study was slightly higher than at the beginning. The spaghetti plot could give more insight of what is happening in the data.

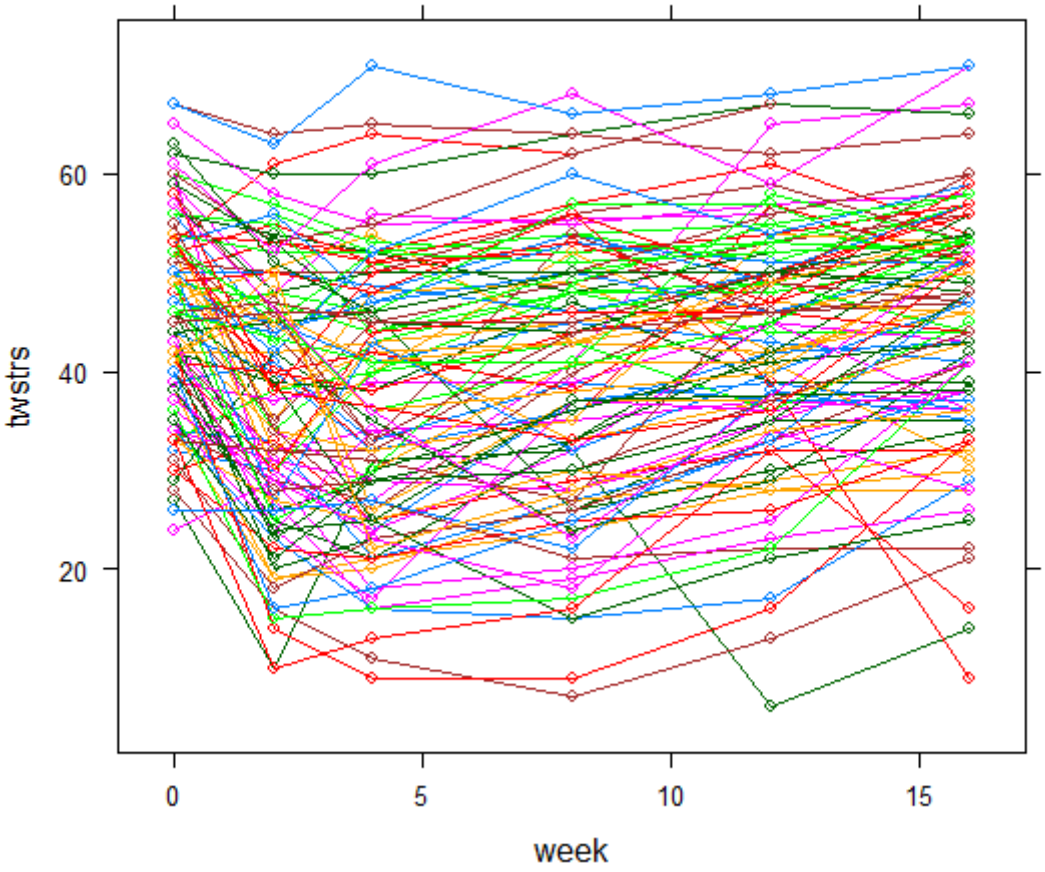

Figure 4.2: Spaghetti plot for all patients.

On Figure 4.2 we can see that people that had very high TWSTRS at the end of study had also high responses on previous measurements. We can also observe that patient who started in the middle of the scale usually ended up also in the middle. Another observation is that variability at the end is significantly higher compared to the beginning of the study.

4.1 Covariance structure

In this subsection we are going to check if we can use a different covariance matrix structure then unstructured. Using a structure that is simpler than unstructured could be more economical in terms of estimation. We are going to test that Compound Symmetry (CS) structure and Autoregressive structure with continuous time (CAR-1) are more fitting to the data than Unstructured (UN). The matrix structures for Unstructured, Compound Symmetry and Autoregressive with continuous time, where occasions of measurement are 0, 2, 4, 8, 12, 16 are presented on the next page.

