FIGHT RP Extension Data

July 02, 2024

# Original Data Received

* On January 26, 2024, Muhammad Jehanzeb Khan sent an Excel sheet containing the FIGHT RP extension study data.
* Total 30 consented patients.
* Total 12 visit times (including Baseline visit and all follow up visit). Month0, Month3, Month6, Month9, Month12, Month15, Month18, Month21, Month24, Month33, Month42, Month51
* Month0 represents the baseline visit, while Month3, Month6, Month9, Month12, Month15, Month18, Month21, Month24, Month33, Month42, and Month51 correspond to follow-up visits.
* Data contains variables: patient\_id, Visit\_Date, VA\_OD, VA\_OS, MAIA\_Avg\_threshold\_OD, MAIA\_Avg\_threshold\_OS, EZ\_width\_OD, EZ\_width\_OS, Time

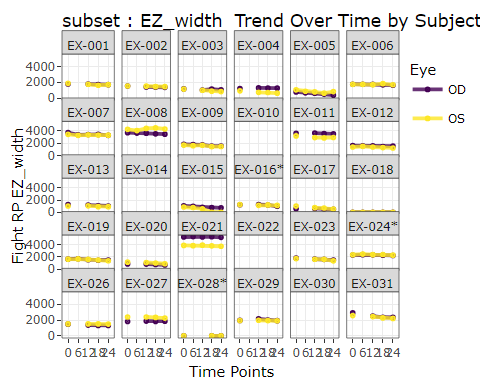
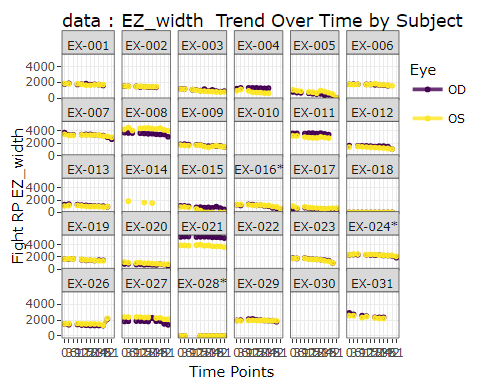
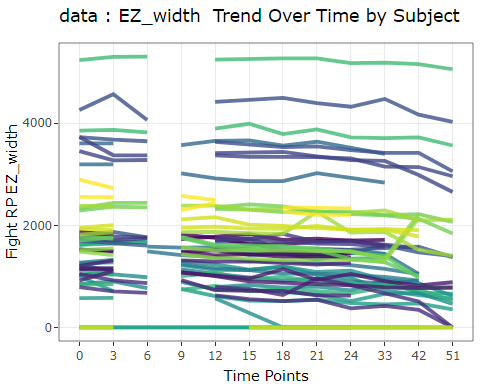
# Data Clarification and Correction: Addressing Outlier Values and Data Errors

* Biostatistician identified potential issues in the data;
* Jehanzeb responded and acknowledged the issues, providing the following updates:
  + For Subject ID EX-004, the correct EZ width in the OS should be 716 μm, as 7161 μm was an error.
  + For Subject ID EX-024, the correct MAIA Avg threshold in the OS should be 14.7 dB, as 2414 was a data entry error.
  + For Patient ID EX-010 initially lacked baseline data but later some data became available, the correction was made to remove the follow-up data based on email received Jan 31,2024.

# Data Cleaning and Processing

* The data was transformed from wide to long format, with each patient having rows for left and right eye measurements.
* Ineligible eyes are excluded in the samples;
  + \*EX-014 Data available for left eye only;
  + \*EX-030 Data available for right eye only;

# Utilize Visulization technics to understand missing for each time point and use this information to deteimine the best time to include in the analysis.



# To understand the missing data patterns in the dataset for EZ\_width, we counted the number of missing as well as the percentage of missing for all consented patients.

* The table shows the missing number and missing percentage for EZ\_width for each time point among all consented patients.

| Time | n\_missing\_EZ\_width | total\_EZ\_width | proportion\_missing |
| --- | --- | --- | --- |
| 0 | 8 | 60 | 13.33% |
| 3 | 9 | 60 | 15% |
| 6 | 38 | 60 | 63.33% |
| 9 | 28 | 60 | 46.67% |
| 12 | 10 | 60 | 16.67% |
| 15 | 9 | 60 | 15% |
| 18 | 8 | 60 | 13.33% |
| 21 | 8 | 60 | 13.33% |
| 24 | 8 | 60 | 13.33% |
| 33 | 15 | 60 | 25% |
| 42 | 21 | 60 | 35% |
| 51 | 34 | 60 | 56.67% |

* Time 0 represent for the baseline visit. other numeric number represent the month of the follow up visit. (i.e. 3 is Month 3 Followup time point).
* Based on the missing data patterns for EZ\_width, We selected baseline to 24 month data to calculate the AAC and change of MAIA as well as VA, 6 month gap between each visit.
* Below is a summary of missing {outcome variable} among eligible patients.

