



STA490: Statistical Practice in Clinical Research

Prof. Ulrike Held

EBPI, Department of Biostatistics

Tissue Bridges Cut-Off Value in Spinal Cord Injury Patients

Analysis for Patrick Freund, Lynn Farner

Supervision by Torsten Hothorn

Mike Krähenbühl (mikeeric.kraehenbuehl@uzh.ch)

Version of December 23, 2024

1 Abstract

We investigated the predictive value of preserved tissue bridges for recovery after a cervical spinal cord injury. This analysis was based on the NISCI study, a randomized, placebo-controlled, multicentre, phase 2b clinical trial with 78 patients in the NG101 (intervention) group and 48 patients in the placebo group. The relationship of baseline tissue bridges with recovery, measured by upper extremity motor score (UEMS), was modelled with a generalized additive model (GAM) on both treatment groups separately. The analysis showed that the amount of preserved midsagittal tissue bridges had a linear positive relationship with the 6-month recovery (Δ -UEMS) in both treatment groups. The parasagittal tissue bridges showed also a positive relationship with recovery. However, in the treatment group the Δ -UEMS did not further improve for parasagittal tissue bridge values higher than 2 mm. Receiving the NG101 treatment was beneficial for patients with parasagittal tissue bridge values lower than 3.5 mm, whereas for higher parasagittal tissue bridge values, the placebo group showed a better 6-month recovery. Hence, in terms of treatment effect, a cut-off value of 3.5 mm in preserved parasagittal tissue bridges could, for example, be used as an inclusion criterion in a follow-up study.

2 Introduction

Traumatic spinal cord injury (SCI) is a severe condition resulting for example from falls and road traffic injuries. Bundles of nerves and nerve fibers are damaged and transmission of nerve signals with the brain is suppressed (NINDS, 2024). SCI can lead to a loss of motor and/or sensory function below the injury site and decrease quality of life (WHO, 2024). The prognosis of recovery is crucial for the patient and the treating physician.

Biomarkers that serve for prognosis are needed. Studies have shown that the amount of preserved tissue bridges around the lesion is an important factor to characterizing and predicting the recovery after a spinal cord injury. Tissue bridges function as preserved neuronal pathways that restore the connection between

the brain and the body. It has been observed that the amount of tissue bridges is positively correlated with the recovery (Huber et al. (2017); Pfyffer et al. (2019)) and that certain cut-off values are prognostic for an improved recovery after 3 months and 12 months (Pfyffer et al., 2024).

Another biomarker that showed some prognostic power is the lesion area. It can be seen as the counterpart to the tissue bridges since it represents the damaged tissue. The lesion area is negatively correlated with the recovery (Huber et al., 2017).

The aim of this analysis is to further investigate the relationship between clinical biomarkers such as preserved midsagittal/parasagittal tissue bridges and lesion volume with the recovery after a spinal cord injury. Depending on the form of this relationship, cut-off values could be determined that help to separate patients into likely and unlikely recovery. Assigning patients to subgroups depending on their predicted outcomes is crucial. It helps improving the planning, individualisation and evaluation of therapies. By segmenting patients in clinical studies into more homogeneous subgroups according to their distinct predicted recovery profile, allows to separately assess the treatment success in the respective subgroups. Finally, a cut-off value could also be used as inclusion criterion in a clinical trial.

3 Research Questions

1. What is the relationship between biomarkers such as preserved midsagittal/parasagittal tissue bridges or lesion volume after a spinal cord injury and the recovery over time measured by the upper extremity motor score?
2. Can a cut-off value on the range of the biomarker be determined that helps to separate patients into poor and good expected recovery?
3. Which biomarker (midsagittal tissue bridges, parasagittal tissue bridges or lesion volume) can be used best to determine a cut-off value?

4 Methods

Study Design

Type of Study

This is an exploratory post-hoc analysis based on the NISCI study (Weidner et al., 2025). Therefore, we will only briefly explain the methods used in the NISCI study, but emphasise those that are relevant to answering the research questions stated in Section 3. The methods of the NISCI study are described in detail in the main study report.

NISCI Study

The NISCI study was a randomised, placebo-controlled, multicentre, phase 2b, clinical trial where the effect of the treatment with an antibody was investigated. The antibody is called NG101 and it binds to the protein Nogo-A. Nogo-A is a protein that inhibits the growth of nerve fibers. The idea is that by targeting Nogo-A, the nerve fibers can regrow better and the recovery is improved.

The study was conducted in specialized SCI centres in Germany, Switzerland, Spain and the Czech Republic. Between 2019 and 2022, 126 patients with acute traumatic cervical spinal cord injury (neurological level of injury C1-C8) were recruited and randomized to treatment (NG101) and placebo in a ratio 2:1.

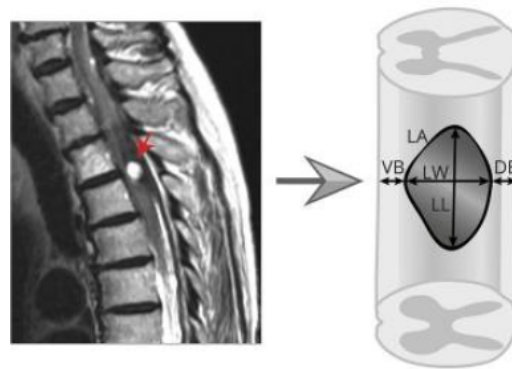


Figure 1: Typical T2W midsagittal slice (left) and schematic lesion segmentation (right) with the quantitative MRI measures analyzed (lesion area [LA], lesion length [LL], lesion width [LW], ventral midsagittal tissue bridges [VB], and dorsal midsagittal tissue bridges [DB]) (Pfyffer et al., 2019).

The primary outcome was the change in upper extremity motor score (UEMS, according to ISNCSCI) in 6 months. The UEMS is a measure that assesses the motor function in multiple muscle groups in the arms and it can take values between 0 and 50 in increments of 1. The score was measured at baseline and after approximately 1, 3 and 6 months. The statistical analysis was mainly conducted with linear mixed-effects models. The overall primary outcome was not significant, but some subgroups showed promising results.

Data collection

Patients underwent MRI scanning at screening. The protocol included T1-weighted, sagittal T2-weighted, and axial T2-weighted anatomic scans of the cervical spinal cord, centered at lesion level. From the scans, the midsagittal tissue bridges, parasagittal tissue bridges and the lesion volume were determined. The amount of preserved midsagittal tissue bridges (measured in mm) were calculated as the sum of the ventral and dorsal bridges that were visible on the midsagittal MRI slice. Figure 1 illustrates the schematic lesion segmentation. Parasagittal tissue bridges were calculated similarly to midsagittal bridges. In this case, however, tissue bridges were also measured in the parasagittal MRI slices. The measurements from all slices where the lesion was visible were added together. To adjust for differences in cord size, this total was divided by the number of slices showing the lesion, giving the average amount of preserved tissue bridges in millimeters. The lesion volume (mm^3) was also determined by MRI screening.

The primary outcome UEMS was measured at baseline and 30, 84 and 168 days post baseline. The time variable in the analysis indicates the measurement timepoint of UEMS in days after start of medication. For example, time=2 means that the UEMS was measured 2 days after the patient received the first dose. The UEMS, assessed according to ISNCSCI, ranges from 0 to 50 in increments of 1. A higher UEMS indicates better motor functions.

Balgrist/Murnau

Additional data from SCI patients including baseline midsagittal tissue bridge values is available from the two centres Balgrist and Murnau. The 157 patients were recruited between 2002-2019 and followed up for up to 12 months (for some even longer). Based on this data, Pfyffer et al. (2024) showed a positive association between the amount of tissue bridges and the recovery in terms of UEMS in a retrospective study. Our results of these additional SCI patients from Balgrist/Murnau will be provided after the results for the NISCI patients in Section 5.

Statistical Analysis

Data Preparation

Most data preparation, such as standardization of time, was already done for the NISCI study. The time variable is standardized so that 0 corresponds to the day of medication start and 1 corresponds to 6 month follow-up (168 days after first medication). The primary dataset contained all variables except the midsagittal tissue bridges and the lesion volume. Those measures were provided in two separate excel files. We joined the baseline measures of the biomarkers with the patient id and added them to the primary dataset.

For the patients from Balgrist/Murnau, no medication timepoint is available. To make the results as comparable as possible to the NSICI data, the time variable is standardized such that 0 corresponds to the average day of first medication in the NISCI study. Therefore, for the Balgrist/Murnau dataset the standardized time variable is defined as:

$$\text{time}_{\text{Balgrist/Murnau}} = \frac{(\text{days since injury}) - (\text{NISCI avg. day of first medication})}{168}$$

Missing Values

There are two types of missing values that are relevant in the analysis of the relationship between baseline biomarkers and the recovery. First, some patients were lost to follow-up for various reasons (for example left the country or no longer wanted to continue). Therefore, we do not have all 4 measurements (baseline and 3 follow-up) from all patients. Second, baseline biomarker values are missing for some patients. The reasons for the missing values will be provided in Section 5. In both cases missing values are not imputed. The primary aim of this exploratory analysis is to investigate the relationship between biomarkers and recovery. No significance testing will be performed. We do not want to introduce additional uncertainty by imputing. We are of the opinion that the lack of strongly related variables with the biomarkers would make imputation of missing values unreliable. The missing biomarker values are therefore handled by excluding the respective patients from the analysis, hence a complete case analysis is performed.

Descriptive Statistics and Simple Methods

The following analysis is conducted for each biomarker individually. Patients are divided into four subgroups according to the quartile of the biomarker range. Then, the UEMS over time is plotted by treatment group and quartile of the biomarker. This should give a first impression of the recovery profiles conditional on the treatment group and different ranges of the biomarker value. Furthermore, boxplots show the distribution of the biomarker per URP-CTREE node. URP-CTREE (Unbiased Recursive Partitioning regression with Conditional Inference Trees) is an algorithm that uses early neurologic impairment to predict the expected motor recovery. This approach was applied during enrollment and patients were stratified into homogeneous subgroups based on the predicted UEMS at 168 days post baseline. Evaniew et al. (2019) validated this approach in an independent study.

The distributions of the biomarkers against the baseline UEMS are visualized by scatter plots. This should uncover potential surprising baseline distributions in the treatment groups.

Next, a generalized additive model (GAM), as will be described in the next Section 4, is fitted on both treatment groups separately. The resulting smooth functions of the predictors are presented, showing how the relationships are modelled. By using different values on the range of the respective biomarker, predictions for the recovery over time are made with the GAM and added to a plot with the raw patient data. Each line in the plot represents a hypothetical patient with a different baseline value for the biomarker. This should give an impression of how the recovery changes over time depending on the value of the biomarker. For

each predicted recovery trajectory, the difference between the UEMS at 6-month follow-up and baseline is calculated (Δ -UEMS). This difference is then plotted against the baseline value of the biomarker and should give an idea of how the predicted 6-month recovery changes depending on the value of the biomarker. If the relationship between the biomarker value at baseline and the 6-month recovery is linear, the Δ -UEMS should be represented as a straight line.

The research question of finding a cut-off value depends on the relationship of Δ -UEMS and the biomarker. In terms of prognostic value, only if there was a non-linear relationship where the Δ -UEMS is higher or lower over a certain range of the biomarker, a cut-off value would be meaningful. In terms of predictive value, crossing curves between the treatment groups would indicate that the treatment was superior only for patients falling into a certain range of the biomarker values.

Generalized Additive Model

Before we consider whether a cut-off value can be determined, we need to investigate the relationship between the tissue bridges and the recovery. Since we want to make as few assumptions as possible, a model is needed that allows for great flexibility. A generalized additive model (GAM) can capture complex, non-linear relationships between predictors and response. It extends the generalized linear model (GLM) by replacing the linear combination of the predictors $\sum \beta_j X_j$ by the sum of smooth functions of the predictors $\sum f_j(X_j)$ (Hastie and Tibshirani, 1986).

A simpler model such as a linear mixed-effects model would require to make the assumption that the relationship between each predictor and the response variable is linear. For illustration; assuming we model the problem by a linear mixed-effects model, a single coefficient for the tissue bridges would be estimated. This coefficient would represent the average effect of the biomarker on the recovery. But this is not the aim of this analysis. Instead, we want to investigate whether the relationship of the biomarker and the response (UEMS) changes depending on the value of the biomarker. A GAM models the problem by fitting a smooth function of the biomarker. This smooth function can have any shape.

Hence, in this analysis we use a GAM to model the recovery. Equation 1 shows the formula of the model that is applied. The model consists of an overall intercept, the main effects of baseline biomarker and time, and additionally an interaction effect for biomarker and time. This interaction term is necessary if the shape of recovery over time is different for varying biomarker values. The patient-specific random effects capture correlations between repeated measurements.

$$\text{UEMS}_{ij} = \beta_0 + f_1(\text{biomarker}_i) + f_2(\text{time}_{ij}) + f_3(\text{biomarker}_i, \text{time}_{ij}) + \alpha_{id} + \beta_{id} \times \text{time}_{ij} + \epsilon_{ij} \quad (1)$$

- UEMS_{ij} is the upper extremity motor score of patient i at time j . It can take values between 0 and 50.
- β_0 is the overall intercept.
- $f_1(\text{biomarker}_i)$ is the smooth function of the biomarker at baseline. It shows the effect on UEMS depending on the value of the biomarker.
- $f_2(\text{time}_{ij})$ is the smooth function of time. It shows the effect on the UEMS over time. The time variable is standardized so that time = 0 corresponds to baseline (day of first medication) and time = 1 corresponds to 6-month follow-up (day 168 after first medication).
- $f_3(\text{biomarker}_i, \text{time}_{ij})$ is the interaction term. It shows how the effect of the biomarker on the UEMS changes over time. This interaction is the effect that is not already included in the main effects of biomarker and time.

- $\alpha_{id} + \beta_{id} \times \text{time}_{ij}$ are the patient-specific random effects. They account for the fact that each patient might have an individual intercept (baseline UEMS value) or trajectory of recovery over time regardless of the value of the predictors used in the model. Without including these random effects, the model would assume that all patients with the same value of the biomarker (e.g. midsagittal tissue bridges) at a given timepoint would have the same baseline UEMS and identical recovery trajectory over time. This would mean that no patient-specific deviations would be allowed and therefore could lead to a biased estimation of the effect on UEMS.
- $\epsilon_{ij} \sim \mathcal{N}(0, \sigma^2)$ is the error term.

Restricted maximum likelihood (REML) is used as fitting algorithm due to its resistance to occasional severe over-fitting (Wood, 2006, p. 267). The smooth functions are approximated using penalized regression splines.

In contrast to random effects in linear mixed-effects models, where explicit parameters are estimated for each individual, the random intercepts and slopes in this GAM are specified as special cases of smooths. Same as the other terms, they are approximated using penalized regression splines. Additionally, the patient-specific slopes can be non-linear.

To avoid making assumptions about the treatment effect of NG101 in any way on the relationship between the biomarkers and the recovery, the model is fitted separately on the placebo and NG101 group.

Implementation

All analyses were performed in the R programming language (R Core Team, 2024) using base packages and the following analysis-specific packages: `mgcv` to fit generalized additive models with `gam()` (Wood, 2011), `lattice` to use `xyplot()` for plotting (Sarkar, 2008).

5 Results

Table 1: Baseline characteristics by treatment group in the NISCI study. The days to first medication as well as the baseline measurements of the UEMS and the three biomarkers are presented as mean (sd).

	Overall	Placebo	NG101
Patients	126	48	78
Time from injury to 1st dose (days)	22.56 (5.19)	23.81 (4.62)	21.78 (5.40)
UEMS	16.39 (9.21)	17.08 (7.94)	15.96 (9.93)
Midsagittal tissue bridges (mm)	1.37 (1.58)	1.40 (1.31)	1.35 (1.75)
Parasagittal tissue bridges (mm)	1.73 (1.45)	1.99 (1.38)	1.56 (1.48)
Lesion volume (mm ³)	585.27 (704.80)	399.28 (234.40)	712.72 (874.94)

NISCI Study and Neuroimaging Substudy

126 patients with acute traumatic cervical spinal cord injury (neurological level of injury C1-C8) were included in the NISCI study. 78 were allocated to the NG101 group and 48 received the placebo. The patients were recruited in 4 countries. The mean age was 46.2 years (SD = 16.7). The percentage of males was 84.9%.

103 out of 126 patients underwent MRI imaging at screening in an exploratory neuroimaging substudy. All 126 patients would theoretically have been included in the substudy. However, some did not want to

participate, were not stable enough, or the MRI machines were out of order. The screening was carried out on average 17.2 days (SD = 7.1) after the injury occurred. Some scans were invalid due to metal or motion artifacts and poor image quality. The midsagittal tissue bridges were calculated based on the midsagittal MRI slice and the parasagittal tissue bridges were calculated based on the parasagittal MRI slices. Table 1 shows the baseline measurements in the treatment groups.