1. Unstructured:

$$cov(Y_i) = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} & \sigma_{14} & \sigma_{15} & \sigma_{16} \\ \sigma_{12} & \sigma_2^2 & \sigma_{23} & \sigma_{24} & \sigma_{25} & \sigma_{26} \\ \sigma_{13} & \sigma_{23} & \sigma_3^2 & \sigma_{34} & \sigma_{35} & \sigma_{36} \\ \sigma_{14} & \sigma_{24} & \sigma_{34} & \sigma_4^2 & \sigma_{45} & \sigma_{46} \\ \sigma_{15} & \sigma_{25} & \sigma_{35} & \sigma_{45} & \sigma_5^2 & \sigma_{56} \\ \sigma_{16} & \sigma_{26} & \sigma_{36} & \sigma_{46} & \sigma_{56} & \sigma_6^2 \end{bmatrix}$$

2. Compound Symmetry:

$$cov(Y_i) = \sigma^2 \begin{bmatrix} 1 & \rho & \rho & \rho & \rho & \rho \\ \rho & 1 & \rho & \rho & \rho & \rho \\ \rho & \rho & 1 & \rho & \rho & \rho \\ \rho & \rho & \rho & 1 & \rho & \rho \\ \rho & \rho & \rho & \rho & 1 & \rho \\ \rho & \rho & \rho & \rho & \rho & 1 \end{bmatrix}$$

3. Autoregressive with continuous time:

$$cov(Y_i) = \sigma^2 \begin{bmatrix} 1 & \rho^2 & \rho^4 & \rho^8 & \rho^{12} & \rho^{16} \\ \rho^2 & 1 & \rho^2 & \rho^6 & \rho^{10} & \rho^{14} \\ \rho^4 & \rho^2 & 1 & \rho^4 & \rho^8 & \rho^{12} \\ \rho^8 & \rho^6 & \rho^4 & 1 & \rho^4 & \rho^8 \\ \rho^{12} & \rho^{10} & \rho^8 & \rho^4 & 1 & \rho^4 \\ \rho^{16} & \rho^{14} & \rho^{12} & \rho^8 & \rho^4 & 1 \end{bmatrix}$$

To test that certain structure is adequate to the data we perform likelihood ratio tests under REML. We will compare different matrix structures with each other. To do so we will build different covariance matrix structures and then we will use ANOVA to perform tests.

Table 4.1: UN, CS and CAR-1 results

Structure	-2 *REML Log-Likelihood	Number of Cov. Parameters
Compound Symmetry	4344.36	2
Unstructured	4263	15
Autoregressive	4325.1	2

In table 4.1 are values of $2 \cdot REMLLog-Likelihood$ and number of parameters used when CS, UN or CAR-1 are assumed to be the structure of covariance matrix. To compare Unstructured with Compound Symmetry we will test the following hypotheses:

$$H_0: \quad Compound\ Symmetry\ fits\ the\ data,$$
$$H_1: \quad Unstructured\ fits\ the\ data\ better.$$

LRT yields that p-value is equal to 0, so we reject the null hypothesis at $\alpha = 0.05$ and conclude that the assumption of Compound Symmetry covariance structure is not adequate for the data.

Now, we are going to compare UN with CAR-1. We will test the following hypotheses:

$$H_0: \quad CAR-1 \ fits \ the \ data,$$
$$H_1: \quad UN \ fits \ the \ data \ better.$$

LRT yields that p-value is equal to 0, so we reject the null hypothesis at $\alpha = 0.05$ and conclude that the assumption of Autoregressive covariance structure with continuous time is not adequate for the data.

Both CS and CAR-1 should be rejected so there is no point in checking which one is more suitable to the data. It means that for remainder of the project we will use Unstructured form of covariance matrix. Knowing that we can finally write down the full model.

