Teeth Asymmetry After MARPE

Written by: Morgan Wood   Client: Richard Andrews

30 April, 2024

Abstract

This report investigates teeth asymmetry resulting from miniscrew-assisted rapid palatal expansion (MARPE) treatment, focusing on changes in left versus right sides of patients' teeth after treatment. Utilizing Euclidean distances and rotational angles derived from landmarks that were tracked pre- and post-treatment, this study employs statistical tests and linear regression models to analyze asymmetry. Findings indicate low asymmetry in movement and angles, with no significant evidence of substantial asymmetry. Notably, further examination using linear regression reveals the effects of teeth extractions on asymmetry in movement, suggesting avenues for future research.

1. Overview

This report focuses on teeth asymmetry caused by MARPE, a non-surgical treatment used to expand the upper jaw. Within this report, we focus on the asymmetry of each patient's left side change versus right side change resulting from treatment. This change is measured in both Euclidean distance and in angles which were provided by the client.

The necessary data for this planned analysis, descriptive variables of each patient, euclidean distances, and angles, were extracted from the excel workbook proved. The data used can be found in the Github repository, link here, under the excel sheet 'demographics.xlxs', 't1_minus_t2.xlsx', and 'angles.xlxs', respectively. The data found here is the corrected data discussed through email. Also as discussed, all R code used in the analysis can be found on the Github repository along with comments on how to adapt the code for future data. A quick description of the Rmarkdown files can be found in Appendix C.

To measure the change after treatment in both the left and right side of each patients' teeth, landmarks (22 on the left side and a corresponding 22 on the right side) were identified and tracked. This results in 22 Euclidean distances and three (rotational) angles for each side. Our plan of analysis was to analyze Euclidean distances and angles. Since the spreadsheet provided was much larger than what we discussed, I extracted the Euclidean distances and the angles that were in this larger spreadsheet and worked with this extracted data.

Data from 26 patients is analyzed which consists of, for each patient,

- descriptive variables: age pre-treatment, age post-treatment, sex, extractions of teeth (converted to: 'yes' or 'no'), No. of turns, PMD left, PMD right, left posterior TAD #cortices, right posterior TAD #cortices,
- euclidean distances for 22 pairs of landmarks (44 Euclidean distances total), and
- three paired angle (rotation) measurements (6 angles total).

We begin with looking at asymmetry in Euclidean distances in Section 2 and then look at asymmetry in angles in Section 3. Each section looks first to measure the amount of average asymmetry using either statistical tests or confidence intervals. Each identifying of a landmark is known to possibly be inaccurate. Thus, some measured asymmetry between the left and right side may be due to measurement error rather than true asymmetry. Thus, we will be testing for *substantial asymmetry* (formally defined in Section 2 for Euclidean distances and Section 3 for angles) where asymmetry will be considered substantial if it is past a certain threshold. Finally, using linear regression, asymmetry is modeled as a linear function of the descriptive variables for each patient using linear regression. This allows us to explore, for each landmark, the effect of the patent's descriptive variables (explanatory variables) on their asymmetry (response variable).

This report concludes with a three-part Appendix. Appendix A contains extra figures which are referenced during the main report. Appendix B contains a discussion of the permutation tests we originally planned and gives reasons why they were not carried out. Also in Appendix B is additional discussion on some concerns that were discussed through email. Finally, Appendix C contains a short description of the files in the Github repository.

2. Asymmetry in Movement Change

For each patient's landmark, the total change in location is recorded as a Euclidean distance. For example, if a landmark was at location (x_0, y_0, z_0) pre-treatment and (x_1, y_1, z_1) post-treatment, the Euclidean distance for that landmark is recorded as

$$d = \sqrt{(x_1 - x_0)^2 + (y_1 - y_0)^2 + (z_1 - z_0)^2},$$

which we refer to as *movement* of that landmark.

As landmarks come in pairs (left and right side), we record 22 total pairs of movement. We write each pair of movement as $(d^l_{j,n}, d^r_{j,n})$ for patient n's jth landmark.

The *asymmetry in movement* of landmark j for patient n is the absolute difference of this movements

$$A_{j,n} = \left| d^l_{j,n} - d^r_{j,n} \right| .$$

Asymmetry is considered substantial if it is more than 2mm. We will test if there is substantial asymmetry in Section 2.1. In Section 2.2, we will model this asymmetry using linear regression as a function of each patient's descriptive variables.

2.1 Testing for Substantial Asymmetry in Movement

As discussed in Section 1, there is believed to be some uncertainty in the identification of landmarks, and thus uncertainty in the measurement of movements. Because of this, we agreed that the threshold for substantial asymmetry would be 2mm.