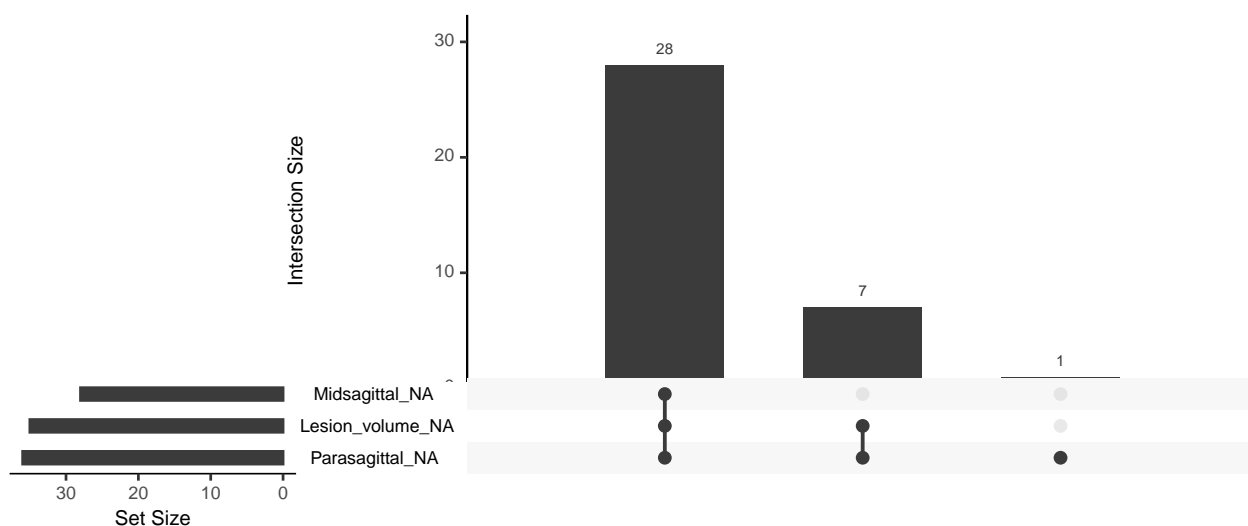


Figure 2: Visualization of the pattern of missing variables in the NISCI dataset.

Figure 2 shows the combinations of missing values of the three biomarkers. It is important to note that only invalid or not recorded values are considered missing values. A biomarker value of 0 is a valid measurement and therefore included in the analysis. For 90 out of 103 patients, all 3 biomarker values were recorded. 8 patients have a missing value for the parasagittal tissue bridges, but a value for midsagittal tissue bridges is recorded. Such a case can happen, for example, when the midsagittal MRI slice is valid, but one or multiple parasagittal slices are invalid because of artifacts that are not visible in the midsagittal slice.

Out of the 126 patients, 111 patients have all 4 measurements for the UEMS recorded. The other patients either discontinued after the first dose or did not receive the full dosing as per protocol. The upper extremity motor score was first measured on average 1.2 days before receiving the first dose. Patients received the first dose on average 22.6 days after the injury.

Figure 3 shows the distributions of the biomarkers at baseline. With regard to the lesion volume, 6 patients are considered outliers and were excluded from the further analysis. All 6 patients were part of the NG101 group. The statistical models would be strongly influenced by these outliers.

Relationship of Biomarkers with Recovery

In this section, the results are shown for each biomarker separately (midsagittal tissue bridges, parasagittal tissue bridges, lesion volume). The results are structured in the following way: First, the raw data is presented for each treatment group and corresponding quartile of the biomarker range. Second, boxplots show the distribution of the biomarker value per URP-CTREE node. Third, the baseline distribution of the biomarker values against the UEMS in the treatment groups is shown. Fourth, the relationship of the biomarker with the recovery is analyzed by the smooth functions that have been fitted by the GAM. It is important to note that the main effects for the biomarker and the time have to be interpreted together with the interaction effect of the two variables. Only looking at the main effects individually does not explain the full

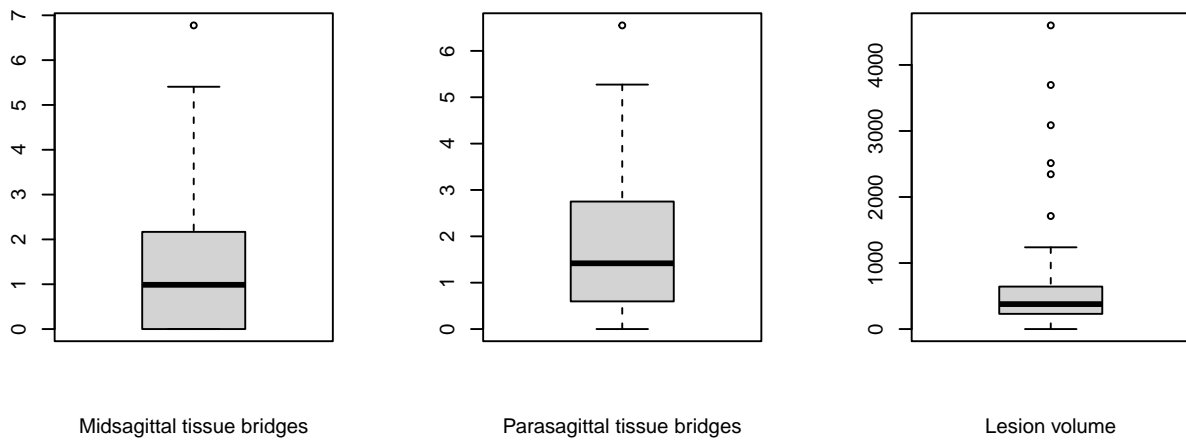


Figure 3: Boxplots of the midsagittal and parasagittal tissue bridges (mm) and the lesion volume (mm^3) at baseline. Concerning the lesion volume, 6 patients are considered outliers and were excluded from the analysis. They are clinically correct observations, but cause problems when fitting the model.

relationship. Fifth, the fitted models are visualized by adding predictions to the raw data plots. Finally, the differences in the predicted UEMS scores between baseline and 6-month follow-up (Δ -UEMS) are calculated for different values of the biomarker.

Midsagittal Tissue Bridges

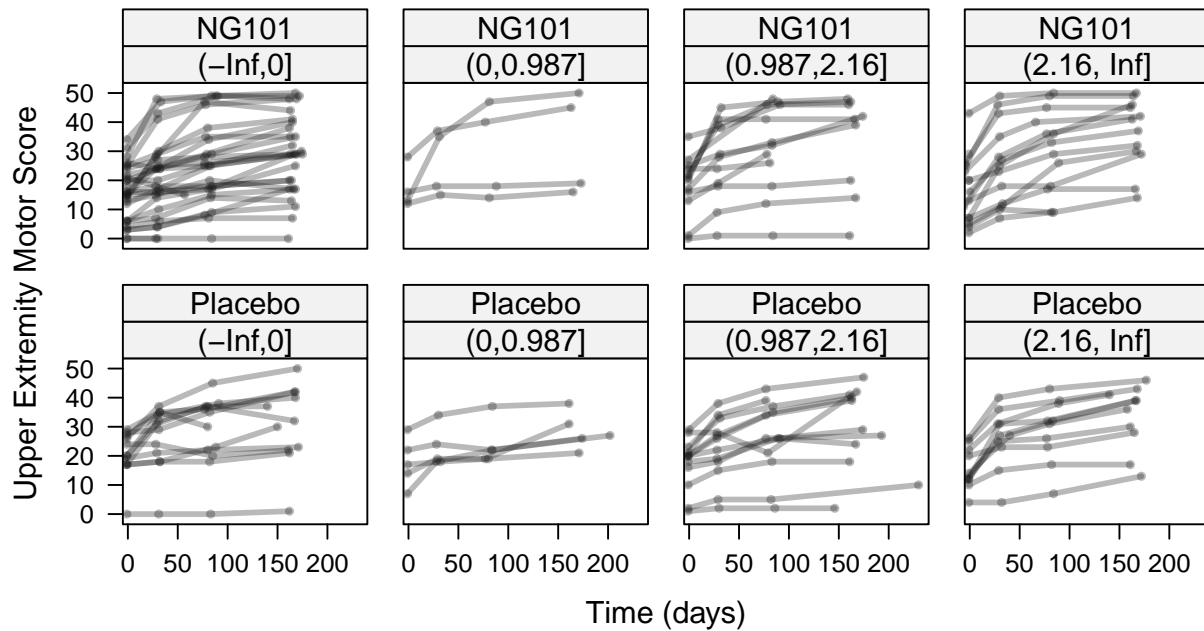


Figure 4: Patient data for each treatment group, categorized into four quartiles based on the overall distribution of midsagittal tissue bridges.

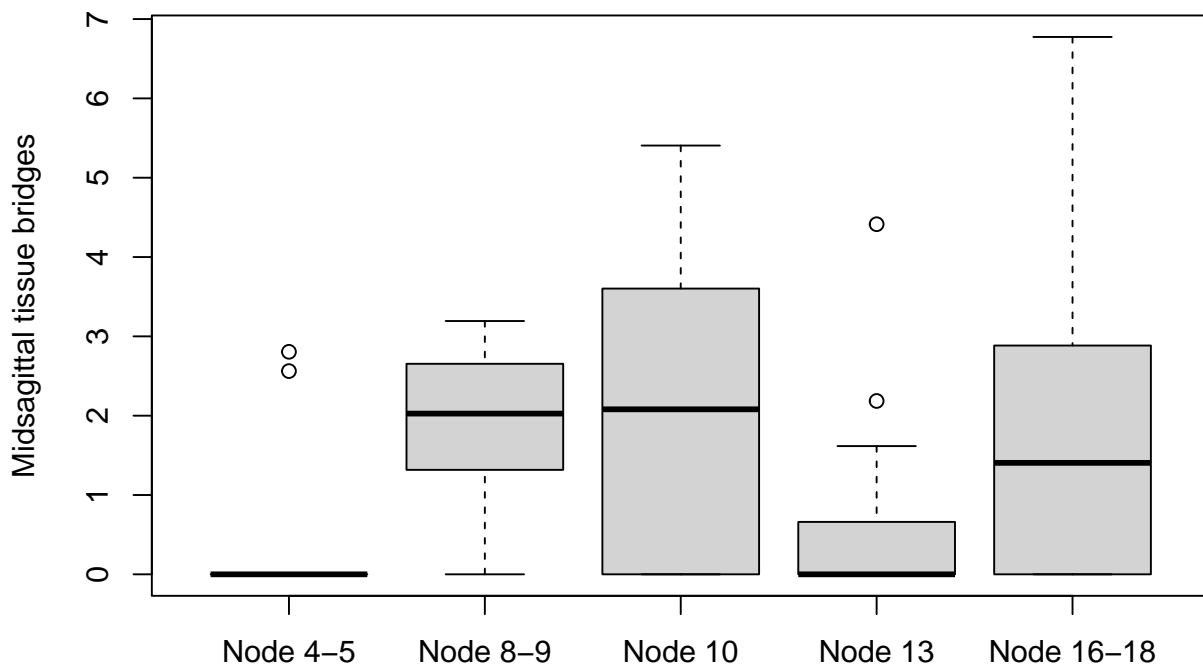


Figure 5: Boxplots of the midsagittal tissue bridges (in mm) by URP-CTREE node.

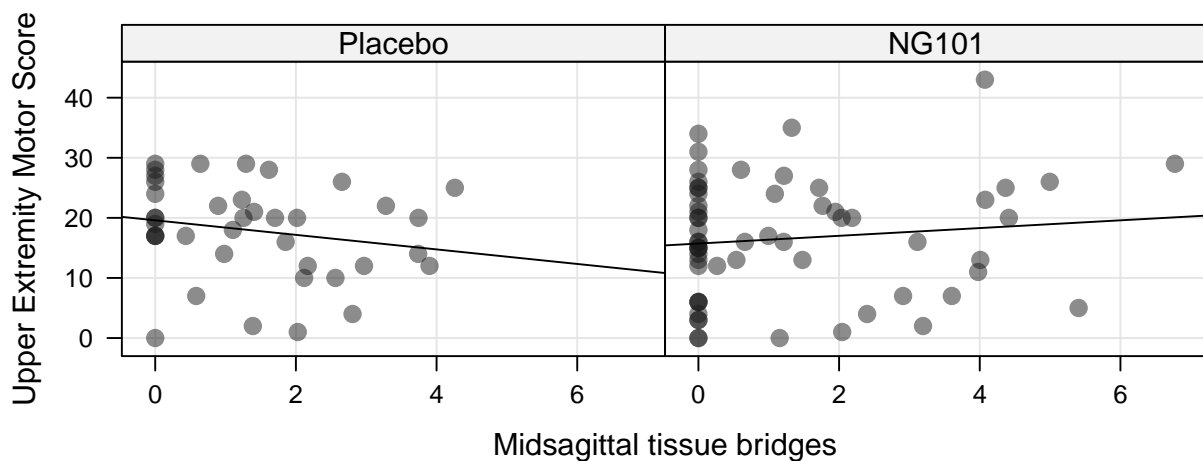


Figure 6: Scatterplot of the midsagittal tissue bridges (in mm) against the baseline Upper Extremity Motor Score in the two treatment groups.

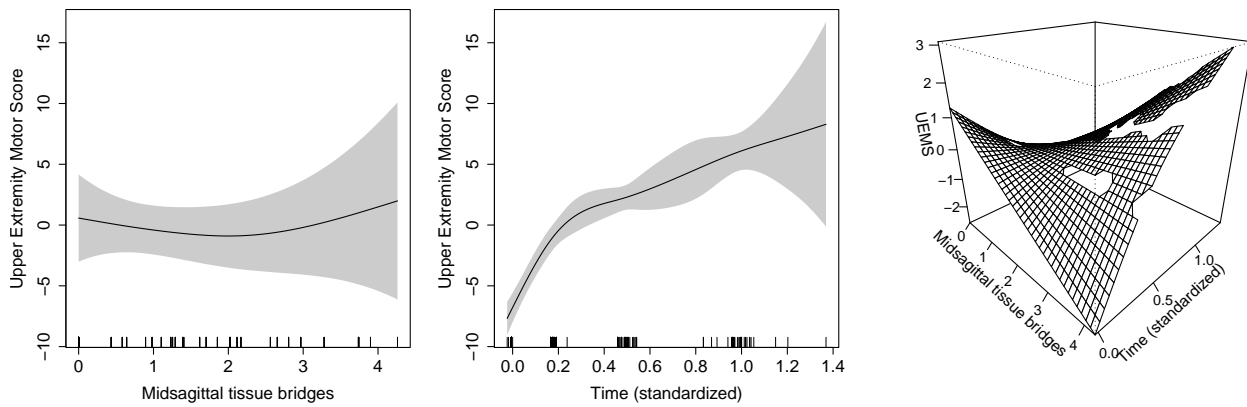


Figure 7: Smooth functions fitted by the GAM for the placebo group.

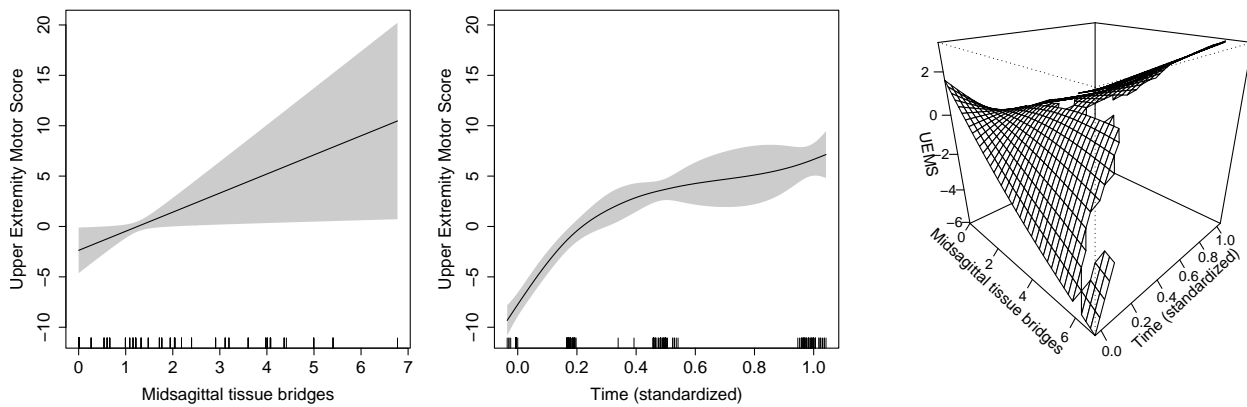


Figure 8: Smooth functions fitted by the GAM for the NG101 group.