4.2 Full model

At the beginning of study the model is

$$Y_{ij} = \beta_1 + \sum_{k=2}^{18} \beta_k X_{kij} + \varepsilon_{ij},$$

where
$X_{1ij} = 1$ for all measurements
$X_{2ij} = 1$ if the jth measurement was taken at time = 2 weeks, 0 otherwise
$X_{3ij} = 1$ if the jth measurement was taken at time = 4 weeks, 0 otherwise
$X_{4ij} = 1$ if the jth measurement was taken at time = 8 weeks, 0 otherwise
$X_{5ij} = 1$ if the jth measurement was taken at time = 12 weeks, 0 otherwise
$X_{6ij} = 1$ if the jth measurement was taken at time = 16 weeks, 0 otherwise
$X_{7ij} = 1$ if the ith patient was treated with 5000 units of BotB, 0 otherwise
$X_{8ij} = 1$ if the ith patient was treated with 10000 units of BotB, 0 otherwise
$X_{9ij} = 1$ if the ith patient was treated with 5000 units of BotB and its jth measurement is at time = 2 weeks, 0 otherwise
$X_{10ij} = 1$ if the ith patient was treated with 5000 units of BotB and its jth measurement is at time = 4 weeks, 0 otherwise
$X_{1ij} = 1$ if the ith patient was treated with 5000 units of BotB and its jth measurement is at time = 8 weeks, 0 otherwise
$X_{12ij} = 1$ if the ith patient was treated with 5000 units of BotB and its jth measurement is at time = 12 weeks, 0 otherwise
$X_{13ij} = 1$ if the ith patient was treated with 5000 units of BotB and its jth measurement is at time = 16 weeks, 0 otherwise
$X_{14ij} = 1$ if the ith patient was treated with 10000 units of and its jth measurement is at time = 2 weeks, 0 otherwise
$X_{15ij} = 1$ if the ith patient was treated with 10000 units of and its jth measurement is at time = 4 weeks, 0 otherwise
$X_{16ij} = 1$ if the ith patient was treated with 10000 units of and its jth measurement is at time = 8 weeks, 0 otherwise
$X_{17ij} = 1$ if the ith patient was treated with 10000 units of and its jth measurement is at time = 12 weeks, 0 otherwise
$X_{18ij} = 1$ if the ith patient was treated with 10000 units of and its jth measurement is at time = 16 weeks, 0 otherwise.

We should notice that week 0 for time and placebo treatment are the baseline for this model. The βs are the parameters of the model. Now we should try to simplify it by testing if some of the parameters are statistically insignificant. First, we will check the null hypothesis about parallelism. If we fail to reject the null hypothesis about parallelism we will test for main effects.

4.3 Test for parallelism

Test for parallelism checks if there is an interaction between time of measurements and type of treatment. It will test the hypotheses:

$$H_0: \quad \forall_{i\in\{9,10,\ldots 18\}} \quad \beta_i = 0,$$

$$H_1: \quad \exists_{i\in\{9,10,\ldots 18\}} \quad \beta_i \neq 0.$$

ANOVA test of parallelism at significance level $\alpha = 0.05$ result with p-value equal to 0.026. It means that we reject the null hypothesis and conclude that interaction $week \times treat$ is statistically significant. Because of that tests for main effects with null hypotheses

$$H_0^1: \quad \beta_7 = \beta_8 = 0,$$

$$H_0^2: \quad \forall_{i\in\{2,3,4,5,6\}} \quad \beta_i = 0$$

are meaningless and we must consider a full model with interactions.

Now that we have final model we can check if at the end of the treatment course the mean responses for placebo, 5000 units of BotB and 10000 units of BotB were significantly different.

