Thing 1 vs. Thing 2 - 2024 Scanner Comparison

Ryan Gallagher

2025-04-11

Analysis Description

The objective of this report is to determine whether there is a measurable difference between the scanners Thing 1 & Thing 2. Three metrics are provided: `MAPD`, `SNPQC`, `Waviness.SD` as well as identifiers for scanner. This report will summarize these metrics per scanner then runs statistical tests to determine if the groups are significantly different.

The data used in this report are:

- `Clinical_604_052824_Thing1.txt` and `Clinical_604_052824_Thing2.txt` for May 28th, 2024
- `Clinical_618_110824.txt` and `Clinical_618_110824_2nd scan.txt` for November 8th, 2024

The data in these files are combined to form data table for this report. The data set consists of two separate array batches, 604 and 618. Each of these batches were run on difference days, however they were scanned consecutively on Thing 1 and Thing 2 on the same day. Each batch consisted of 11 patient samples and 1 control (`CytoRef103`).

Summary Statistics

The objective of this section is to describe the data and determine the distribution within each group to then allow us to compare the data across scanners.

For each metric within each group, we summarize the data and test whether they approximately follow a normal distribution using the Shapiro-Wilks Test. Our interpretation is that if the p-value is **less than 0.05 then we reject the null that our data is approximately normal**.

Table 1: Summary of Metrics by Scanner with Shapiro-Wilk Test Results

		MAPD			**SNPQC**			**Waviness SD**		
Scanner	N	Mean	SD	Shapiro p-value	Mean	SD	Shapiro p-value	Mean	SD	Shapiro p-value
Thing 1	24	0.17	0.01	0.65	22.09	2.81	0.07	0.09	0.01	0.92
Thing 2	24	0.17	0.01	0.82	22.27	2.97	0.09	0.09	0.01	0.87

We find that all metrics report a Shapiro-Wilk p-value greater than 0.05. Thus, **we accept the null** that our data is approximately normal. This will aid in accepting assumptions necessary to compare our scanner groups - Thing 1 vs. Thing 2.

We'll run a Bartlett test for homogeneity of variances to determine whether the variance is roughly equal for the same metric across groups. Our interpretation is that if the Bartlett p-value is **less than 0.05 then we reject the null that the variances in each of the groups are the same**.

Table 2: Bartlett's Test for Equality of Variances

Variable	Bartlett Statistic	p-value
MAPD	0.02	0.88
SNPQC	0.07	0.79
Waviness.SD	0.00	0.99

We find that our p-values are greater than 0.05, and **we can accept the null** hypothesis that the variances for each metric are the same across groups. This will allow us to accept assumptions of the statistical test in the next section.

Two-Sample T-Test for Difference of Groups

We have satisfied the following assumptions for each metric in our data:

- Data is normally distributed within each group
- Observations are independent
- Variance is homogeneous between groups

Thus, the Two-Sample T-Test will be employed to determine whether the the means of the two groups are equal. Our interpretation is that if the Two-Sample T-Test p-value is **less than 0.05 then we reject the null that there is no significant difference between the means of the two groups**.

Table 3: Two-Sample T-Test Results

Variable	T Statistic	P-Value
MAPD	0.78	0.44
SNPQC	-0.21	0.83
Waviness.SD	-0.11	0.91

We find that our p-values are greater than 0.05, and **we can accept the null hypothesis that there is no significant difference between the means of the two groups**.

Conclusion

Based on the metrics `MAPD`, `SNPQC`, and `Waviness SD`, **we are confidence that Thing 1 and Thing 2 demonstrate comparable performance.**