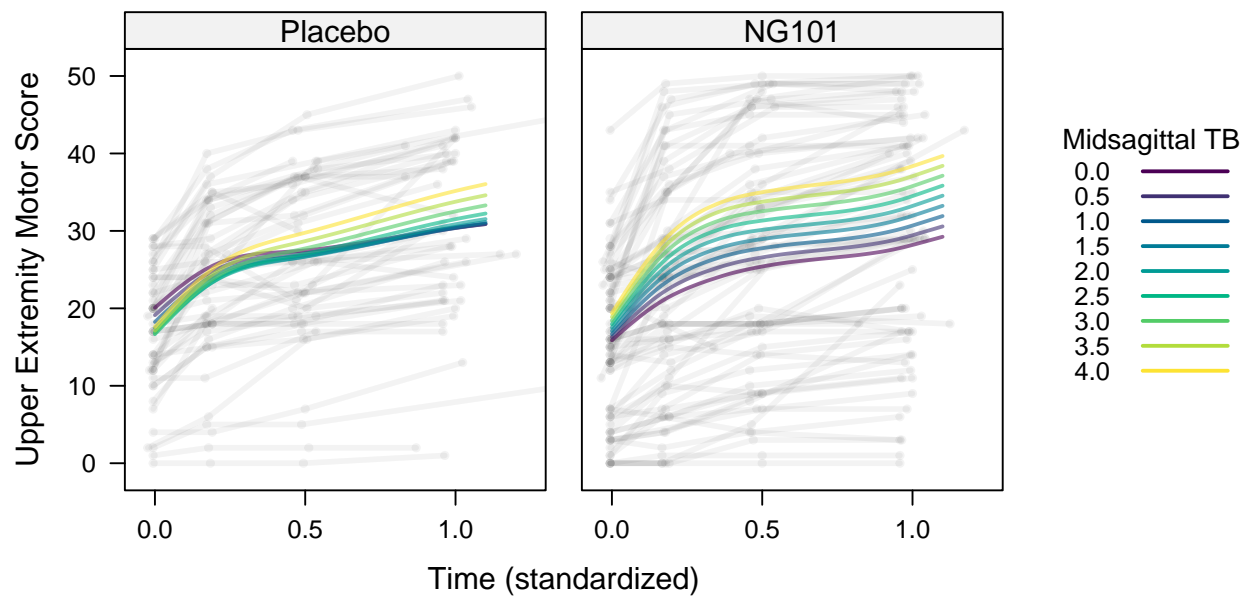


Figure 9: Predictions made with different values for midsagittal tissue bridges (in mm). For each treatment group the respective GAM was fitted separately. The only parameter that changes for the different predictions is the amount of midsagittal tissue bridges.

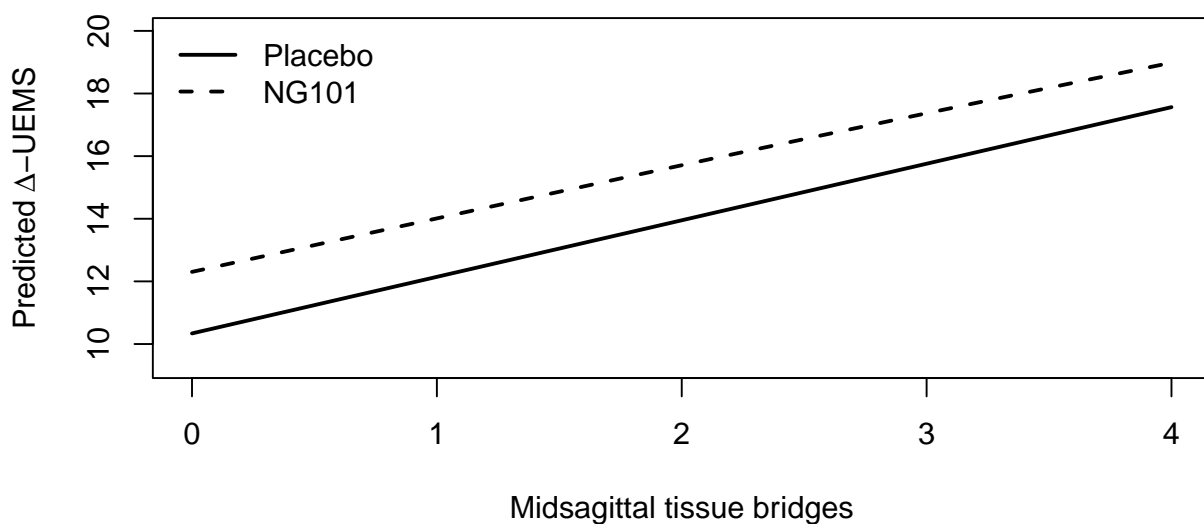


Figure 10: Differences of the predicted 6-month follow-up UEMS scores and the predicted baseline scores for different values of midsagittal tissue bridges (in mm).

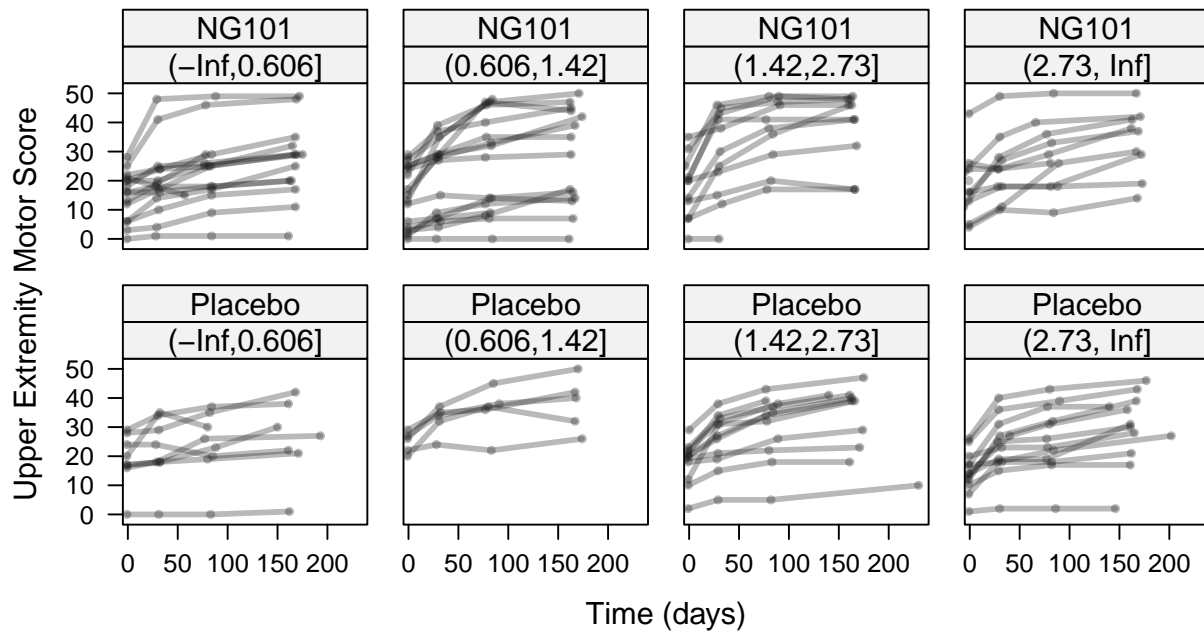
Parasagittal Tissue Bridges

Figure 11: Patient data for each treatment group, categorized into four quartiles based on the overall distribution of parasagittal tissue bridges.

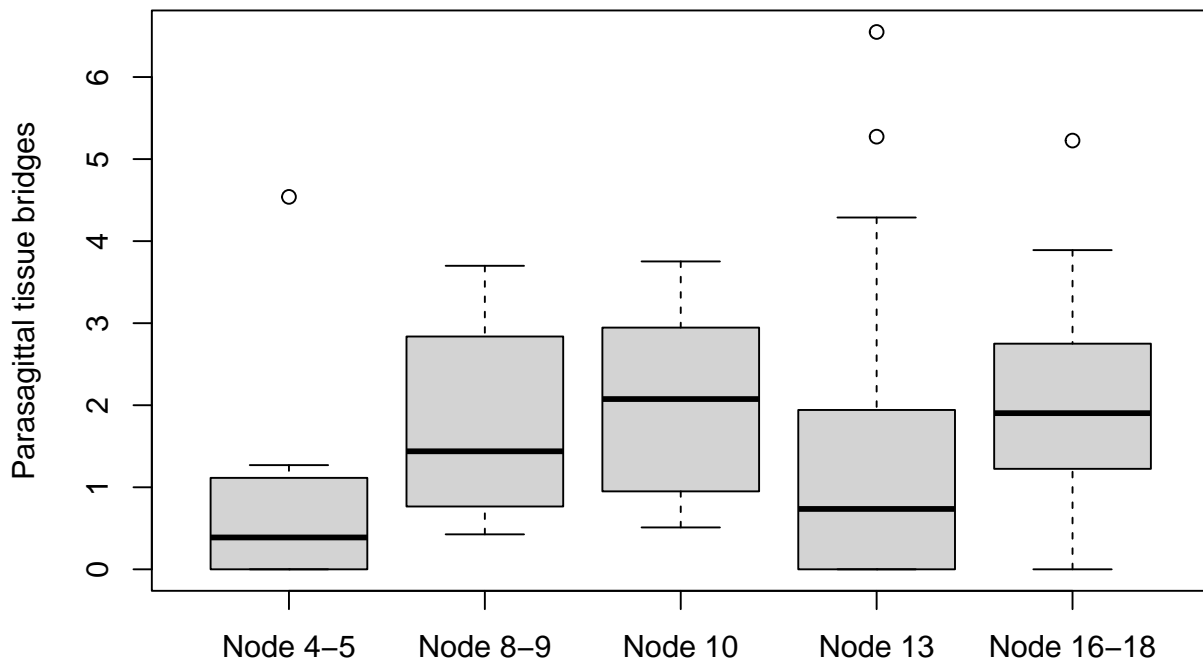


Figure 12: Boxplots of the parasagittal tissue bridges (in mm) by URP-CTREE node.

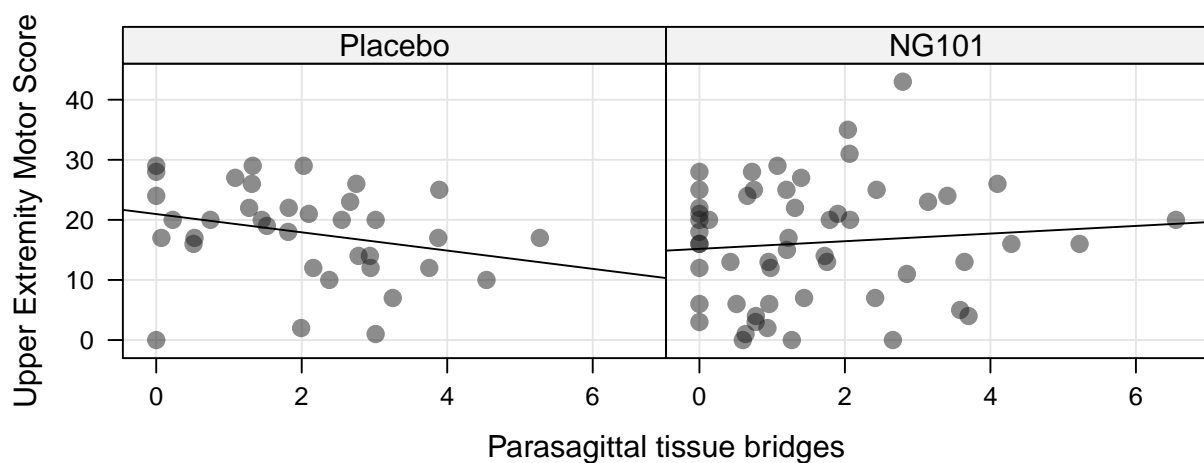


Figure 13: Scatterplot of the parasagittal tissue bridges (in mm) against the baseline Upper Extremity Motor Score in the two treatment groups.

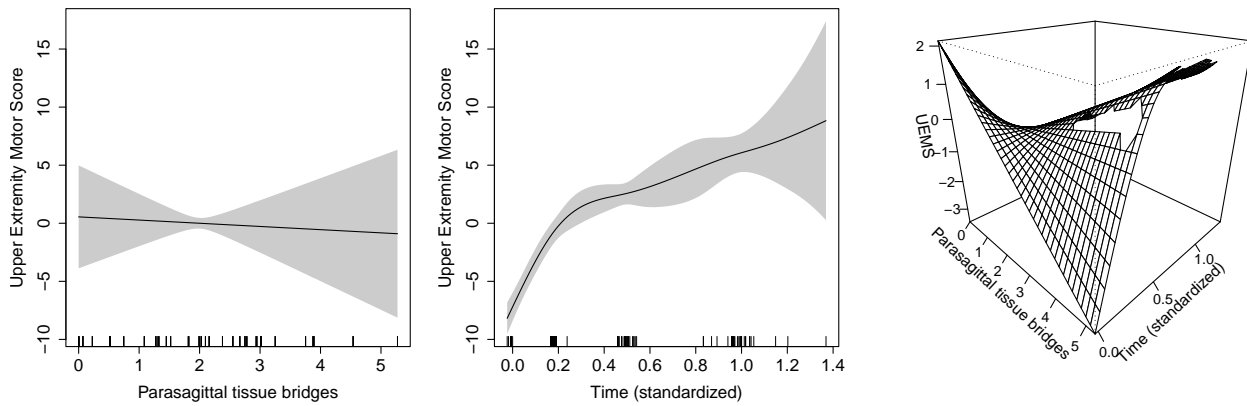


Figure 14: Smooth functions fitted by the GAM for the placebo group.

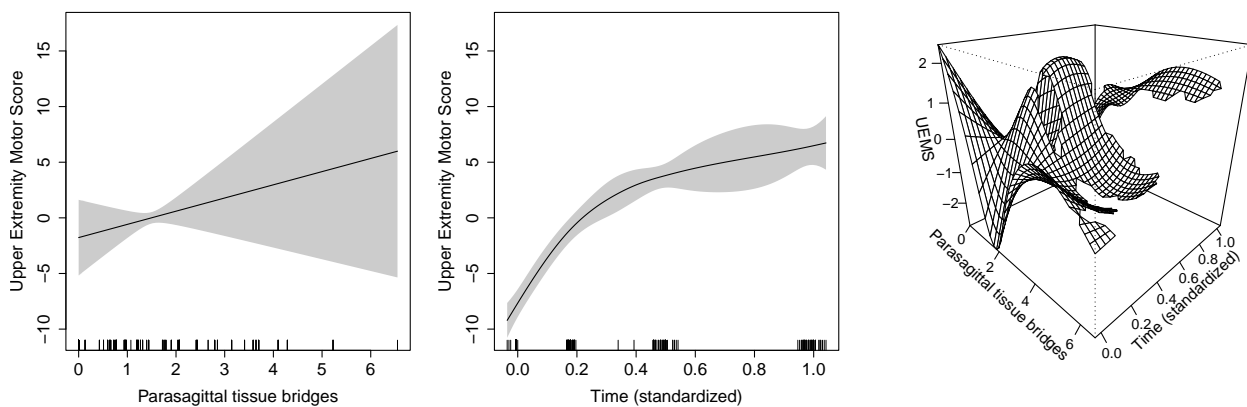


Figure 15: Smooth functions fitted by the GAM for the NG101 group.

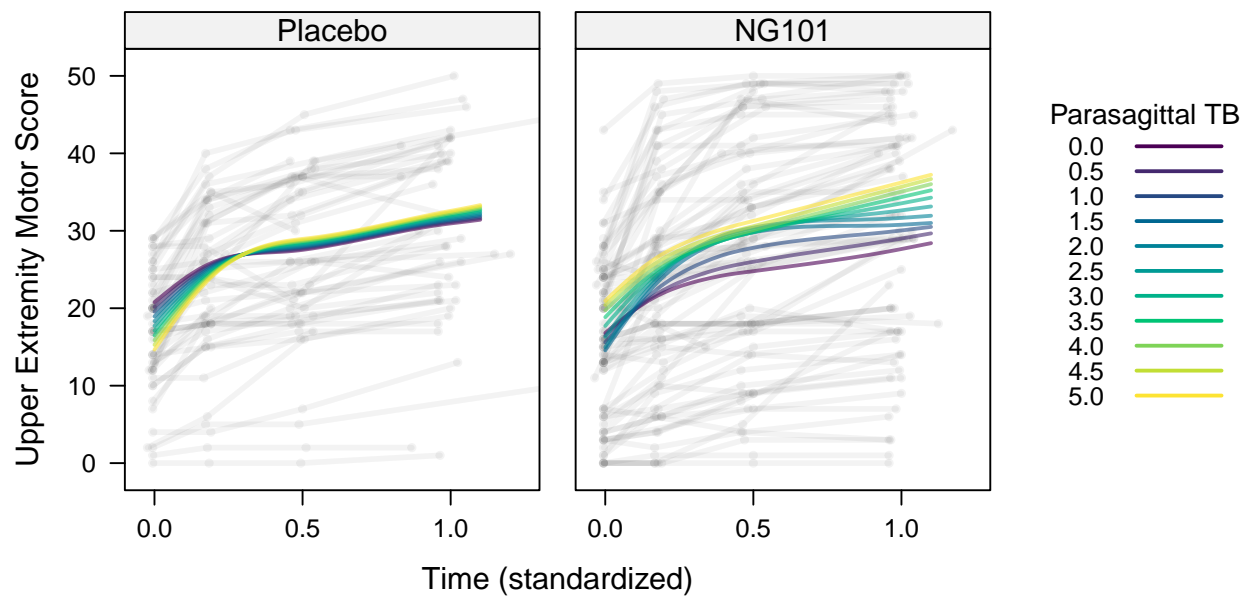


Figure 16: Predictions made with different values for parasagittal tissue bridges (in mm). For each treatment group the respective GAM was fitted separately. The only parameter that changes for the different predictions is the amount of parasagittal tissue bridges.