4.4 Post-trial estimated response means comparison

We will start with checking if all three estimated means at 16th week are equal, namely

$$\mu_{placebo} = \mu_{5000U} = \mu_{10000U}.$$

Using the final model described at previous page and fixing it at week 16th we can simplify the equality of $\mu_{placebo}$, μ_{5000U}, μ_{10000U} into terms of betas. By writing

$$\mu_{placebo} = \beta_1 + \beta_6$$

$$\mu_{5000U} = \beta_1 + \beta_6 + \beta_7 + \beta_{13}$$

$$\mu_{10000U} = \beta_1 + \beta_6 + \beta_8 + \beta_{18}$$

one can obtain

$$\beta_7 - \beta_8 + \beta_{13} - \beta_{18} = 0.$$

With this expression we can use T-test to check the significance of difference between all three groups.

$$H_0: \quad \beta_7 - \beta_8 + \beta_{13} - \beta_{18} = 0$$

$$H_1: \quad \beta_7 - \beta_8 + \beta_{13} - \beta_{18} \neq 0$$

T-test yields that p-value is equal to 0.27, so we fail reject the null hypothesis at $\alpha = 0.05$ and conclude that the difference of response means between all groups at week 16th is statistically insignificant.

We can also compare both treatments to placebo separately. We will start by testing that estimated means of response at 16th week are statistically the same for placebo and 5000 units of BotB. To do so, we have to obtain similar expression of betas as before. The hypotheses of that test are

$$H_0: \quad \beta_7 + \beta_{13} = 0$$
$$H_1: \quad \beta_7 + \beta_{13} \neq 0.$$

T-test yields that p-value is equal to 0.43, so we fail reject the null hypothesis at $\alpha = 0.05$ and conclude that the difference of response means between groups of placebo and 5000 units of BotB treatment at week 16th is statistically insignificant.

Now we test that estimated means of response at 16th week are statistically the same for placebo and 10000 units of BotB. The hypotheses of that test are

$$H_0: \quad \beta_8 + \beta_{18} = 0$$
$$H_1: \quad \beta_8 + \beta_{18} \neq 0.$$

T-test yields that p-value is equal to 0.06, so we fail reject the null hypothesis at $\alpha = 0.05$ and conclude that the difference of response means between groups of placebo and 5000 units of BotB treatment at week 16th is statistically insignificant.

4.5 Summary

With T-tests we conclude the profile analysis. We used Unstructured covariance matrix and built a model of mean response over time. Using it we learnt that there is a significant interaction between time and used treatment. It means that not only both time and treatment are significant variables but also they interact with each other. We also learnt that there is no real difference in mean response of pain and disability in all groups at the last week of treatment course.

5 Data analysis - patients with 5000 unit of BotB treatment

In previous pages we focused only on treatment and time. We could not use the information about sex or age of patient because we are supposed to have the same numbers of men and women in each groups of treatment or the same age structure in every group. The only group of treatment where there are similar numbers of men and women is 5000 units of BotB group. In this treatment group are 18 women and 18 men. In group with 10000 units there are 28 women and 9 men and in placebo group there are 21 women and 15 men. We can again perform a profile analysis just for that group and compare the results of treatment when sex is taken into consideration.

Like before, it does seem like missing values are completely at random. The highest number of missing values is 2 and it happened to be on the second week of the medical course. Thus, we will ignore missingness of data.

As previous, we are going to repeat all the steps we did in section 4. We shall start with plot of means of TWSTRS when grouped by sex for all patients treated with 5000 units of BotB.