In Figure 1 below, asymmetry is explored graphically. The landmark pairs with the lowest and highest average asymmetry are displayed. This corresponds to the landmarks labeled "l3_cusp" and "u4_ectopremolare", respectively. Each point on the figure corresponds to a patient with the x-coordinate depicting left side movement and the y-coordinate depicting right side movement.

Within Figure 1, dotted lines represent different thresholds of asymmetry. The green line represents 0mm of asymmetry (no asymmetry), the orange line represents 1mm of asymmetry, and the red line represents 2mm of asymmetry (the agreed upon threshold for substantial asymmetry). Corresponding figures for all landmarks can be found in Appendix A.

Landmarks that have observations that are close to the green line (observations with high correlation) correspond to landmarks with low asymmetry. An interesting future analysis would be to look at the average asymmetry of each landmark (or the correlation values of the corresponding scatterplots above) versus the location of each landmark. Through this analysis, the impact of landmark location (for example, near the front of the mouth or near the back of the mouth) on the landmark's asymmetry could be explored.

We see that few observations make the 2mm threshold for substantial asymmetry. In fact, we find that every landmark pair has average asymmetry less than 2mm. The highest average asymmetry was from the landmark pair "u4_ectopremolare" with average asymmetry of 1.53mm and the lowest average asymmetry was from the landmark pair "l3_cusp" with average asymmetry 0.34mm. Additionally, as asymmetry in movement is very low, this does suggest that the measurement errors are also low, perhaps lower than expected.

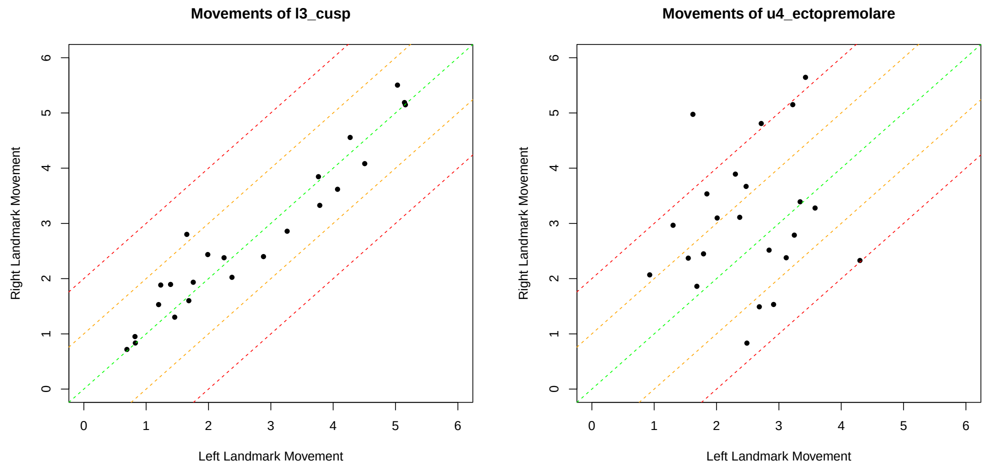


Figure 1: Left versus right side movement for landmarks with lowest average asymmetry (left) and highest average asymmetry (right). The green line represents 0mm of asymmetry, the orange line 1mm, and the red line 2mm. There are very few observations of substantial asymmetry, even when looking at the landmark of largest average asymmetry which is on the right.

As the average asymmetry for each sample is less than 2mm, we will not find significant evidence that there is substantial asymmetry for any landmark pair.

Because of this, we decide to test if the average asymmetry is substantial where the threshold for substantial asymmetry is reduced to 1mm. However, we find that there is still not significant evidence. The lowest p-value when testing if the average asymmetry is above 1mm is 0.0048. While this p-value is below the typical threshold of 0.05, as we are conducting 22 tests (each corresponding to a different landmark), we must adjust the typical threshold p-value for significance from 0.05 to 22 times 0.05 which is 0.0023 if we wish to test at the 95% significance level. Thus, any test with a p-value less than 0.0023 would result in the conclusion of significant asymmetry in our setting. As no p-value is below 0.0023, we do not find significant evidence of substantial asymmetry at the 1mm level for any landmark.

2.2 Modeling Movement Asymmetry

We next consider modeling asymmetry in movement as a function of patients' descriptive variables. To do this, each landmark pair's asymmetry of movement is modeled separately using a linear model. In these linear models, we consider the response variable of a patient's landmark's movement asymmetry and explanatory variables corresponding to a patient's descriptive variables. Below are the descriptive variables considered.