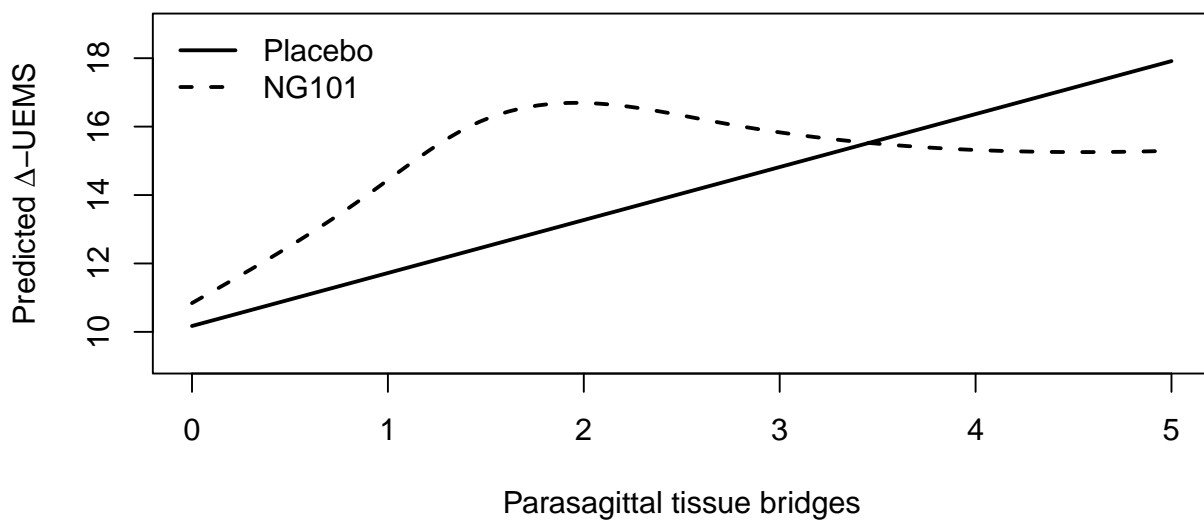


Figure 17: Differences of the predicted 6-month follow-up UEMS scores and the predicted base-line scores for different values of parasagittal tissue bridges (in mm).

Lesion Volume

The 6 patients with lesion volume > 1200 were excluded from the analysis as described earlier.

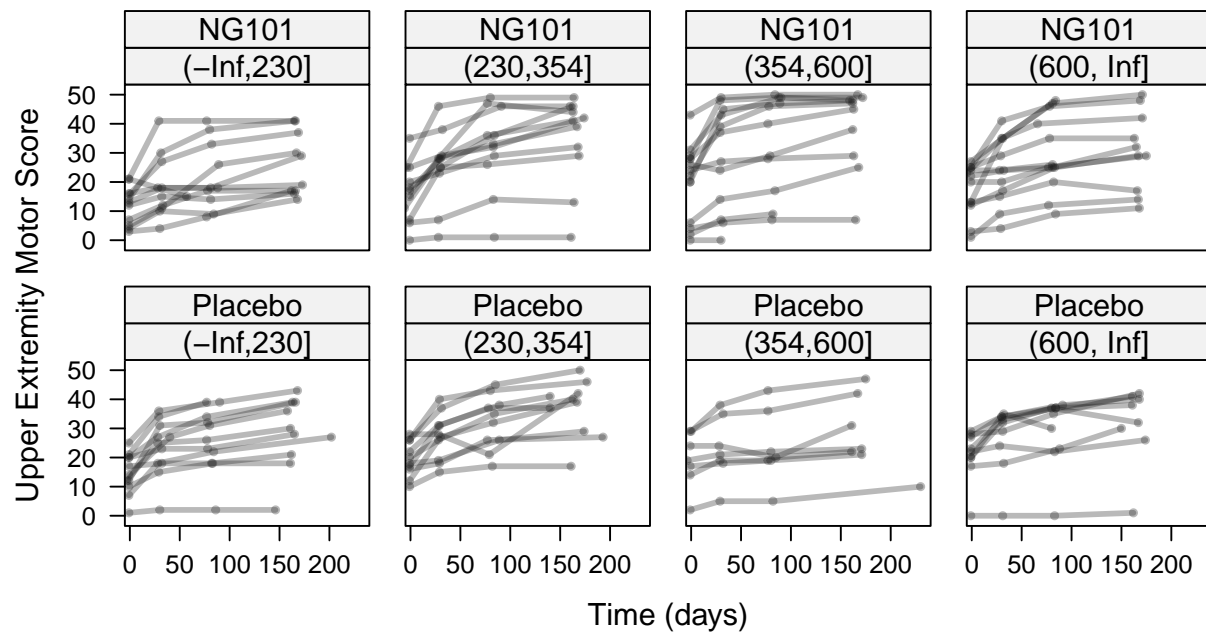


Figure 18: Patient data for each treatment group, categorized into four quartiles based on the overall distribution of lesion volume.

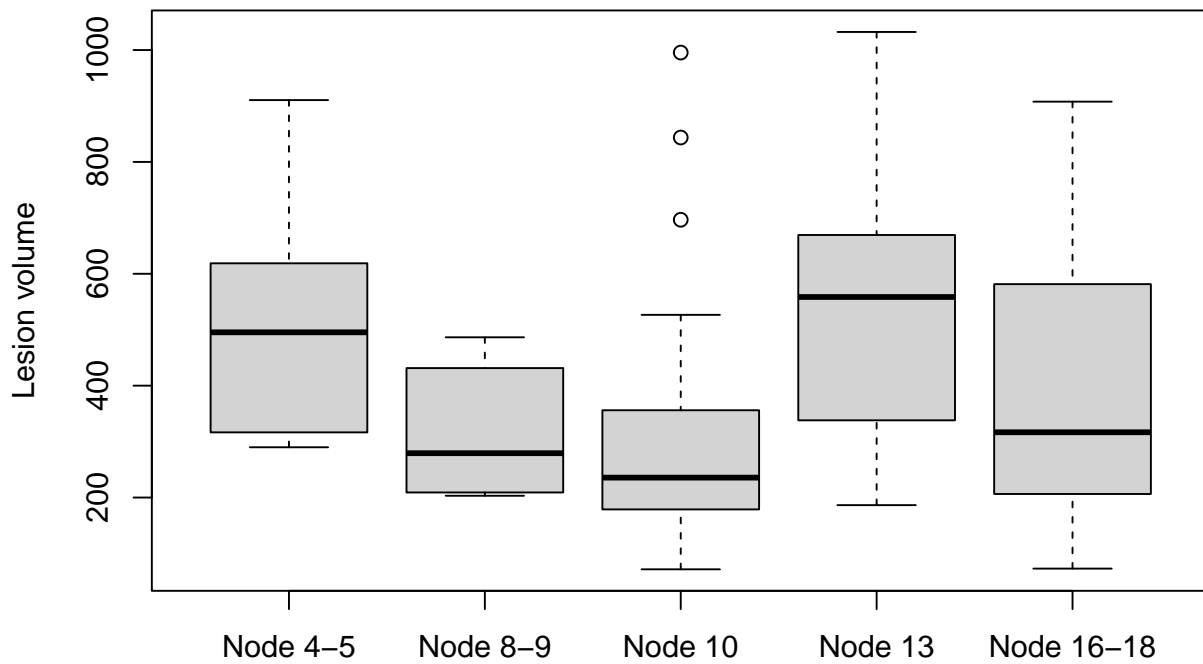


Figure 19: Boxplots of the lesion volume (in mm³) by URP-CTREE node.

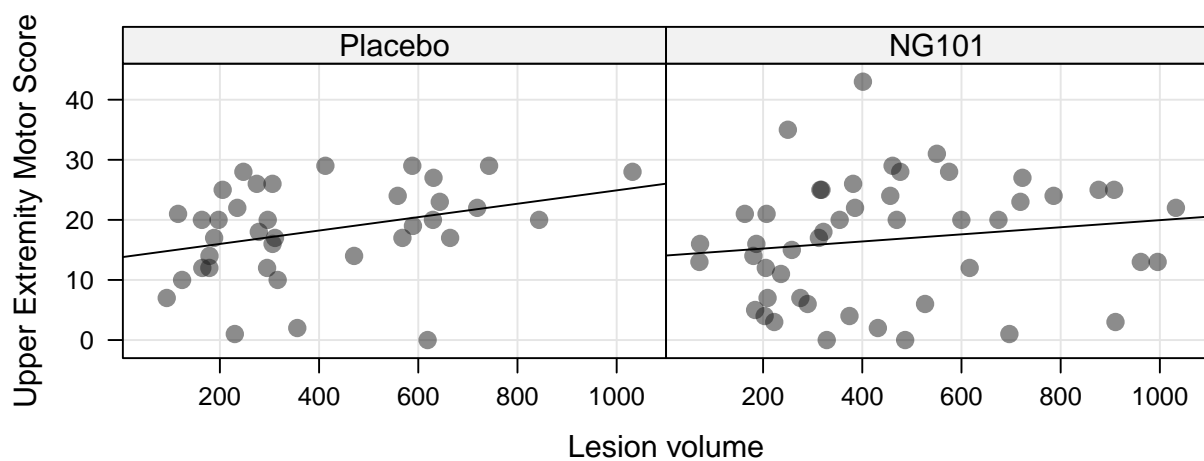


Figure 20: Scatterplot of the lesion volume (in mm³) against the baseline Upper Extremity Motor Score in the two treatment groups.

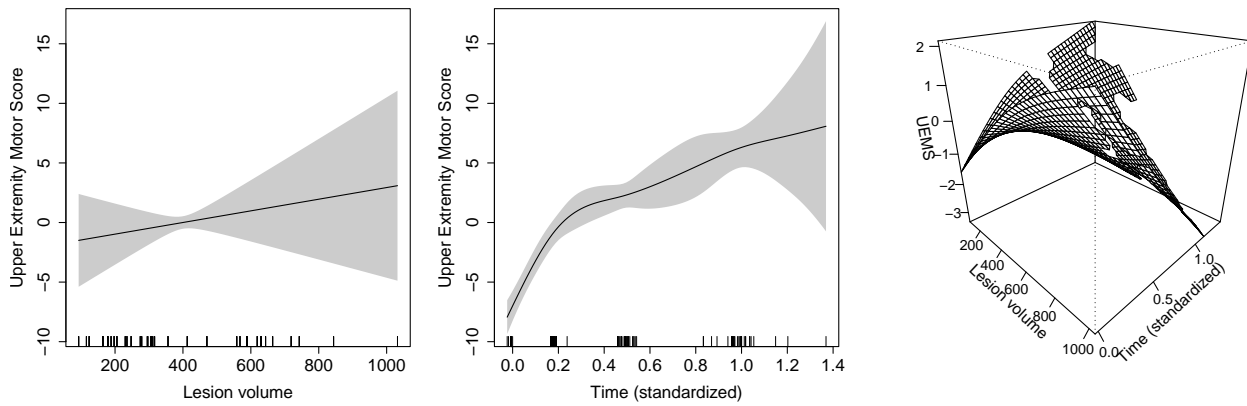


Figure 21: Smooth functions fitted by the GAM for the placebo group.

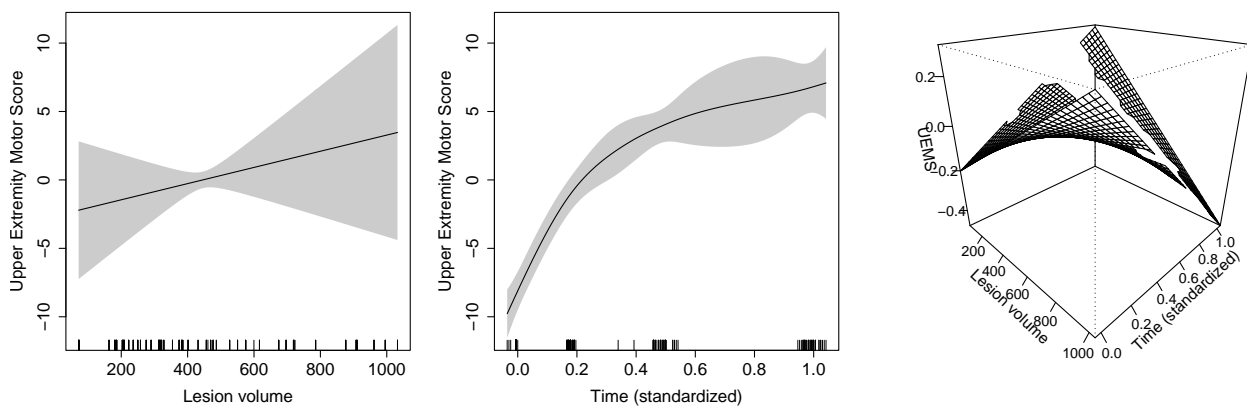


Figure 22: Smooth functions fitted by the GAM for the NG101 group.

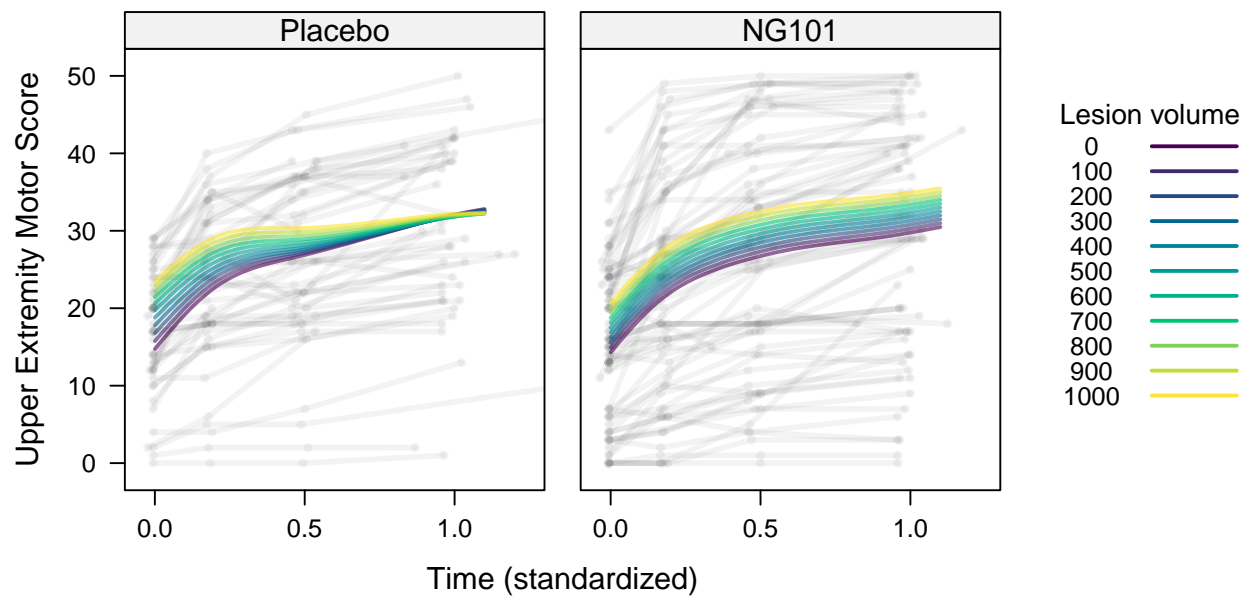


Figure 23: Predictions made with different values for lesion volume (in mm^3). For each treatment group the respective GAM was fitted separately. The only parameter that changes for the different predictions is the lesion volume.

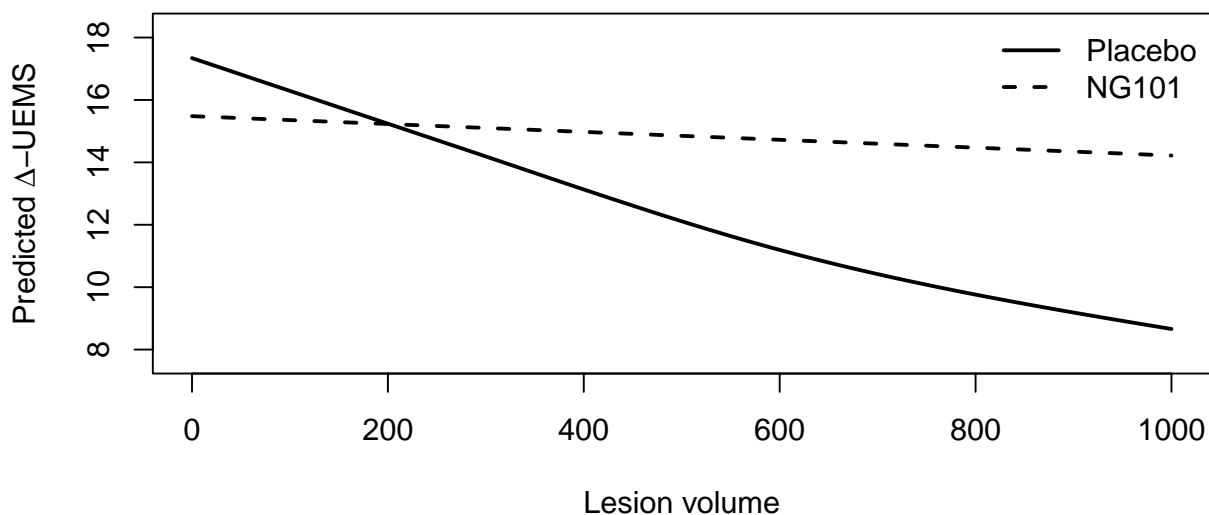


Figure 24: Differences of the predicted 6-month follow-up UEMS scores and the predicted baseline scores for different values of lesion volume (in mm^3).