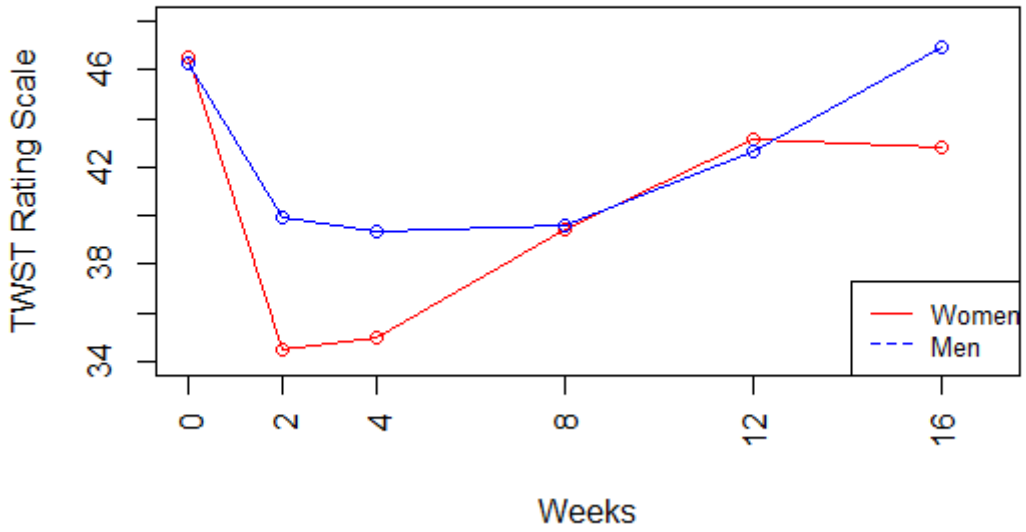

Figure 5.1: Means of TWSTRS over time grouped by sex when treatment was 5000 units of BotB.

Figure 5.1 shows that women were more sensitive than men. The drop for women in TWSTRS at week 2 was noticeable greater in comparison to men. At weeks 8 and 12 mean responses were almost the same. In the end men's mean response was 4 points higher than women's. We are not going to show the spaghetti plot here, but it shows that variability at the beginning is smaller than in next times of measurement.

5.1 Covariance matrix structure

Now we are going to test if there is a better covariance matrix structure than Unstructured. We are going to use the same candidates as we did in pages 4 and 5. Havin all described there we can jump right into results.

Table 5.1: UN, CS and CAR-1 results

Structure	-2 *REML Log-Likelihood	Number of Cov. Parameters
Compound Symmetry	1354.96	2
Unstructured	1323.86	15
Autoregressive	1374.23	2

In table 5.1 are values of 2· REML Log-Likelihood and number of parameters used when CS, UN or CAR-1 are assumed to be the structure of covariance matrix. To compare Unstructured with Compound Symmetry we will test the following hypotheses:

$$H_0: \quad Compound\ Symmetry\ fits\ the\ data,$$
$$H_1: \quad Unstructured\ fits\ the\ data\ better.$$

LRT yields that p-value is equal to 0.003, so we reject the null hypothesis at $\alpha = 0.05$ and conclude that the assumption of Compound Symmetry covariance structure is not adequate for the data.

Now, we are going to compare UN with CAR-1. We will test the following hypotheses:

$$H_0: \quad CAR-1\ fits\ the\ data,$$
$$H_1: \quad UN\ fits\ the\ data\ better.$$

LRT yields that p-value is equal to 0, so we reject the null hypothesis at $\alpha = 0.05$ and conclude that the assumption of Autoregressive covariance structure with continuous time is not adequate for the data.

Both CS and CAR-1 should be rejected so there is no point in checking which one is more suitable to the data. It means that for remainder of the project we will use Unstructured form of covariance matrix. Knowing that we can finally write down the full model for patients with 5000 units of BotB as treatment.