- age pre-treatment, age post-treatment,
- sex,
- extractions of teeth (converted to: 'yes' or 'no'),
- No. of turns,
- PMD left, PMD right,
- left posterior TAD #cortices, and right posterior TAD #cortices.

It is worth noting that the age pre-treatment and age post-treatment are highly correlated (r=0.999). To deal with this correlation within the model, an interaction term between the ages was added. An interaction term was added to help capture the length of treatment along with the age at time of treatment. An alternative fix to the correlation would be to transform the variables in order to provide measurements of age pre-treatment and treatment length while removing the variable corresponding to age post-treatment.

Additionally, the categorical variables PMD left & right (in agreement in 25 out of 26 patients), and left & right posterior TAD #cortices (in agreement in 20 out of 26 patients) are also correlated. For these variables, the right measurements were removed from the model.

We find that, in general, a linear model based on these descriptive variables is not a good fit for modeling asymmetry. There were no models with a p-value under 0.01, with the majority of p-values over 0.4. A full table of p-values is listed below in Table 1. After adjusting for multiple tests, a p-value of 0.0023 is desirable evidence for a good fit. Full details on the model can be found in the "planned_analysis_euclidean_distances.Rmd" file.

landmark	p-values of linear model
ag_l_dist	0.81995
ans_l_dist	0.19632
jg_l_dist	0.58327
k_l_dist	0.098
l3_l_cusp_dist	0.59749
l4_l_cg_dist	0.40237
l6_l_cg_dist	0.01148
or_l_dist	0.09993
pns_l_dist	0.16317
u1_l_apex_dist	0.53189
u1_l_mi_dist	0.17594
u3_l_apex_dist	0.47617
u3_l_cusp_dist	0.3122
u3_l_ectocanine_dist	0.21058
u4_l_pcusp_dist	0.69407
u4_l_ectopremolare_dist	0.64815
u4_l_papex_dist	0.56995
u6_l_ectomolare_dist	0.56842
u6_l_furc_dist	0.89185
u6_l_mbcusp_dist	0.13214
u6_l_pcusp_dist	0.11901
zf_l_dist	0.52048

Table 1: P-values for each linear model with response being movement asymmetry and explanatory variables corresponding to patient descriptive variables. None of the linear models are significant at the 95% level after adjusting for multiple tests.

Given the insignificance of the previous models, we next attempt to remove variables that do not explain the asymmetry in order to get a better model. To perform variable selection, note that it is intuitive that asymmetry should have common significant predictors regardless of the specific landmark. That is, the descriptive variables that are predictive for the asymmetry of one landmark should also be predictive for the other landmarks. This suggests that we should look for a common model where the covariates are allowed to differ, but the variables used should be the same. To do this, we use a backwards selection technique based on the average p-values of covariates across all models. This technique suggests removing the age variables.

After modeling without the age variables, the linear model was still not significant for the majority of landmarks. However, for one landmark pair, "l6_cg", the linear model explains the asymmetry well (according to a model p-value of 0.0008). A list of all p-values can be found below in Table 2.

landmark	p-values of linear model
ag_l_dist	0.60182
ans_l_dist	0.54135
jg_l_dist	0.43815
k_l_dist	0.03217

l3_l_cusp_dist	0.4474
l4_l_cg_dist	0.31117
l6_l_cg_dist	0.00084
or_l_dist	0.54671
pns_l_dist	0.03901
u1_l_apex_dist	0.51783
u1_l_mi_dist	0.29206
u3_l_apex_dist	0.35667
u3_l_cusp_dist	0.11153
u3_l_ectocanine_dist	0.06646
u4_l_pcusp_dist	0.43301
u4_l_ectopremolare_dist	0.53456
u4_l_papex_dist	0.43028
u6_l_ectomolare_dist	0.39743
u6_l_furc_dist	0.66893
u6_l_mbcusp_dist	0.34087
u6_l_pcusp_dist	0.31139
zf_l_dist	0.37063

Table 2: P-values for each linear model with response being movement asymmetry and explanatory variables corresponding to patient descriptive variables not including age variables. Only the model corresponding to the landmark "l6_cg" is significant at the 95% level after adjusting for multiple tests.