Balgrist/Murnau: Midsagittal tissue bridges

Midsagittal tissue bridges were also recorded for patients from Balgrist/Murnau. On those patients, the same analysis as for the NISCI study is performed. In the NISCI study however, there are two treatment groups and the time variable is standardized such that 0 corresponds to day of first medication (on average 22.6 days after injury) and 1 corresponds to 6-month follow-up. To make the data from Balgrist/Murnau as comparable as possible to the NISCI data, the time variable is defined such that 0 corresponds to 22.6 days after injury and 1 corresponds to 6-month follow-up or 190.6 days after injury ($22.6 + 168$). In the NISCI study, each patient had its individual start of medication and therefore a different day after injury as baseline timepoint. Therefore, the results are not exactly comparable to Balgrist/Murnau. Another important point is that the first UEMS measurement for NISCI was taken on average on day 21.4 compared to day 10.5 for Balgrist/Murnau.

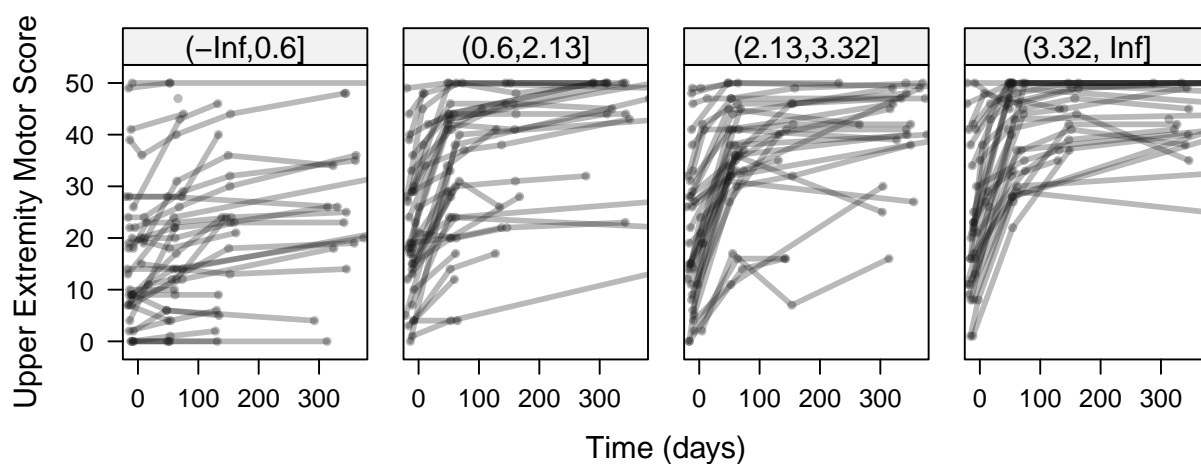


Figure 25: Patient data for Balgrist/Murnau, categorized into four quartiles based on the distribution of midsagittal tissue bridges. Patients were followed up for up to 12 months (or more) compared to the 6 months in the NISCI study. The time variable is defined such that 0 corresponds to day 22.6 after injury (which is equivalent to the average day after injury of first medication in the NISCI study).

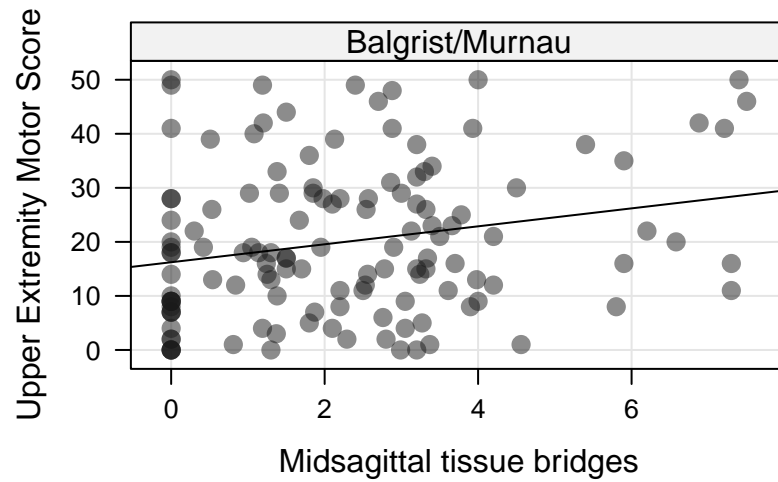


Figure 26: Scatterplot of the midsagittal tissue bridges (in mm) against the baseline Upper Extremity Motor Score for the Balgrist/Murnau patients.

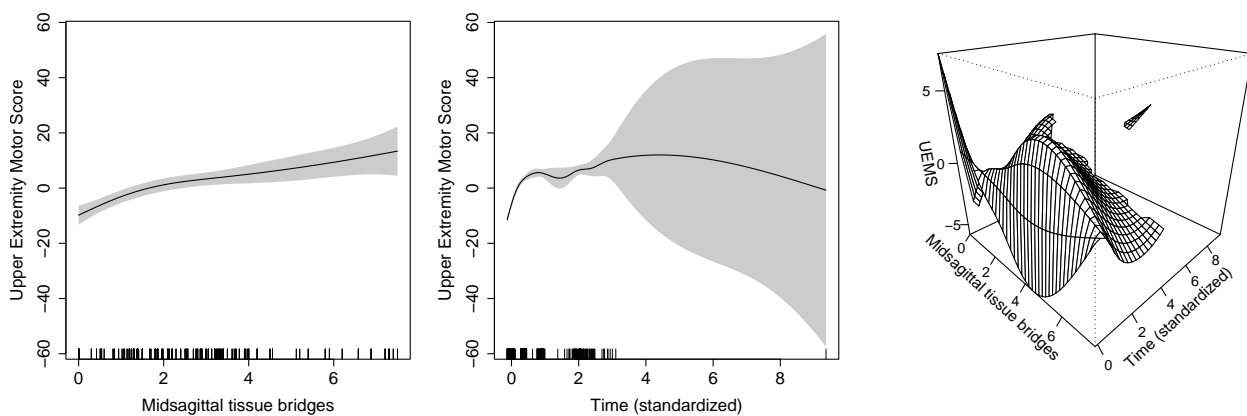


Figure 27: Smooth functions fitted by the GAM for the Balgrist/Murnau patients.

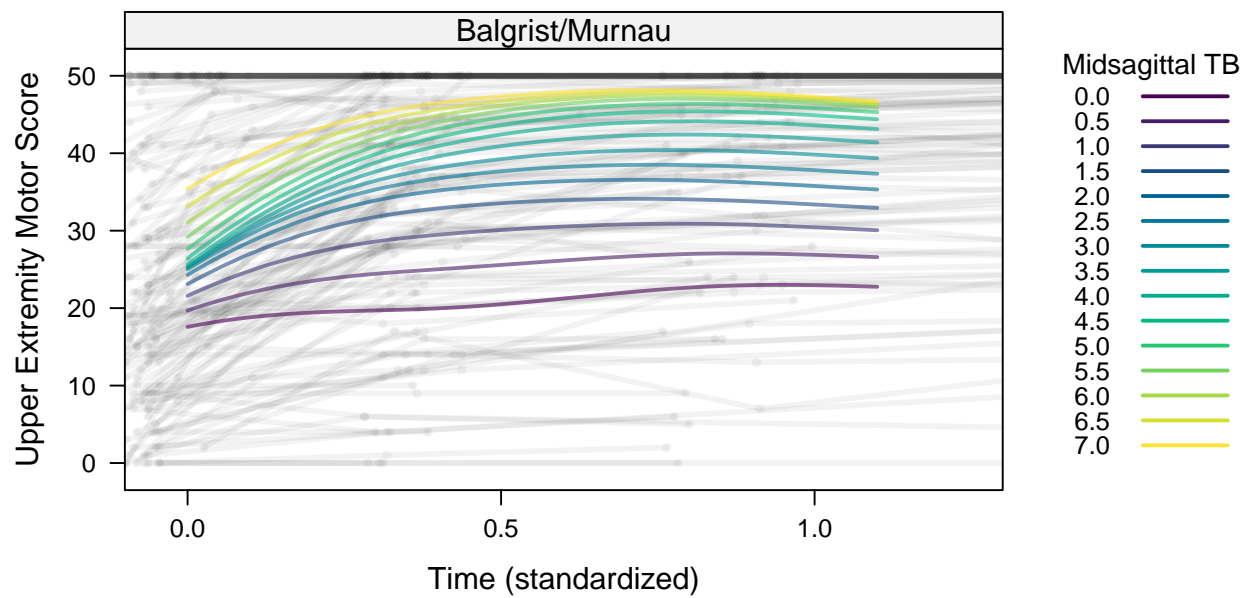


Figure 28: Predictions made by the fitted model for different values of midsagittal tissue bridges (in mm). The only parameter that changes for the different predictions is the amount of midsagittal tissue bridges.

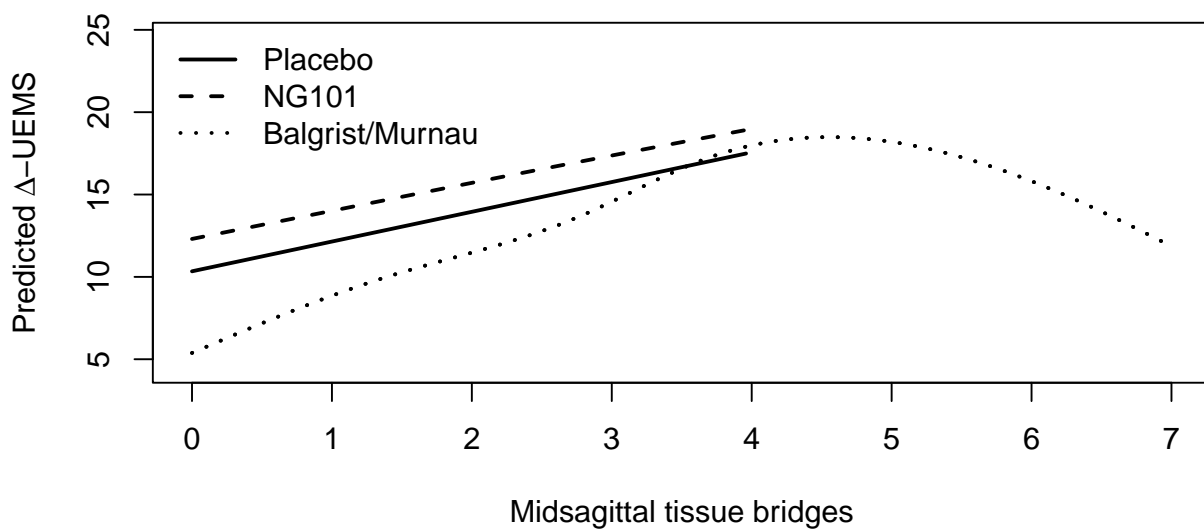


Figure 29: Differences of the predicted 6-month follow-up UEMS scores and the predicted baseline scores for different values of midsagittal tissue bridges (in mm). In the NISCI study, each patient had its individual start of medication (on average day 22.6 after injury) and therefore a different baseline timepoint. For Balgrist/Murnau, day 22.6 was set as baseline timepoint for all patients. Therefore, the results between the studies are not exactly comparable.

5.1 Cut-Off Values

The Δ -UEMS for the midsagittal tissue bridges is shown in Figure 10. The 6-month recovery linearly increases with the amount of midsagittal tissue bridges in both treatment groups. Therefore, it is not meaningful to define a distinct cut-off value for the midsagittal tissue bridges.

The Δ -UEMS for the parasagittal tissue bridges is shown in Figure 17. In terms of 6-month recovery, the placebo group shows a similar relationship as for the midsagittal tissue bridges. In the treatment group, the recovery levels off at 2 mm. Furthermore, the lines of the treatment groups cross at 3.5 mm. Patients with less than 3.5 mm parasagittal tissue bridges seem to benefit from the NG101 treatment, while patients with more preserved parasagittal tissue bridges show a better 6-month recovery under the placebo treatment.

The Δ -UEMS for the lesion volume in the placebo group is shown in Figure 24. The curves cross at 202 mm³. Patients with a lesion volume below 202 mm³ seem to benefit from the placebo treatment, while patients with a higher lesion volume show a better 6-month recovery under the NG101 treatment.

6 Conclusion

Our findings regarding the association of the midsagittal tissue bridges with the recovery support the results of previous studies. More preserved midsagittal tissue bridges correspond to an improved recovery as seen in Figure 9. Figure 10 suggests that the relationship between the amount of midsagittal tissue bridges and the 6-month recovery is linear in both treatment groups. The baseline distribution of the midsagittal tissue bridges against the UEMS in the treatment groups is shown in Figure 6. It slightly differs among the treatment groups. In the placebo group, smaller tissue bridge values seem to be associated with higher UEMS values, while in the NG101 group, the opposite is the case. The surprising baseline distribution in the placebo group is possibly the reason for the crossing recovery curves in Figure 9.

The baseline distribution of the parasagittal tissue bridges against the UEMS in the treatment groups, as shown in Figure 13, is similar to the midsagittal tissue bridges. Also the recovery profile in terms of the Δ -UEMS in the placebo group is comparable, as can be seen in Figure 17. However, in the NG101 group the recovery reaches a plateau at 2 mm and then slightly decreases for higher parasagittal tissue bridge values. Receiving the NG101 treatment seemed to be beneficial for patients with parasagittal tissue bridge values lower than 3.5 mm, whereas for higher parasagittal tissue bridge values, the placebo group showed a better 6-month recovery. In this case, it could be concluded to define a cut-off value at 3.5 mm and only include patients with less than this value in a follow up study, since they seem to benefit from the NG101 treatment.

The lesion volume in the placebo group shows a rather surprising baseline distribution as visualized in Figure 20. Patients with higher lesion volumes seem to have higher baseline UEMS values. This is contradictory to what one would expect. In the NG101 group, this effect is less extreme. Figure 23 shows that independent of the lesion volume, all recovery trajectories converge to the same UEMS level. In the NG101 group, the Δ -UEMS is almost constant over the range of lesion volumes (Figure 24). In the placebo group, the Δ -UEMS decreases with increasing lesion volume. As with the parasagittal tissue bridges, the lines of the predicted Δ -UEMS cross at some point (here 202 mm³). This makes intuitively sense, because low parasagittal tissue bridge values are generally associated with a high lesion volume. For high lesion volume, the NG101 group recovered better. This agrees with the observed worse performance under NG101 for higher parasagittal tissue bridge values.

Figure 25 of the Balgrist/Murnau patients already suggest some positive relationship between the amount of midsagittal tissue bridges and an improved recovery. Such a strong relationship was not visible in the NISCI data. However, it has to be kept in mind that the first UEMS measurement for NISCI was taken on average on day 21.4 compared to day 10.5 for Balgrist/Murnau. Figure 27 shows the smooth functions fitted by the GAM. The main effect of the midsagittal tissue bridges is positive and the uncertainty is lower compared

to the NISCI data. Figures 28 and 29 both suggest an improved recovery for higher midsagittal tissue bridge values. However, the Δ -UEMS decreases for midsagittal tissue bridge values above 4.6 mm. This is probably due to a ceiling effect, because the recovery potential is limited for patients with an already high baseline UEMS.

This analysis is subject to multiple limitations. The estimation of the relationship between biomarker and recovery includes high uncertainty. Especially in the NISCI data, as can be seen in the wide confidence intervals for the estimated smooth functions of the main effect of the biomarkers. Furthermore, the relatively high number of missing values could bias our inference. Patients also had varying medication and measurement timepoints. The sometimes different baseline distributions of the biomarkers against UEMS between the treatment groups make a reliable assessment of the relationships even more difficult. We analyzed different biomarkers that should all represent more or less the same, that is the physical state of the patient or the severity of the injury. However, the results were sometimes different depending on the biomarker (for example the Δ -UEMS for midsagittal vs. parasagittal tissue bridges). In Figure 28, it is visible that many of the Balgrist/Murnau patients have an UEMS of 50 points, which is problematic for the applied model which assumes normally distributed errors. In this case it might be more appropriate to take a more suitable distribution or to apply a transformation model with bounded outcomes. Because of all these limitations, the results should be viewed with caution.

To summarize, we gained an impression of the relationship between the biomarkers and the recovery. Midsagittal and parasagittal tissue bridges, as well as lesion volume, seem to have some prognostic value for the recovery. A cut-off value could be set at 3.5 mm for the parasagittal tissue bridges. This value seems to be predictive for the treatment effect. There is no difference in treatment effect visible for varying midsagittal tissue bridges. Therefore, the parasagittal tissue bridges seem to be most suitable to set a cut-off value. A cut-off value in terms of prognostic separation of patients into poor and good expected recovery does not seem to be meaningful, because there is no distinct range of biomarker values that indicates a poor or good recovery.

It is recommended to also stratify patients based on tissue bridges in future studies. This should prevent unfavourable baseline distributions from impacting the estimated treatment effect.

7 Generative AI Declaration

During the preparation of this report, I used Github Copilot and ChatGPT in order to be more efficient with coding. After using these tools/services, I reviewed and edited the content as needed and I take full responsibility for the content of the report.