5.2 Full model

At the beginning of study the model is

$$Y_{ij} = \beta_1 + \sum_{k=2}^{12} \beta_k X_{kij} + \varepsilon_{ij},$$

where
$X_{1ij} = 1$ for all measurements
$X_{2ij} = 1$ if the jth measurement was taken at time = 2 weeks, 0 otherwise
$X_{3ij} = 1$ if the jth measurement was taken at time = 4 weeks, 0 otherwise
$X_{4ij} = 1$ if the jth measurement was taken at time = 8 weeks, 0 otherwise
$X_{5ij} = 1$ if the jth measurement was taken at time = 12 weeks, 0 otherwise
$X_{6ij} = 1$ if the jth measurement was taken at time = 16 weeks, 0 otherwise
$X_{7ij} = 1$ if the ith patient was male, 0 otherwise
$X_{8ij} = 1$ if the ith patient was male and its jth measurement is at time = 2 weeks, 0 otherwise
$X_{9ij} = 1$ if the ith patient was male and its jth measurement is at time = 4 weeks, 0 otherwise
$X_{10j} = 1$ if the ith patient was male and its jth measurement is at time = 8 weeks, 0 otherwise
$X_{11ij} = 1$ if the ith patient was male and its jth measurement is at time = 12 weeks, 0 otherwise
$X_{12ij} = 1$ if the ith patient was male and its jth measurement is at time = 16 weeks, 0 otherwise

We should notice that week 0 for time and female sex are the baseline for this model. Now we should try to simplify it by testing if some of the parameters are statistically insignificant. First, we will check the null hypothesis about parallelism.

5.3 Test for parallelism

Test for parallelism checks if there is an interaction between time of measurements and sex. It will test the hypotheses:

$$\begin{aligned} H_0: &\quad \forall_{i\in\{8,9,10,11,12\}} \quad \beta_i = 0, \\ H_1: &\quad \exists_{i\in\{8,9,10,11,12\}} \quad \beta_i \neq 0. \end{aligned}$$

ANOVA test of parallelism at significance level $\alpha = 0.05$ result with p-value equal to 0.169. It means that we fail reject the null hypothesis and conclude that interaction $week \times sex$ is statistically insignificant. Because of that we have to simplify our model and then perform tests for main effects.

5.4 Tests for main effects

After simplification our model is

$$Y_{ij} = \beta_1 + \sum_{k=2}^{7} \beta_k X_{kij} + \varepsilon_{ij},$$

with the same X_{kij} interpretation as before.

No we are going to see if the sex of patient or week of measurement has any impact on mean response of TWSTRS when patient with cervical dystonia is treated with 5000 units of Botulinum neurotoxin type B.

We tests for main effects with hypotheses:

$$\begin{aligned} H_0^1: &\quad \beta_7 = 0, \\ H_1^1: &\quad \beta_7 \neq 0, \end{aligned}$$

and

$$\begin{aligned} H_0^2: &\quad \forall_{i\in\{2,3,4,5,6\}} \quad \beta_i = 0 \\ H_1^2: &\quad \exists_{i\in\{2,3,4,5,6\}} \quad \beta_i \neq 0 \end{aligned}$$

ANOVA test of sex main effect at significance level $\alpha = 0.05$ result with p-value equal to 0.998. It means that we fail reject the null hypothesis and conclude that effect of sex is statistically insignificant.

Now we do the same with time measurements effect.

ANOVA test of week main effect at significance level $\alpha = 0.05$ result with p-value equal to 0. It means that we reject the null hypothesis and conclude that effect of time is statistically significant.

It means we should not take the sex of patient into the consideration but we should take time of measurement.

Test form main effects conclude with another simplification of the model:

$$Y_{ij} = \beta_1 + \sum_{k=2}^{6} \beta_k X_{kij} + \varepsilon_{ij},$$

with the same X_{kij} interpretation as before.

Having those results, there is no point in doing the test for difference in TWSTRS between men and women. Because sex effect is insignificant for TWSTRS there is no significant difference in means between men and women at any timepoint of measurements.

5.5 Summary

In this section we restricted our data to patients with treatment of 5000 units of BotB. Instead of taking treatment as a variable we used sex of patient. Tests about covariance matrix concluded that the best option out of Unstructured, Compound Symmetry and Autoregressive with continuous time was Unstructured. Then we tested the full model about significance of $week \times sex$ interactions. The result was that interaction is insignificant. In order to further simplification of the model we tested about main effects. Tests concluded that main effect of sex is insignificant and main effects of time is significant. That means statistically there is no difference in mean response of TWSTRS between men and women with cervical dystonia when treated with 5000 units of botulinum neurotoxin type B.