Coefficient values along with their p-values for the linear model corresponding to landmark "l6_cg" can be found in Table 3. For categorical variables, the coefficient can be interpreted as the average increase in asymmetry if the variable's value was the one listed versus the unlisted value. For example, the coefficient corresponding to "'Ext or Non'yes" gives the average increase in asymmetry of a value of "yes" (i.e., an extraction occurred) versus a value of "no" in the variable "Ext or Non". Similarly, the coefficient corresponding to "'PMD left'1" gives the average increase in asymmetry of a value of "1" versus a value of "0" and "'PMD left'2" gives the average increase in asymmetry of a value of "2" versus a value of "0" in the variable "PMD left". Finding the average effect of having a value of "1" versus "2" can be found by subtracting these two coefficients. The coefficient corresponding to "sex1" gives the average increase in asymmetry when male ("1") versus female ("0"). Finally, the coefficient corresponding to "'left posterior TAD #cortices'2" gives the average increase in asymmetry of a value of "2" versus a value of "1 in the variable"left posterior TAD #cortices".

Significant effects were found from extractions and number of turns, with an extraction increasing asymmetry on average by 0.5 mm and number of turns decreasing asymmetry by 0.03 mm per turn. However, across the linear models for all landmarks, the modeled effect of an extraction increased asymmetry on average by 0.28 mm and number of turns increased asymmetry by 0.01 mm per turn. Table 3 lists the average values of each coefficient across all models along with coefficients from the model with response of asymmetry in landmark "l6_cg". Based on the agreement on effect of extractions and disagreement on effect of number of turns, this suggests the effect of extractions on asymmetry in movement is worth further investigation.

explanatory variable	estimate of coefficent	p-values	average coefficent across all models
sex1	0.14	0.20652	-0.18
'Ext or Non'yes	0.5	0.00027	0.28
'No. of turns'	-0.03	0.00092	0.01
'PMD left'1	0.41	0.05762	-0.13
'PMD left'2	0.18	0.15782	0.06
'left posterior TAD #cortices'2	0.07	0.55532	-0.22

Table 3: Estimates of the coefficient and p-value for each explanatory variable according to the linear model of movement asymmetry in landmark "l6_cg" along with the average value of each estimated coefficient across

all linear models. We find extractions and of number of turns to each have a significant effect on asymmetry in landmark “l6_cg”. However, the coefficient estimate using this model and the average coefficient across all models disagree in direction of effect (decrease versus increase in asymmetry, respectively) for number of turns. They agree on the direction of effect (increase in asymmetry) of extractions.

3. Asymmetry in Angle Change

We next look at the asymmetry in angles. To do this, we use the three pairs of angles provided, which are each based on three landmarks

- angle for each side based on ans, pns, & jg,
- angle for each side based on u6_furc, u6_mbcusp, & u6_ectomolare, and
- angle for each side based on u6_furc, u6_pcusp, & u6_ectomolare.

We refer to the absolute difference between the left and right corresponding angle as *angle asymmetry*.

3.1 Significant Angle Asymmetry

We do not have a predetermined threshold for substantial angle symmetry (as we do in Section 2 for movement asymmetry). However, there is still error assumed to be in the angle measurements which will cause a non-negative asymmetry even for angles that are perfectly symmetric. Thus, in this section, we define substantial angle asymmetry to be asymmetry more than 5 degrees and, along with testing for substantial angle asymmetry, we also give confidence intervals for the average asymmetry for each angle. We define this arbitrary threshold of substantial asymmetry, 5 degrees, as a starting point in the analysis. The code provided in the Github repository can be adapted to a more appropriate threshold, if one exists.

Figure 2 gives a plot of the angle measurements for the left side versus the right side. As before, dashed lines represent different levels of asymmetry, with the green line represents 0 degrees of asymmetry (no asymmetry), the orange line represents 5 degrees of asymmetry (substantial asymmetry), and the red line represents 15 degrees of asymmetry.

As seen in Figure 2, there are two outliers. We suggest checking these values to be sure that they are true extremes in the data rather than data entry errors.

It is also worth noting that the magnitude of asymmetry appears to differ depending on the angle pair. In fact, we have the following 95% confidence interval for each respective mean,

- asymmetry of angles based on ans, pns, & jg: (0.88, 1.96),
- asymmetry of angles based on u6_furc, u6_mbcusp, & u6_ectomolare: (2.49, 6.45), and
- asymmetry of angles based on u6_furc, u6_pcusp, & u6_ectomolare: (2.48, 8.33).

Finally, we can test to see if the average asymmetry for any angle pair is substantial (i.e., more than 5 degrees). We find that there is not sufficient evidence to claim that the average angle asymmetry is above 5 degrees for any of the three angle pairs with p-values of 1, 0.708, and 0.390 for each respective angle pair.