References

- EVANIEW, N. ET AL. (2019). Unbiased recursive partitioning to stratify patients with acute traumatic spinal cord injuries: External validity in an observational cohort study. *Journal of Neurotrauma* **36** 2732–2742.
URL <https://doi.org/10.1089/neu.2018.6335>
- HASTIE, T. and TIBSHIRANI, R. (1986). Generalized Additive Models. *Statistical Science* **1** 297 – 310.
URL <https://doi.org/10.1214/ss/1177013604>
- HUBER, E. ET AL. (2017). Are midsagittal tissue bridges predictive of outcome after cervical spinal cord injury? *Annals of Neurology* **81** 740–748.
URL <https://onlinelibrary.wiley.com/doi/abs/10.1002/ana.24932>

- NINDS (2024). Spinal cord injury. Accessed: 2024-11-18.
URL <https://www.ninds.nih.gov/health-information/disorders/spinal-cord-injury>
- PFYFFER, D. ET AL. (2019). Tissue Bridges Predict Recovery after Traumatic and Ischemic Thoracic Spinal Cord Injury. *Neurology* **93** 10.1212/WNL.0000000000008318.
- PFYFFER, D. ET AL. (2024). Prognostic value of tissue bridges in cervical spinal cord injury: a longitudinal, multicentre, retrospective cohort study. *The Lancet Neurology* **23** 816–825.
URL [https://doi.org/10.1016/S1474-4422\(24\)00173-X](https://doi.org/10.1016/S1474-4422(24)00173-X)
- R CORE TEAM (2024). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.
URL <https://www.R-project.org/>
- SARKAR, D. (2008). *Lattice: Multivariate Data Visualization with R*. Springer, New York.
URL <http://lmdvr.r-forge.r-project.org>
- WEIDNER, N. ET AL. (2025). Safety and efficacy of intrathecal antibodies to nogo-a in patients with acute cervical spinal cord injury: a randomised, double-blind, multicentre, placebo-controlled, phase 2b trial. *The Lancet Neurology* **24** 42–53.
URL [https://doi.org/10.1016/S1474-4422\(24\)00447-2](https://doi.org/10.1016/S1474-4422(24)00447-2)
- WHO (2024). Spinal cord injury: Fact sheet. Accessed: 2024-11-18.
URL <https://www.who.int/news-room/fact-sheets/detail/spinal-cord-injury>
- WOOD, S. (2006). *Generalized Additive Models: An Introduction With R*, vol. 66.
- WOOD, S. (2011). Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. *Journal of the Royal Statistical Society (B)* **73** 3–36.

8 Appendix

8.1 Computational Details

This document was generated on Dezember 23, 2024 at 21:08. R version and packages used to generate this report:

- R version 4.4.1 (2024-06-14 ucrt), x86_64-w64-mingw32
- Locale: LC_COLLATE=German_Germany.utf8, LC_CTYPE=German_Germany.utf8, LC_MONETARY=German_Germany.utf8, LC_NUMERIC=C, LC_TIME=German_Germany.utf8
- Time zone: Europe/Zurich
- TZcode source: internal
- Running under: Windows 11 x64 (build 22631)
- Matrix products: default
- Base packages: base, datasets, graphics, grDevices, methods, stats, utils
- Other packages: biostatUZH 2.2.7, colorspace 2.1-1, dplyr 1.1.4, ggplot2 3.5.1, knitr 1.48, lattice 0.22-6, MASS 7.3-60.2, mgcv 1.9-1, naniar 1.1.0, nlme 3.1-164, RColorBrewer 1.1-3, readxl 1.4.3, stringr 1.5.1, survival 3.6-4, tableone 0.13.2, xtable 1.8-4
- Loaded via a namespace (and not attached): boot 1.3-30, cellranger 1.1.0, class 7.3-22, cli 3.6.3, cmprsk 2.2-12, codetools 0.2-20, compiler 4.4.1, data.table 1.16.2, DBI 1.2.3, digest 0.6.37, e1071 1.7-16, evaluate 1.0.0, fansi 1.0.6, farver 2.1.2, forcats 1.0.0, future 1.34.0, future.apply 1.11.2, generics 0.1.3, globals 0.16.3, glue 1.7.0, grid 4.4.1, gridExtra 2.3, gtable 0.3.5, haven 2.5.4, highr 0.11, hms 1.1.3, labeling 0.4.3, labelled 2.13.0, lava 1.8.0, lifecycle 1.0.4, listenr 0.9.1, lme4 1.1-35.5, magrittr 2.0.3, Matrix 1.7-0, minqa 1.2.8, mitools 2.4, munsell 0.5.1, nloptr 2.1.1, parallel 4.4.1, parallelly 1.38.0, pillar 1.9.0, pkgconfig 2.0.3, plyr 1.8.9, prodlim 2024.06.25, proxy 0.4-27, psy 1.2, R6 2.5.1, Rcpp 1.0.13, ReplicationSuccess 1.3.3, rlang 1.1.4, rstudioapi 0.17.1, scales 1.3.0, splines 4.4.1, stringi 1.8.4, survey 4.4-2, tibble 3.2.1, tidyr 1.2.1, tools 4.4.1, UpSetR 1.4.0, utf8 1.2.4, vctrs 0.6.5, visdat 0.6.0, withr 3.0.1, xfun 0.47

8.2 Code

```
#####
# code for packages, settings
#####
## Import external functions
## -----

## Packages
## -----
library(RColorBrewer) # colors for plots
library(tableone) # for Table 1 functions
library(xtable) # formatting tables and generating the tex code
library(biostatUZH) # EBPI-written package, if not installed, uncomment code below
# devtools::install_github(repo = "felix-hof/biostatUZH")
library(ggplot2) # customizable plots
library(stringr) # to prettify tables
library(readxl) # to import data
library(lattice) # for xyplot
library(mgcv) # for GAMs
library(colorspace) # sequential colors
library(dplyr)
library(naniar) # for missing data visualization (gg_miss_upset)
# library(excel.link)
# library(ILConnect) # to read password protected file
# library(gtsummary) # for table creation
### include project-specific packages here as well (e.g., lme4 for linear mixed effects models)
### if possible do not load libraries in chunks further below or in scripts that you source
```

```
#####
## only load libraries that you really use!###
#####

## Additional settings
## -----
cols <- brewer.pal(3, "Set1")
options(width = 85, digits = 4, show.signif.stars = FALSE)

#####
# code for preparation: data import and joining
#####
## -----
## Import Data
## -----

# Import primary dataset
# (it already includes parasagittal tissue bridges)
source("../data/NISCI_data.R")

# Import midsagittal tissue bridges
tissue_mid <- read_excel("../data/lesion_parameters_LF.xlsx")
# Subset only first visit
tissue_mid <- subset(tissue_mid, ExamStage == "NISCI01")
# add total_bridges from tissue_mid to datc (where tissue_mid(NISCI_ID) == datc(id))
datc$tissue_mid <- tissue_mid$total_bridges[match(datc$id, tissue_mid$NISCI_ID)]

# Import lesion volume
lesion_volume <- read_excel("../data/Lesion_volume_LF.xlsx")
# rename Slice Thickness
names(lesion_volume)[4] <- "Slice_Thickness"
# Subset only first visit
lesion_volume <- subset(lesion_volume, scan == 1)
# According to Lynn Farner (Zoom 31.10.24), if Slice_Thickness is NA, the lesion volume
# should also be NA and not 0. Hence, these values are set to NA here:
lesion_volume$Volume2[is.na(lesion_volume$Slice_Thickness)] <- NA
# add Volume2 to datc (where lesion_volume(id) == datc(id))
datc$lesion_volume <- lesion_volume$Volume2[match(datc$id, lesion_volume$id)]

# amount of lesion_volume > 1200 (7 patients)
# sum(lesion_volume$Volume2 > 1200, na.rm = TRUE)

# amount of visit_id == 2 with lesion_volume > 1200 (joined 6 patients,
# one patient was not in the main dataset)
# sum(datc$lesion_volume > 1200 & datc$visit_id == 2, na.rm = TRUE)

# Set 2 patients to NA where lesion volume not correct in dataset:

# Email by Lynn Farner: "Ich haben die Daten überprüft und es stimmt leider nicht ganz.
# BSL-001 und BYH-025 sollten ausgeschlossen werden, das habe ich beim Volumen falsch
# vermerkt sorry. BYH-002 ist in Ordnung für das Volumen aber nicht für die
# parasagittalen Tissue Bridges."
datc[datc$id == "BSL-001",]$lesion_volume <- NA
datc[datc$id == "BYH-025",]$lesion_volume <- NA
```

```

# store dataset including outliers
datc_with_outliers <- datc

# remove outliers for lesion volume
# set lesion_volume to NA where lesion_volume > 1200
datc$lesion_volume[datc$lesion_volume > 1200] <- NA

# subset data for treatment and placebo
datc_placebo <- subset(datc, trtplot=="Placebo")
datc_treatment <- subset(datc, trtplot=="NG101")

#####
# code for preparation: Balgrist/Murnau
#####

#library("mgcv")
#library("lme4")
#library("multcomp")
#library("lattice")
#library("readxl")
#library("dplyr")
#library("excel.link")

## read in data
# source("../data/NISCI_data.R")

# read midsagittal tissue bridges
# SR_mid <- read_excel("../data/lesion_parameters_LF.xlsx")

# Midsagittal tissue bridges (NISCI)

# SR_mid <- subset(SR_mid, ExamStage == "NISCI01")

# add total_bridges from SR_mid to datc (where SR_mid(NISCI_ID) == datc(ID))
# datc$SR_mid <- SR_mid$total_bridges[match(datc$id, SR_mid$NISCI_ID)]

# Load data: Balgrist & Murnau

# install.packages("XLConnect")

# Balgrist clinical scores (password protected)
# wb_balgrist <- loadWorkbook("../data/Balgrist_Clinical_Scores.xlsx", password = pw)
# balgrist <- readWorksheet(wb_balgrist, sheet = 1, header = TRUE)
#
#
# # Murnau clinical scores (password protected)
# wb_murnau <- loadWorkbook("../data/Murnau_Clinical_Scores.xlsx", password = pw)
# murnau <- readWorksheet(wb_murnau, sheet = 1, header = TRUE)
#
#
# # Midsagittal tissue bridges (Balgrist & Murnau)
# wb_bridges <- loadWorkbook("../data/TB_for_Lynn_Patrick.xlsx", password = pw)
#

```

```
# bridges <- readWorksheet(wb_bridges, sheet = 1, header = TRUE)

# save as .R to get rid off password protection.
# then source the .R files

# setwd("../data/")

# save as .R in report folder
# dump("balgrist", file = "balgrist.R")
# dump("murnau", file = "murnau.R")
# dump("bridges", file = "bridges.R")

# source the saved files:

source("../data/balgrist.R")
source("../data/murnau.R")
source("../data/bridges.R")

## Prepare Balgrist

# get names of variables
# names(balgrist)

balgrist_reduced <- balgrist %>%
  select('PatNr_Klinik', 'ExamStage', 'DOI',
        'ISNCSCI_TestDate', 'UEMS')

# create new column with time since injury (days)
balgrist_reduced$Interval_injury_UEMS <- balgrist_reduced$ISNCSCI_TestDate - balgrist_reduced$DOI

# change to factor
balgrist_reduced$PatNr_Klinik <- as.factor(balgrist_reduced$PatNr_Klinik)
balgrist_reduced$ExamStage <- as.factor(balgrist_reduced$ExamStage)

# change to numeric
balgrist_reduced$Interval_injury_UEMS <- as.numeric(balgrist_reduced$Interval_injury_UEMS)
balgrist_reduced$UEMS <- as.numeric(balgrist_reduced$UEMS)

# 64 patients
# length(unique(balgrist_reduced$PatNr_Klinik))

# 5 entries for each patient
# table(balgrist_reduced$ExamStage)
#
# head(balgrist_reduced)

balgrist_reduced$Study <- "Balgrist"

# assuming DOI is date of injury
# first_visit <- subset(balgrist_reduced, ExamStage == "very acute")
```

```
# mean time from injury to first visit (uems) : 8 days
# mean(first_visit$ISNCSCI_TestDate - first_visit$DOI, na.rm=TRUE)

## Prepare Murnau

# get names of variables
# names(murnau)

murnau_reduced <- murnau %>%
  select('ID', 'Interval_injury_mri_1', 'Interval_injury_isncsci_1', 'UEMS_1',
         'Interval_injury_isncsci_2', 'UEMS_2',
         'Interval_injury_isncsci_3', 'UEMS_3')

# change id to factor
murnau_reduced$ID <- as.factor(murnau_reduced$ID)
# change rest to numeric
murnau_reduced[,2:8] <- sapply(murnau_reduced[,2:8], as.numeric)

# # 93 patients
# length(unique(murnau_reduced$ID))
#
# # days from injury to first measurement (uems) : 9.8 days
# mean(murnau_reduced$Interval_injury_isncsci_1)
#
# # days from injury to second measurement (uems) : 78 days
# mean(murnau_reduced$Interval_injury_isncsci_2)
#
#
# # days from injury to third measurement (uems) : 336 days
# mean(murnau_reduced$Interval_injury_isncsci_3, na.rm=TRUE)

## Murnau wide to long

# Reshape the data using base R's reshape function
murnau_reduced_long <- reshape(
  murnau_reduced,
  direction = "long",
  idvar = "ID",
  varying = list(
    Interval = c("Interval_injury_isncsci_1",
                  "Interval_injury_isncsci_2", "Interval_injury_isncsci_3"),
```



```
    UEMS = c("UEMS_1", "UEMS_2", "UEMS_3")
  ),
  v.names = c("Interval_injury_UEMS", "UEMS"),
  timevar = "Stage",
  times = c(1, 2, 3)
)

# delete rownames
rownames(murnau_reduced_long) <- NULL

# long-format dataset with (days since injury)
# head(murnau_reduced_long)

murnau_reduced_long$Study <- "Murnau"

## Combine Balgrist and Murnau

names(balgrist_reduced)[1] <- "ID"

balgrist_murnau <- rbind(balgrist_reduced[c("ID", "Study", "Interval_injury_UEMS", "UEMS")],
                        murnau_reduced_long[c("ID", "Study", "Interval_injury_UEMS", "UEMS")])

# 157 patients
# length(unique(balgrist_murnau$ID))

## Prepare tissue bridges

# head(bridges)
#
# nrow(bridges)
# # get names of variables
# names(balgrist)

# 64 patients
# length(unique(balgrist_reduced$PatNr_Klinik))

# 43 Balgrist, 93 Murnau
# table(bridges$`center (0=Balgrist)`)

# add tb_tot to balgrist_murnau (where PID == balgrist_murnau(ID))
balgrist_murnau$tissue_mid <- bridges$tb_tot[match(balgrist_murnau$ID, bridges$PID)]

# # change Interval_injury_UEMS, UEMS and tissue_mid to numeric
# balgrist_murnau <- balgrist_murnau %>%
#   mutate(across(c(Interval_injury_UEMS, UEMS, tissue_mid), as.numeric))
```

```
# head(balgrist_murnau)

## Plot raw data: Balgrist & Murnau

# Most patients were recorded well beyond 6 months

# library(lattice)
#
# xyplot(UEMS ~ Interval_injury_UEMS | Study,
#        data = balgrist_murnau,
#        group = ID,
#        xlim = c(-0.1, 1.9),
#        type = "b",
#        between = list(x = 1, y = 1),
#        pch = 20,
#        cex = .7,
#        lwd = 3,
#        col = rgb(.1, .1, .1, alpha = 0.05), # Fade the primary data lines using transparency (alpha)
#        xlab = "Time (standardized)",
#        ylab = "UEMS",
#        main = "Predictions for different SR values (placebo vs treatment group)"
# )

### Check time variable in NISCI and Murnau/Balgrist

## Transform datc (NISCI) to days since injury

# Create a column "Interval since injury to measurement of UEMS" in the datc.

datc$Interval_injury_UEMS <- datc$medistart + datc$tm

visit_2 <- subset(datc, visit_id == 2)
visit_11 <- subset(datc, visit_id == 11)

## NISCI: descriptive statistics

# Average medistart for NISCI was 22.5 days after injury. The first UEMS measurement was
# taken on average 21.4 days after injury. For the standardized time (tmstd),
# 0 corresponded to day of first dosing and 1 corresponds to day 168 after
# dosing (6 months). Last measurement was taken before tmstd=1 (168 days
# after dosing) for more than 75% of patients.
```

```

# average medistart (== tmst 0): 22.55
# mean(visit_2$medistart)
#
# # average time of first measurement since injury (uems): 21.4 days
mean_first_UEMS_NISCI <- round(mean(visit_2$Interval_injury_UEMS),1)

# # check date calculation with an example patient
# patient1 <- datc %>% filter(id == "BCA-001")
#
#
# # start of medication (days after injury) = 17
# patient1$medistart
#
# # study day of test = -1 30 84 168
# patient1$tm
#
# # study day of test (standardized) = -0.005952381 0.178571429 0.500000000 1.000000000
# patient1$tmstd
#
# # calculated by dividing tm by 168 (equal to tmstd)
# patient1$tm /168

# check average time (tmstd) of last measurement

# summary(visit_11$tmstd)
# boxplot(visit_11$tmstd)

## Balgrist & Murnau: descriptive statistics

# In the combined dataset of Balgrist and Murnau, the average time of first measurement since injury
# was 10.48 days.

# average time of first measurement since injury (uems): 10.48 days

mean_day_first_measurement <- balgrist_murnau %>%
  group_by(ID) %>%
  # sort from smallest to largest interval
  arrange(Interval_injury_UEMS) %>%
  # get the first measurement for each patient
  slice(1) %>%
  # ungroup
  ungroup() %>%
  # get average time of first measurement since injury (uems)
  summarise(mean_interval = mean(Interval_injury_UEMS, na.rm = TRUE))

# mean_day_first_measurement[[1]]