3.2 Modeling Angle Asymmetry

Similar to before, we can model the angle asymmetry as a linear function of patient descriptive variables using linear regression. However, as before, the linear models are not significant at predicting angle asymmetry. In fact, even after removing age from the model, no model is significant (model p-values are 0.147, 0.263, and 0.376, and no dependent variables are significant at the 0.05 level in any model). Details on all linear regression models can be found in the Rmarkdown file “planned_analysis_angles.Rmd”.

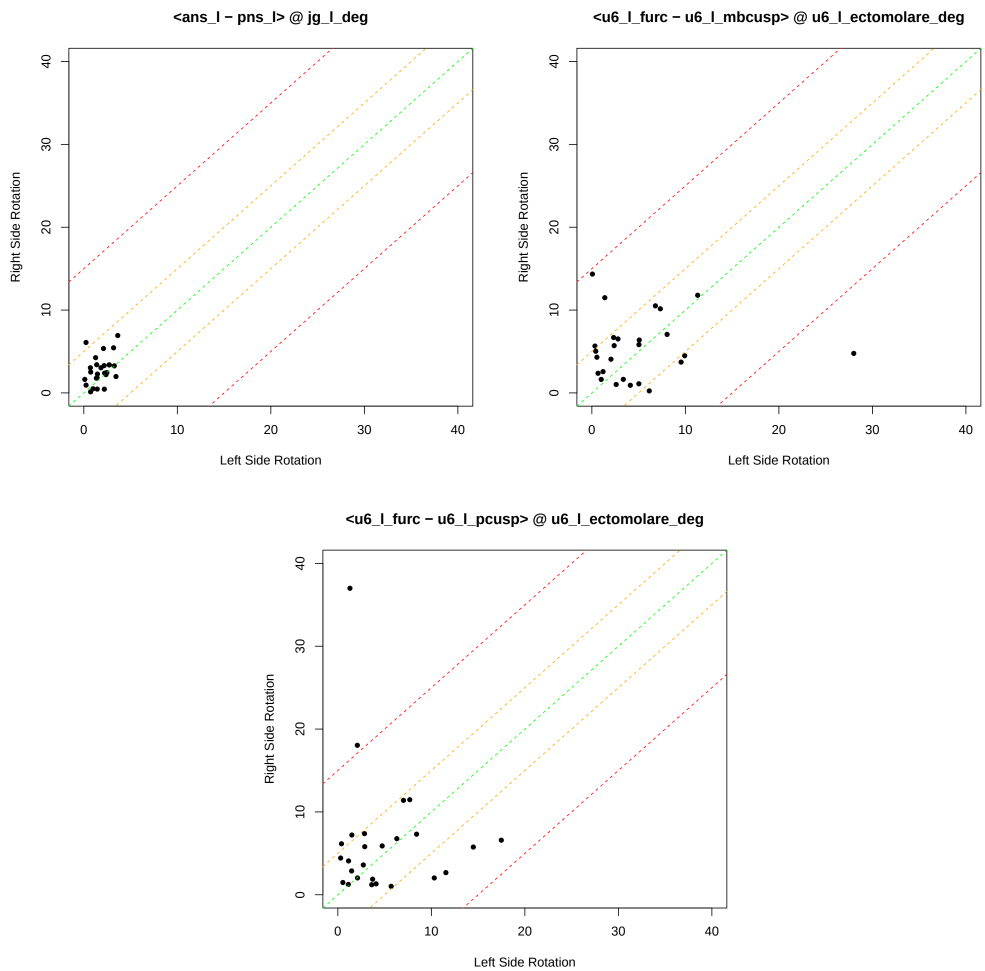


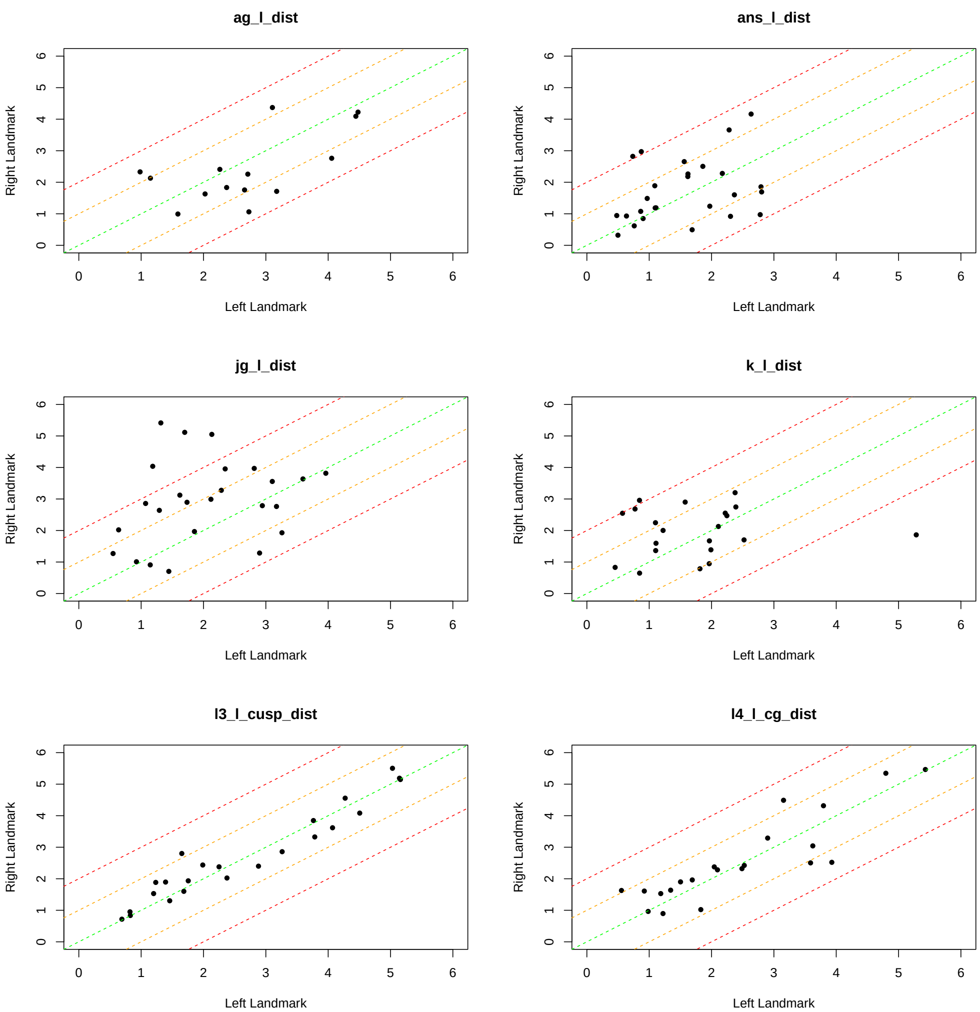
Figure 2: Left versus right side angles. The green line represents 0 degrees of asymmetry, the orange line 5 degrees, and the red line 15 degrees.

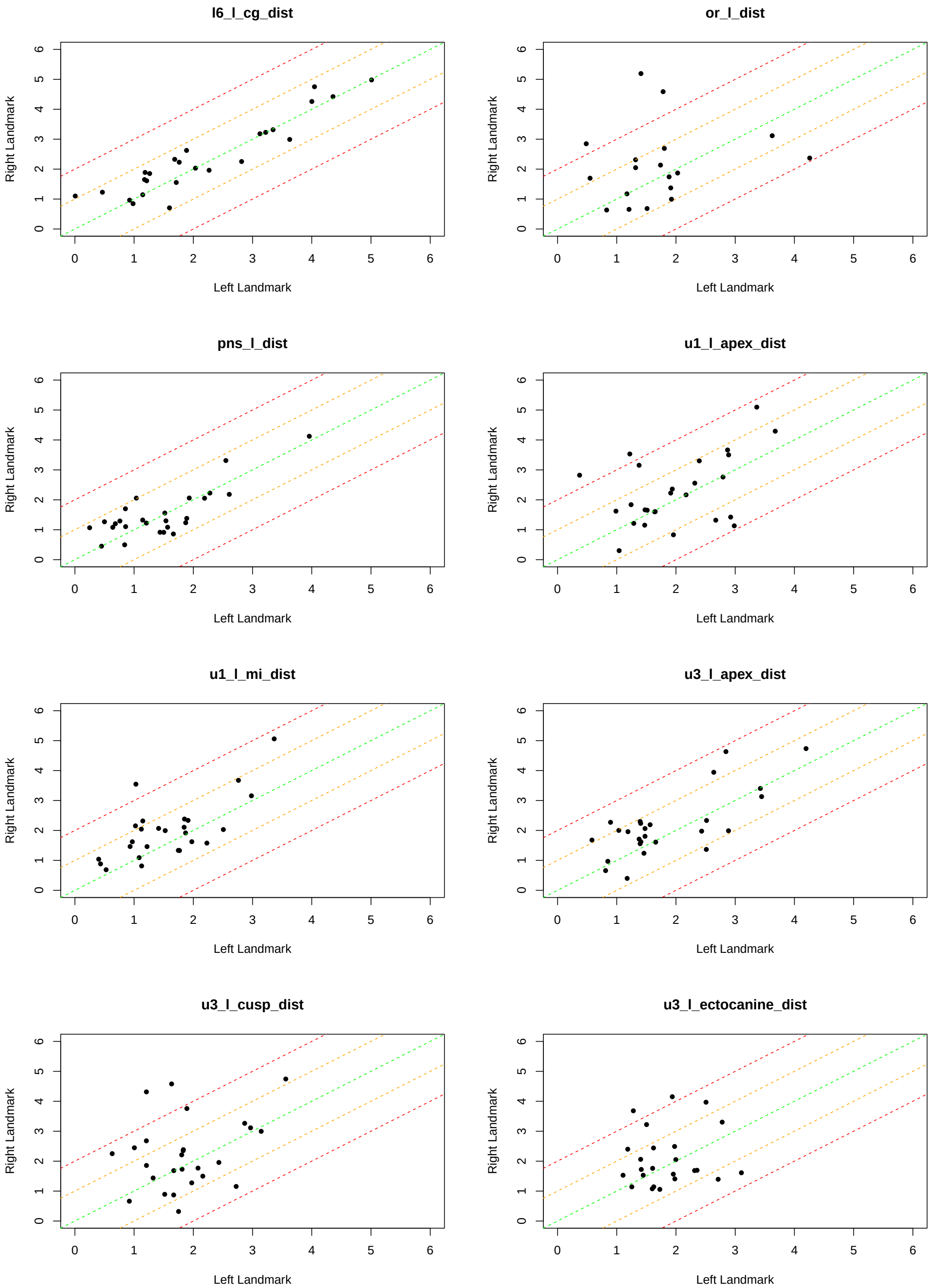
4. ChatGPT Statement

ChatGPT was used to develop a draft for the abstract.

I asked ChatGPT "Will you help me write an abstract for a statistical report?", and then I fed it the entire report. It didn't seem to mind that I copy-and-pasted from a pdf and the structure was messed up. It produced a long (three paragraph) abstract. I asked "Can you give me a shorter abstract that is only a paragraph". Which it did, and then I edited it to get the abstract here. The general abstract structure was provided by ChatGPT, however I changed the wording in most of the sentences. Interestingly, ChatGPT made up what MARPE stood for using context clues, "Maxillary Skeletal Expansion" instead of the correct "miniscrew-assisted rapid palatal expansion".

Appendix A




l6_l_cg_dist
or_l_dist
pns_l_dist
u1_l_apex_dist
u1_l_mi_dist
u3_l_apex_dist
u3_l_cusp_dist
u3_l_ectocanine_dist
Right Landmark
Left Landmark