```

```
## time standardization for Balgrist & Murnau:

#####
# Set Balgrist/Murnau baseline to the average medistart of the NISCI study (22.6 days): hence tmstd==0
# will be day 22.6 after injury. Note that many measurements of the uems were taken well before day 23.
# These measurements are fitted in the model, but for later calculation of the delta-UEMS the
# difference between day 22.6 and 6 month follow up are taken to make the estimate comparable
# to the NISCI study.
#####

baseline <- round(mean(visit_2$medistart),1)

# define the new time variable tm as time since injury to measurement - baseline (22.6 days)
balgrist_murnau$tm <- balgrist_murnau$Interval_injury_UEMS - baseline

# define the new time variable tmstd as tm divided by 168 (6 months)
balgrist_murnau$tmstd <- balgrist_murnau$tm / 168

names(balgrist_murnau)[1] <- "id"
names(balgrist_murnau)[4] <- "uems"

#####
# code for results: data collection
#####

# mean time and sd for MRI timepoint (days after injury)

MRI_timepoint_mean <- mean(tissue_mid$time, na.rm = TRUE)

MRI_timepoint_sd <- sd(tissue_mid$time, na.rm = TRUE)

# to get the baseline values, we analyze the dataset including outliers

# subset only baseline
visit_2 <- subset(datc_with_outliers, visit_id == 2)

# Number of total patients (126)
total_patients <- length(unique(visit_2$id))

# total countries
total_countries <- length(unique(visit_2$country))

# total patients in treatment and placebo
total_treatment <- sum(visit_2$trtplot == "NG101")
total_placebo <- sum(visit_2$trtplot == "Placebo")

# age mean and sd
mean_age <- mean(visit_2$bl_age, na.rm = TRUE)
sd_age <- sd(visit_2$bl_age, na.rm = TRUE)

# male percentage
percentage_male <- table(visit_2$bl_sex)[1]/total_patients

# total patients with tissue_mid
```

```

total_tissue_mid <- sum(!is.na(visit_2$tissue_mid))

# table of number of records per patient (0, 1, 2, 3, 4)
number_UEMS_records <- table(table(datc$id))

# mean medication start
mean_medistart <- mean(visit_2$medistart, na.rm = TRUE)

# table 1

# subset only baseline measures
visit_2 <- subset(datc_with_outliers, visit_id == 2)

## Vector of variables to summarize
myVars <- c("medistart", "uems", "tissue_mid", "SR", "lesion_volume")

# Create a TableOne object with an overall column
tab_one <- CreateTableOne(vars = myVars, strata = "trt", data = visit_2,
                          addOverall = TRUE, test = FALSE)

# Convert TableOne to data frame for LaTeX formatting
tab_one_df <- print(tab_one, printToggle = FALSE, noSpaces = TRUE)

rownames(tab_one_df) <- c("Patients", "Time from injury to 1st dose (days)", "UEMS",
                          "Midsagittal tissue bridges (mm)", "Parasagittal tissue bridges (mm)",
                          "Lesion volume (mm3)")

colnames(tab_one_df) <- c("Overall", "Placebo", "NG101")

tab_one_xt <- xtable(tab_one_df, caption = "Baseline characteristics by treatment group in the
NISCI study.

The days to first medication as well as the baseline measurements of
the UEMS and the three biomarkers are presented as mean (sd).",
                    label = "tab:table_one_results")

# Print xtable
print(tab_one_xt, include.rownames = TRUE, caption.placement = "top", type = "latex",
      sanitize.text.function=function(x){x})

# <<missing_table>>

# create a plot of the missing value combinations (same as table above)

missing_set <- visit_2[, c("tissue_mid", "SR", "lesion_volume")]
colnames(missing_set) <- c("Midsagittal", "Parasagittal", "Lesion_volume")

# Generate a missing data heatmap
gg_miss_upset <- gg_miss_upset(missing_set,
                              nsets = 5)

# Display the plot

```

```

gg_miss_upset

# Select individuals with missing parasagittal but recorded midsagittal tissue bridges
# show them in a table
# -> according to Lynn Farner, Patient BHY-002 indeed has a missing value for
# parasagittal tissue but a valid lesion volume

missing_SR <- visit_2 %>%
  filter(is.na(SR) & !is.na(tissue_mid)) %>%
  select(id, tissue_mid, SR, lesion_volume)

# Replace NA values with "Missing"
missing_SR[is.na(missing_SR)] <- "Missing"

colnames(missing_SR) <- c("Patient-ID", "Tissue midsagittal", "Tissue parasagittal", "Lesion volume")

tab_SR_missing <- xtable(missing_SR, caption = "8 Patients that have recorded midsagittal tissue bridge
                        values but no parasagittal tissue bridges. This should not be possible.
                        Check with clinicians.",
                        label = "tab:table_SR_missing")

# Print xtable
print(tab_SR_missing, include.rownames = FALSE,
      caption.placement = "top", type = "latex")

# mean time of UEMS measurements in the respective visit_id (in days after first dose)
mean_time <- datc %>%
  group_by(visit_id) %>%
  summarize(mean_time = mean(tm, na.rm = TRUE))

# Boxplots of the biomarkers at baseline

par(mfrow=c(1,3))
boxplot(visit_2$tissue_mid, xlab = "Midsagittal tissue bridges")
boxplot(visit_2$SR, xlab = "Parasagittal tissue bridges")
boxplot(lesion_volume$Volume2, xlab = "Lesion volume")

# amount of patients with lesion volume > 1200 (they are excluded from the analysis)
patients_lesion_1200 <- sum(datc_with_outliers[datc_with_outliers$visit_id == 2,]$lesion_volume > 1200,
                          na.rm = TRUE)

#####
# code for results: relationship of biomarkers with recovery (code-chunk-reuse)
#####

# Define quartiles for the selected biomarker, independent of treatment groups and based on visit 2
quantiles <- quantile(subset(datc, visit_id == 2)[[biomarker]], prob = 1:3 / 4, na.rm = TRUE)

# Categorize the biomarker based on quartiles
datc[[paste0(biomarker, "_c")]] <- cut(datc[[biomarker]], breaks = c(-Inf, quantiles, Inf),
                                     as.ordered = TRUE)

```

```

# Plot the raw data of recovery separately per treatment group and quartile of biomarker
xyplot(
  uems ~ tm | get(paste0(biomarker, "_c")) + trtplot, data = datc,
  group = id, xlim = c(-5, 240), type = "b", between = list(x = 1, y = 1),
  pch = 20, cex = 0.7, lwd = 3, col = rgb(0.1, 0.1, 0.1, 0.3),
  xlab = "Time (days)", ylab = "Upper Extremity Motor Score",
  scales = list(
    x = list(alternating = 1), # Show x-axis tick labels only at the bottom,
    y = list(alternating = 1), # Show y-axis tick labels only at the left
    tck = c(1, 0) # Remove top ticks
  )
  #, main = paste("UEMS over time by", biomarker_name, "quartiles and treatment group")
)

# Boxplots of the biomarker in each node (treatment groups combined)

boxplot(as.formula(paste(biomarker, " ~ nodeCplot")), data = subset(datc, visit_id == 2), na.rm = TRUE,
  ylab = biomarker_name_upper,
  xlab = "")

# Plot the baseline UEMS against the baseline biomarker value

xyplot(as.formula(paste("uems ~", biomarker, "| trtplot")),
  data = subset(datc, visit_id == 2),
  group = id,
  type = "b",
  pch = 16,
  cex = 1.2,
  lwd = 2,
  col = rgb(.1, .1, .1, .5),
  xlab = biomarker_name_upper,
  ylab = "Upper Extremity Motor Score",
  #main = paste("Baseline", biomarker_name, "vs UEMS"),
  grid = TRUE,
  scales = list(
    x = list(alternating = 1), # Show x-axis tick labels only at the bottom
    tck = c(1, 0) # Remove top ticks
  ),
  panel = function(x, y, ...) {
    panel.xyplot(x, y, ...) # Plot the points
    panel.lmline(x, y, col = "black", lwd = 1) # Add a linear regression line
  }
)

# fit GAM for each treatment group separately

# Define the formula
gam_formula <- as.formula(paste("uems ~ s(", biomarker, ") + s(tmstd) + ti(", biomarker, ", tmstd) +
  s(id, bs = 're') + s(id, tmstd, bs = 're')"))

# Fit the GAM
gam_placebo <- gam(gam_formula,
  data = datc_placebo, method = "REML")

gam_treatment <- gam(gam_formula,

```

```
data = datc_treatment, method = "REML")

# <<summary_table>>

# plot the smooth functions fitted by the GAM for the placebo group

par(mfrow = c(1, 3))

plot(gam_placebo, scheme = 1, select = 1,
     seWithMean = TRUE, rug = TRUE,
     xlab = biomarker_name_upper, ylab = "Upper Extremity Motor Score",
     cex.lab = 1.5, cex.axis = 1.5)

plot(gam_placebo, scheme = 1, select = 2,
     seWithMean = TRUE, rug = TRUE,
     xlab = "Time (standardized)", ylab = "Upper Extremity Motor Score",
     cex.lab = 1.5, cex.axis = 1.5)

plot(gam_placebo, scheme = 1, select = 3,
     seWithMean = TRUE, rug = TRUE, main = "UEMS",
     xlab = biomarker_name_upper, ylab = "Time (standardized)", ticktype = "detailed",
     cex.lab = 1.45, cex.axis = 1.3, cex.main = 1.45, theta = 45, # Rotate around the z-axis
     phi = 25) # Adjust elevation angle (tilt the plot vertically)

# plot the smooth functions fitted by the GAM for the treatment group

par(mfrow = c(1, 3))

plot(gam_treatment, scheme = 1, select = 1,
     seWithMean = TRUE, rug = TRUE,
     xlab = biomarker_name_upper, ylab = "Upper Extremity Motor Score",
     cex.lab = 1.5, cex.axis = 1.5)

plot(gam_treatment, scheme = 1, select = 2,
     seWithMean = TRUE, rug = TRUE,
     xlab = "Time (standardized)", ylab = "Upper Extremity Motor Score",
     cex.lab = 1.5, cex.axis = 1.5,)

plot(gam_treatment, scheme = 1, select = 3,
     seWithMean = TRUE, rug = TRUE, main = "UEMS",
     xlab = biomarker_name_upper, ylab = "Time (standardized)", ticktype = "detailed",
     cex.lab = 1.45, cex.axis = 1.3,
     cex.main = 1.45, theta = 45, # Rotate around the z-axis
     phi = 25) # Adjust elevation angle (tilt the plot vertically)

# predict recovery trajectories based on different values of the biomarker

models <- list(gam_placebo, gam_treatment)
```



```

# Prediction function using biomarker variable
make_prediction <- function(biomarker_value, model) {
  # new_data includes time at baseline up to 6 months
  # (used standardized time = 1.1 just for visualization)
  new_data <- expand_grid(id = 1, tmstd = seq(0, 1.1, length = 100))
  # current biomarker value is added to the new data
  # current "biomarker" is defined before the code chunk
  new_data[[biomarker]] <- biomarker_value

  # Make predictions excluding random effects
  prediction <- data.frame(
    tmstd = new_data$tmstd,
    pred = predict(model, newdata = new_data, exclude = "s(id)")
  )
  return(prediction)
}

# Set up colors for plotting
colors_plot <- sequential_hcl(length(biomarker_values), palette = "Viridis")

legend_key <- list(
  text = list(labels = paste0(format(biomarker_values,
                                     nsmall = ifelse(biomarker == "lesion_volume", 0, 1)))),
  lines = list(col = colors_plot, lwd = 2), # Corresponding colors
  space = "right",
  columns = 1,
  title = biomarker_name_legend,
  cex.title = 0.9,
  cex = 0.8
)

# Plot the raw recovery data per treatment group and
# add the predictions for different values of the biomarker
xyplot(uems ~ tmstd | trtplot,
  data = datc,
  group = id,
  xlim = c(-0.1, 1.3),
  type = "b",
  between = list(x = 1, y = 1),
  pch = 20,
  cex = .7,
  lwd = 3,
  col = rgb(.1, .1, .1, alpha = 0.05), # Fade primary data lines
  xlab = "Time (standardized)",
  ylab = "Upper Extremity Motor Score",
  #main = paste("Predictions for different", biomarker_name, "values"),
  key = legend_key,
  scales = list(
    x = list(alternating = 1), # Show x-axis tick labels only at the bottom
    tck = c(1, 0) # Remove top ticks
  ),
  # add predictions
  panel = function(x, y, subscripts, ...) {
    panel.xyplot(x, y, subscripts = subscripts, ...)
    # panel_index indicating treatment group (1 = placebo, 2 = NG101)
    panel_index <- panel.number()
  }
)

```

```

    # for each biomarker value, make predictions and add to plot
    for (i in seq_along(biomarker_values)) {
      biomarker_value <- biomarker_values[i]
      # Generate predictions for each treatment group and biomarker value
      preds <- make_prediction(biomarker_value = biomarker_value, model = models[[panel_index]])
      panel.lines(preds$tmstd, preds$pred, col = colors_plot[i], alpha = 0.6, lwd = 2)
    }
  })

# this chunk generates predictions for baseline UEMS and 6-month UEMS for different values
# of the biomarker and plots the delta-UEMS over the range of the biomarker

# number of predicted delta-UEMS
N <- 100

# predictions made along the sequence of min(biomarker_value) to max(biomarker_value) of the biomarker
prediction_sequence <- seq(from = min(biomarker_values), to = max(biomarker_values), length.out = N)

# Create new data for predictions using biomarker variable dynamically
#new_data <- expand.grid(id = 1, tmstd = c(0, 1), biomarker_value = biomarker_values)

#new_data <- data.frame(id = rep(1, 2 * length(biomarker_values)),
# tmstd = rep(c(0, 1), each = length(biomarker_values)))

# new_data for tmstd=0 and tmstd=1
new_data <- data.frame(id = rep(1, 2 * N), tmstd = rep(c(0,1), each = N))

# add biomarker value sequence to new data
new_data[[biomarker]] <- rep(prediction_sequence, 2)

# Generate predicted UEMS for placebo and treatment groups at tmstd=0 and tmstd=1
predictions <- data.frame(
  tmstd = new_data$tmstd,
  biomarker = new_data[[biomarker]],
  pred_placebo = predict(gam_placebo, newdata = new_data, exclude = "s(id)",
  pred_treatment = predict(gam_treatment, newdata = new_data, exclude = "s(id)"
)

# Calculate the predicted delta change between tmstd = 1 and tmstd = 0 for each biomarker value
predicted_delta <- data.frame(
  biomarker_value = prediction_sequence,
  delta_placebo = predictions[predictions$tmstd == 1,]$pred_placebo
- predictions[predictions$tmstd == 0,]$pred_placebo,
  delta_treatment = predictions[predictions$tmstd == 1,]$pred_treatment
- predictions[predictions$tmstd == 0,]$pred_treatment
)

# Plot the predicted delta-UEMS for each biomarker value
plot(predicted_delta$biomarker_value, predicted_delta$delta_placebo, type = "l",
  #col = "#69b3a2", lwd = 2,
  col = "black", lwd = 2, lty = 1,
  ylim = c(min(predicted_delta[,2:3])-1, max(predicted_delta[,2:3])+1),

```

```

xlab = biomarker_name_upper,
ylab = expression("Predicted " * Delta * "-UEMS")
#, main = bquote("Predicted " * Delta * "-UEMS for different" ~ .(biomarker_name) ~ "values")
)

# Add line for the treatment group
lines(predicted_delta$biomarker_value, predicted_delta$delta_treatment, #col = "#f4a582", lwd = 2,
      col = "black", lwd = 2, lty = 2)

# Add legend
legend(legend_location, legend = c("Placebo", "NG101"), lty = c(1, 2), lwd = 2, bty = "n")

#####
# Parasagittal tissue bridges: find turning and crossing point

legend_location <- "topleft"

# this chunk generates predictions for baseline UEMS and 6-month UEMS for different values
# of the biomarker and plots the delta-UEMS over the range of the biomarker

# number of predicted delta-UEMS
N <- 100

# predictions made along the sequence of min(biomarker_value) to max(biomarker_value) of the biomarker
prediction_sequence <- seq(from = min(biomarker_values), to = max(biomarker_values), length.out = N)

# Create new data for predictions using biomarker variable dynamically
#new_data <- expand.grid(id = 1, tmstd = c(0, 1), biomarker_value = biomarker_values)

#new_data <- data.frame(id = rep(1, 2 * length(biomarker_values)),
# tmstd = rep(c(0, 1), each = length(biomarker_values)))

# new_data for tmstd=0 and tmstd=1
new_data <- data.frame(id = rep(1, 2 * N), tmstd = rep(c(0,1), each = N))

# add biomarker value sequence to new data
new_data[[biomarker]] <- rep(prediction_sequence, 2)

# Generate predicted UEMS for placebo and treatment groups at tmstd=0 and tmstd=1
predictions <- data.frame(
  tmstd = new_data$tmstd,
  biomarker = new_data[[biomarker]],
  pred_placebo = predict(gam_placebo, newdata = new_data, exclude = "s(id)",
  pred_treatment = predict(gam_treatment, newdata = new_data, exclude = "s(id)"
)

# Calculate the predicted delta change between tmstd = 1 and tmstd = 0 for each biomarker value
predicted_delta <- data.frame(
  biomarker_value = prediction_sequence,
  delta_placebo = predictions[predictions$tmstd == 1,]$pred_placebo
- predictions[predictions$tmstd == 0,]$pred_placebo,
  delta_treatment = predictions[predictions$tmstd == 1,]$pred_treatment