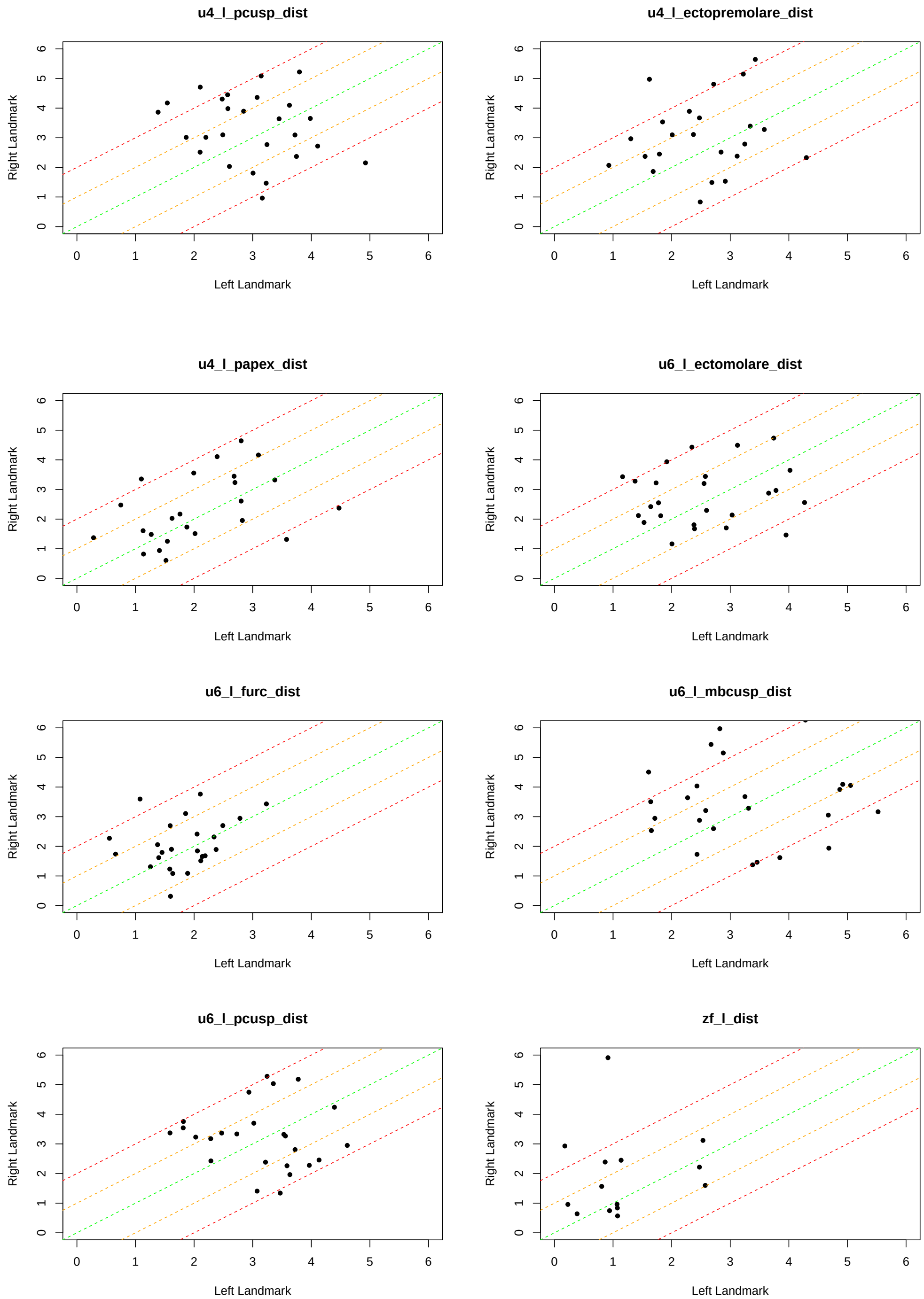


Figure 3: Left versus right side movement for all landmarks. The green line represents 0mm of asymmetry, the orange line 1mm, and the red line 2mm.

Appendix B

The original plan of analysis listed testing to see if the average movement of the left side is equal to the average movement of the right side (as a vector analysis). However, this analysis would not capture the desired characteristic of ''asymmetry". It is possible for the average movement of the left and right side to be the same (i.e. $(d^l_{j,\cdot})_{j=1,\ldots,22} = (d^r_{j,\cdot})_{j=1,\ldots,22}$) , but for them to not be aligned in such a way that the movement is symmetric (i.e., $(A_{j,\cdot})_{j=1,\ldots,22} = 0$) which is what we would like to test for. This is the "washing out" that we discussed in emails. We avoid this washing out by testing on observations of asymmetries $A_{j,n}$ rather than observations of movements $d^l_{j,\cdot}$ and $d^r_{j,\cdot}$.

Further, there is ambiguity in what should be tested. Two options would be to either test if the movement is asymmetric (i.e. $(A_{j,\cdot})_{j=1,\ldots,22} \neq 0$) or to test if the asymmetry is more than 2mm different (substantial asymmetry) in every landmark (i.e. $(A_{j,\cdot})_{j=1,\ldots,22} \geq \bar{2}$). By the same argument in Section 1, because of measurement errors and practical significance, I would consider the second a better option. However, as the sample averages for each of the landmarks are all under 2, testing if there is substantial asymmetry will result in not enough evidence to conclude that there is asymmetry just as in Section 2.

The same discussion above can be applied to the angle measurements. However, it is worth noting that most tests that look at these measurements as a vector (rather than separately as we did in the main report), gain their power from assuming equal variance. However, looking at Figure 2, it appears that this assumption is not valid when it comes to the angle asymmetry measurements.

Appendix C

The Github repository can be found at https://github.com/smithmor/STOR765_Spring2024_Asymmetry _Project.git.

This repository contains the cleaned data as three excel files 'demographics.xlxs', 't1_minus_t2.xlsx', and 'angles.xlxs' which are referred to and used in the r-code.

The r-code used for analysis can be found in the files 'eda_euclidean_distances.Rmd', 'planned_analysis_euclidean_distances.R 'eda_angles.Rmd', 'planned_analysis_angles.Rmd'. Each ''eda" file (exploratory data analysis) contains a quick exploratory analysis of the data including some summary statistics and comments on how the formal analysis can be expected to perform. The formal analysis can be found in the"planned_analysis" files. All files contain comments on why certain steps were done along and should be a quick adaption to new datasets.

Finally, there are files used to create and display this report within the repository 'asymmetry_report.Rmd' and 'asymmetry_report.pdf', respectively.