```

```

- predictions[predictions$tmstd == 0,]$pred_treatment
)

# Plot the predicted delta-UEMS for each biomarker value
plot(predicted_delta$biomarker_value, predicted_delta$delta_placebo, type = "l",
      #col = "#69b3a2", lwd = 2,
      col = "black", lwd = 2, lty = 1,
      ylim = c(min(predicted_delta[,2:3])-1, max(predicted_delta[,2:3])+1),
      xlab = biomarker_name_upper,
      ylab = expression("Predicted " * Delta * "-UEMS")
      #, main = bquote("Predicted " * Delta * "-UEMS for different" ~ .(biomarker_name) ~ "values")
)

# Add line for the treatment group
lines(predicted_delta$biomarker_value, predicted_delta$delta_treatment, #col = "#f4a582", lwd = 2,
      col = "black", lwd = 2, lty = 2)

# Add legend
legend(legend_location, legend = c("Placebo", "NG101"), lty = c(1, 2), lwd = 2, bty = "n")

# Find the index where treatment group levels off
diff <- (c(predicted_delta$delta_treatment,0)-c(0, predicted_delta$delta_treatment))

first_negative_index <- which(diff < 0)[1]
parasagittal_turnpoint <- predicted_delta$biomarker_value[first_negative_index]

# Find the index of the first negative value
diff <- (predicted_delta$delta_placebo - predicted_delta$delta_treatment)

first_positive_index <- which(diff > 0)[1]

parasagittal_crossing <- predicted_delta$biomarker_value[first_positive_index]

#####
# Lesion volume: find crossing point

legend_location <- "topright"

# this chunk generates predictions for baseline UEMS and 6-month UEMS for different values
# of the biomarker and plots the delta-UEMS over the range of the biomarker

# number of predicted delta-UEMS
N <- 100

# predictions made along the sequence of min(biomarker_value) to max(biomarker_value) of the biomarker
prediction_sequence <- seq(from= min(biomarker_values), to= max(biomarker_values), length.out = N)

# Create new data for predictions using biomarker variable dynamically
#new_data <- expand.grid(id = 1, tmstd = c(0, 1), biomarker_value = biomarker_values)

#new_data <- data.frame(id = rep(1, 2 * length(biomarker_values)),

```

```

# tmstd = rep(c(0, 1), each = length(biomarker_values))

# new_data for tmstd=0 and tmstd=1
new_data <- data.frame(id = rep(1, 2 * N), tmstd = rep(c(0,1), each = N))

# add biomarker value sequence to new data
new_data[[biomarker]] <- rep(prediction_sequence, 2)

# Generate predicted UEMS for placebo and treatment groups at tmstd=0 and tmstd=1
predictions <- data.frame(
  tmstd = new_data$tmstd,
  biomarker = new_data[[biomarker]],
  pred_placebo = predict(gam_placebo, newdata = new_data, exclude = "s(id)",
  pred_treatment = predict(gam_treatment, newdata = new_data, exclude = "s(id)"
)

# Calculate the predicted delta change between tmstd = 1 and tmstd = 0 for each biomarker value
predicted_delta <- data.frame(
  biomarker_value = prediction_sequence,
  delta_placebo = predictions[predictions$tmstd == 1,]$pred_placebo
- predictions[predictions$tmstd == 0,]$pred_placebo,
  delta_treatment = predictions[predictions$tmstd == 1,]$pred_treatment
- predictions[predictions$tmstd == 0,]$pred_treatment
)

# Plot the predicted delta-UEMS for each biomarker value
plot(predicted_delta$biomarker_value, predicted_delta$delta_placebo, type = "l",
  #col = "#69b3a2", lwd = 2,
  col = "black", lwd = 2, lty = 1,
  ylim = c(min(predicted_delta[,2:3])-1, max(predicted_delta[,2:3])+1),
  xlab = biomarker_name_upper,
  ylab = expression("Predicted " * Delta * "-UEMS")
  #, main = bquote("Predicted " * Delta * "-UEMS for different" ~ .(biomarker_name) ~ "values")
)

# Add line for the treatment group
lines(predicted_delta$biomarker_value, predicted_delta$delta_treatment, #col = "#f4a582", lwd = 2,
  col = "black", lwd = 2, lty = 2)

# Add legend
legend(legend_location, legend = c("Placebo", "NG101"), lty = c(1, 2), lwd = 2, bty = "n")

# Find the index of the first negative value
diff <- (predicted_delta$delta_placebo - predicted_delta$delta_treatment)

first_negative_index <- which(diff < 0)[1]

lesion_volume_crossing <- predicted_delta$biomarker_value[first_negative_index]

```

```
#####
# Balgrist/Murnau: Midsagittal tissue bridges

## quartile plot for Balgrist/Murnau

biomarker <- "tissue_mid"
biomarker_name <- "midsagittal tissue bridges"
biomarker_name_upper <- "Midsagittal tissue bridges"
biomarker_name_legend <- "Midsagittal TB"

# Define quartiles
quantiles <- quantile(balgrist_murnau[[biomarker]], prob = 1:3 / 4, na.rm = TRUE)

# Categorize the biomarker based on quartiles
balgrist_murnau[[paste0(biomarker, "_c")]] <- cut(balgrist_murnau[[biomarker]],
                                                  breaks = c(-Inf, quantiles, Inf), as.ordered = TRUE)

# Plot the raw data of recovery separately per treatment group and quartile of biomarker
xyplot(
  uems ~ tm | get(paste0(biomarker, "_c")), data = balgrist_murnau,
  group = id, xlim = c(-25, 380), type = "b", between = list(x = 1, y = 1),
  layout = c(4, 1),
  pch = 20, cex = 0.7, lwd = 3, col = rgb(0.1, 0.1, 0.1, 0.3),
  xlab = "Time (days)", ylab = "Upper Extremity Motor Score",
  scales = list(
    x = list(alternating = 1), # Show x-axis tick labels only at the bottom,
    y = list(alternating = 1), # Show y-axis tick labels only at the left
    tck = c(1, 0) # Remove top ticks
  )
)

## baseline plot for Balgrist/Murnau

balgrist_murnau$studyPlot <- "Balgrist/Murnau"

balgrist_murnau_baseline <- balgrist_murnau %>%
  group_by(id) %>%
  # sort from smallest to largest interval
  arrange(Interval_injury_UEMS) %>%
  # get the first measurement for each patient
  slice(1) %>%
  # ungroup
  ungroup()

xyplot(as.formula(paste("uems ~", biomarker, "| studyPlot")),
  data = balgrist_murnau_baseline,
  group = id,
  type = "b",
  pch = 16,
  cex = 1.2,
  lwd = 2,
  col = rgb(.1, .1, .1, .5),
```

```
xlab = biomarker_name_upper,
ylab = "Upper Extremity Motor Score",
grid = TRUE,
      scales = list(
    x = list(alternating = 1), # Show x-axis tick labels only at the bottom
    tck = c(1, 0) # Remove top ticks
  ),
      panel = function(x, y, ...) {
    panel.xyplot(x, y, ...) # Plot the points
    panel.lmline(x, y, col = "black", lwd = 1) # Add a linear regression line
  }
)

## fit GAM Balgrist/Murnau

par(mfrow = c(1, 3))

# Define the formula
gam_formula <- as.formula(paste("uems ~ s(", biomarker, ") + s(tmstd) + ti(", biomarker, ", tmstd) +
                               s(id, bs = 're') + s(id, tmstd, bs = 're')"))

# Fit the GAM
gam_balgrist_murnau <- gam(gam_formula,
                          data = balgrist_murnau, method = "REML")

plot(gam_balgrist_murnau, scheme = 1, select = 1,
     seWithMean = TRUE, rug = TRUE,
     xlab = biomarker_name_upper, ylab = "Upper Extremity Motor Score",
     cex.lab = 1.5, cex.axis = 1.5)

plot(gam_balgrist_murnau, scheme = 1, select = 2,
     seWithMean = TRUE, rug = TRUE,
     xlab = "Time (standardized)", ylab = "Upper Extremity Motor Score",
     cex.lab = 1.5, cex.axis = 1.5)

plot(gam_balgrist_murnau, scheme = 1, select = 3,
     seWithMean = TRUE, rug = TRUE, main = "UEMS",
     xlab = biomarker_name_upper, ylab = "Time (standardized)", ticktype = "detailed",
     cex.lab = 1.45, cex.axis = 1.3, cex.main = 1.45, theta = 45, # Rotate around the z-axis
     phi = 25) # Adjust elevation angle (tilt the plot vertically)

## make predictions Balgrist/Murnau

biomarker_values <- seq(0, 7, by = 0.5)

# predict recovery trajectories based on different values of the biomarker
```

```

models <- list(gam_balgrist_murnau)

# Prediction function using biomarker variable
make_prediction <- function(biomarker_value, model) {
  # new_data includes time at baseline up to 6 months
  # (used standardized time = 1.1 just for visualization)
  new_data <- expand.grid(id = 1, tmstd = seq(0, 1.1, length = 100))
  # current biomarker value is added to the new data
  # current "biomarker" is defined before the code chunk
  new_data[[biomarker]] <- biomarker_value

  # Make predictions excluding random effects
  prediction <- data.frame(
    tmstd = new_data$tmstd,
    pred = predict(model, newdata = new_data, exclude = "s(id)")
  )
  return(prediction)
}

# Set up colors for plotting
colors_plot <- sequential_hcl(length(biomarker_values), palette = "Viridis")

legend_key <- list(
  text = list(labels = paste0(format(biomarker_values, nsmall = 1))), # Add descriptive labels
  lines = list(col = colors_plot, lwd = 2), # Corresponding colors and line widths
  space = "right", # Place the legend to the right of the plot
  columns = 1, # Single column layout for vertical alignment
  title = "Midsagittal TB", # Clear and concise title
  cex.title = 0.9, # Slightly smaller title for compactness
  cex = 0.8 # Slightly smaller text for the labels
)

# Plot the raw recovery data per treatment group and
# add the predictions for different values of the biomarker
xyplot(uems ~ tmstd | studyPlot,
  data = balgrist_murnau,
  group = id,
  xlim = c(-0.1, 1.3),
  type = "b",
  between = list(x = 1, y = 1),
  pch = 20,
  cex = .7,
  lwd = 3,
  col = rgb(.1, .1, .1, alpha = 0.05), # Fade primary data lines
  xlab = "Time (standardized)",
  ylab = "Upper Extremity Motor Score",
  #main = paste("Predictions for different", biomarker_name, "values"),
  key = legend_key,
  scales = list(
    x = list(alternating = 1), # Show x-axis tick labels only at the bottom
    tck = c(1, 0) # Remove top ticks
  ),
  # add predictions
  panel = function(x, y, subscripts, ...) {

```



```

panel.xyplot(x, y, subscripts = subscripts, ...)
# panel_index: only 1 in this case
panel_index <- panel.number()

# for each biomarker value, make predictions and add to plot
for (i in seq_along(biomarker_values)) {
  biomarker_value <- biomarker_values[i]
  # Generate predictions
  preds <- make_prediction(biomarker_value = biomarker_value, model = models[[panel_index]])
  panel.lines(preds$tmstd, preds$pred, col = colors_plot[i], alpha = 0.6, lwd = 2)
}
})

## delta UEMS for NISCI & Balgrist Murnau

legend_location <- "topleft"

# Fit GAM for NISCI again

# fit GAM for each treatment group separately

# Define the formula
gam_formula <- as.formula(paste("uems ~ s(", biomarker, ") + s(tmstd) + ti(", biomarker, ", tmstd) +
                               s(id, bs = 're') + s(id, tmstd, bs = 're')"))

# Fit the GAM
gam_placebo <- gam(gam_formula,
                  data = datc_placebo, method = "REML")

gam_treatment <- gam(gam_formula,
                   data = datc_treatment, method = "REML")

# this chunk generates predictions for baseline UEMS and 6-month UEMS for different values
# of the biomarker and plots the delta-UEMS over the range of the biomarker

# number of predicted delta-UEMS
N <- 100

# predictions made along the sequence of min(biomarker_value) to max(biomarker_value) of the biomarker
prediction_sequence <- seq(from = min(biomarker_values), to = max(biomarker_values), length.out = N)

# new_data for tmstd=0 and tmstd=1
new_data <- data.frame(id = rep(1, 2 * N), tmstd = rep(c(0,1), each = N))

# add biomarker value sequence to new data
new_data[[biomarker]] <- rep(prediction_sequence, 2)

# Generate predicted UEMS for placebo, treatment and Balgrist/Murnau groups at tmstd=0 and tmstd=1

```

```

predictions <- data.frame(
  tmstd = new_data$tmstd,
  biomarker = new_data[[biomarker]],
  pred_placebo = predict(gam_placebo, newdata = new_data, exclude = "s(id)",
  pred_treatment = predict(gam_treatment, newdata = new_data, exclude = "s(id)",
  pred_balgrist_murnau = predict(gam_balgrist_murnau, newdata = new_data, exclude = "s(id)")
)

# Calculate the predicted delta change between tmstd = 1 and tmstd = 0 for each biomarker value
predicted_delta <- data.frame(
  biomarker_value = prediction_sequence,
  delta_placebo = predictions[predictions$tmstd == 1,]$pred_placebo
- predictions[predictions$tmstd == 0,]$pred_placebo,
  delta_treatment = predictions[predictions$tmstd == 1,]$pred_treatment
- predictions[predictions$tmstd == 0,]$pred_treatment,
  delta_balgrist_murnau = predictions[predictions$tmstd == 1,]$pred_balgrist_murnau
- predictions[predictions$tmstd == 0,]$pred_balgrist_murnau
)

# Plot the predicted delta-UEMS for each biomarker value
# first empty plot to set the limits
plot(predicted_delta$biomarker_value, predicted_delta$delta_balgrist_murnau, type = "l",
  col = "black", lwd = 2, lty = 3,
  ylim = c(min(predicted_delta[,2:4])-1, max(predicted_delta[,2:4])+1),
  xlab = biomarker_name_upper,
  ylab = expression("Predicted " * Delta * "-UEMS")
  #, main = bquote("Predicted " * Delta * "-UEMS for different" ~ .(biomarker_name) ~ "values")
)

# select only values < 4 for NISCI, because no observations above 4 in placebo
filtered_data <- predicted_delta[predicted_delta$biomarker_value < 4, ]

# Add line for the placebo group for TB values < 4
lines(filtered_data$biomarker_value, filtered_data$delta_placebo, col = "black", lwd = 2, lty = 1)

# Add line for the treatment group
lines(filtered_data$biomarker_value, filtered_data$delta_treatment, col = "black", lwd = 2, lty = 2)

# Add line for balgrist murnau
lines(predicted_delta$biomarker_value, predicted_delta$delta_balgrist_murnau, col = "black",
  lwd = 2, lty = 3)

# Add legend
legend(legend_location,
  legend = c("Placebo", "NG101", "Balgrist/Murnau"),
  col = "black",
  lwd = 2,
  lty = c(1, 2, 3),
  bty = "n")

# alternatively with colors: col = c("#69b3a2", "#f4a582", "#8c6bb1"), lwd = 2

# Find the turningpoint of Balgrist/Murnau

```

```
diff <- (c(predicted_delta$delta_balgrist_murnau,0)-c(0, predicted_delta$delta_balgrist_murnau) )

first_negative_index <- which(diff < 0)[1]
balgrist_murnau_turnpoint <- predicted_delta$biomarker_value[first_negative_index]

#####
# code for results: relationship of biomarkers with recovery (call of the code chunks)
#####

# not explicitly showed here because the prepared code chunks would be displayed each time again.

# midsagittal
# <<quartile_plot_other_biomarker>>
# <<baseline_uems_plot_tissue_mid>>
# <<plot_gam_fit>>
# <<plot_gam_fit_NG101>>
# <<plot_gam_predictions>>
# <<plot_gam_delta>>

# parasagittal
# <<quartile_plot_para>>
# <<baseline_uems_plot_tissue_para>>
# <<plot_gam_fit_para>>
# <<plot_gam_fit_NG101_para>>
# <<plot_gam_predictions_para>>
# <<plot_gam_delta_para>>

# lesion volume
# <<quartile_plot_lesion>>
# <<baseline_uems_plot_tissue_lesion>>
# <<plot_gam_fit_lesion>>
# <<plot_gam_fit_NG101_lesion>>
# <<plot_gam_predictions_lesion>>
# <<plot_gam_delta_lesion>>
```