**Summary of missing EZ among eligible patients**

| **Characteristic** | **OD**, N = 29 | **OS**, N = 29 |
| --- | --- | --- |
| **time0\_missing** |  |  |
| Missing | 3 (10%) | 3 (10%) |
| Non-Missing | 26 (90%) | 26 (90%) |
| **time24\_missing** |  |  |
| Missing | 3 (10%) | 3 (10%) |
| Non-Missing | 26 (90%) | 26 (90%) |
| **time\_both\_missing** |  |  |
| Missing | 3 (10%) | 3 (10%) |
| Non-Missing | 26 (90%) | 26 (90%) |

**Summary of missing MAIA among eligible patients**

| **Characteristic** | **OD**, N = 29 | **OS**, N = 29 |
| --- | --- | --- |
| **time0\_missing** |  |  |
| Missing | 3 (10%) | 3 (10%) |
| Non-Missing | 26 (90%) | 26 (90%) |
| **time24\_missing** |  |  |
| Missing | 3 (10%) | 3 (10%) |
| Non-Missing | 26 (90%) | 26 (90%) |
| **time\_both\_missing** |  |  |
| Missing | 3 (10%) | 3 (10%) |
| Non-Missing | 26 (90%) | 26 (90%) |

**Summary of missing VA among eligible patients**

| **Characteristic** | **OD**, N = 29 | **OS**, N = 29 |
| --- | --- | --- |
| **time0\_missing** |  |  |
| Missing | 0 (0%) | 1 (3.4%) |
| Non-Missing | 29 (100%) | 28 (97%) |
| **time24\_missing** |  |  |
| Missing | 0 (0%) | 1 (3.4%) |
| Non-Missing | 29 (100%) | 28 (97%) |
| **time\_both\_missing** |  |  |
| Missing | 0 (0%) | 1 (3.4%) |
| Non-Missing | 29 (100%) | 28 (97%) |

# Missing Data Imputation for for calculating EZ width AAC

* Filling Missing EZ Width Values: We used linear interpolation to fill in missing Ellipsoid Zone (EZ) width values in the dataset.
* The key steps were:
* Identified patients with available data group by (patient\_id,Eye).
* Applied na.approx() function to linearly interpolate missing EZ width values.
* Updated the EZ width column with the imputed values. Notes: For Subject EX-015 as an example, the last two visits had an EZ width of 0, which was kept as the actual measurement, not imputed.
* After filling in missing values for EZ, the following patients EX-001, EX-002, EX-003, EX-004, EX-005, EX-006, EX-007, EX-008, EX-009, EX-010, EX-011, EX-012, EX-013, EX-014, EX-015, EX-016*, EX-017, EX-018, EX-019, EX-020, EX-021, EX-022, EX-023, EX-024*, EX-026, EX-027, EX-028\*, EX-029, EX-030, EX-031 who don’t have baseline and follow-up EZ measurements are considered ineligible samples for calculating the AAC.Thus, we exclude these patients for AAC calculation.
* For patients ‘EX-018’ and ’EX-028\*’ where EZ baseline until all follow-up visits have a value of 0, we exclude these ineligible samples from the AAC calculation as well.

# Area Above Curve with outcome measurements observed at pre-specified follow-up visits. The AAC can be empirically and non-parametrically estimated as the sum of all the trapezoids determined by the measurements at all visits.

To estimate the AAC, the area can be non-parametrically estimated as the sum of trapezoids (Let y\_i be the outcome variable measured at visit j, j = 0,1,2…K ; The interval length, denoted as t;

# General calculations for the AAC are conducted. Following this, the between-eye correlation analysis for the AAC of EZ-width is performed.

Protocol Addendum to be added as 6.1.2.4

Although EZ loss is expected to be a degenerative process, it is unknown whether EZ width at a follow-up visit may be greater than that at baseline due to measurement error or due to intervention effect. The AAC is the cumulative loss of EZ width overtime. Its calculation needs to consider if there was gain of EZ width during follow-up (i.e. negative loss). Section 6.1.2.3 derives the AAC calculation if y1 is increased compared to y0 . Here we provide an algorithm that derives the general AAC calculation without assuming whether a follow-up EZ-width measurement is greater than y0. general\_AAC\_calculation is a function will escalate for AAC

[1] “Between-eye correlation of the AAC of EZ-width:” Correlation Coefficient: 0.346757 95% Confidence Interval - Lower Bound: -0.0468935 95% Confidence Interval - Upper Bound: 0.6471815

| correlation | lower\_bound | upper\_bound |
| --- | --- | --- |
| 0.346757 | -0.0468935 | 0.6471815 |

[1] “Summary statistics for EZ width : Baseline until Month24”

| **Eligible Sample** | **N = 48***1* |
| --- | --- |
| **y0** | 1,921 (1,034); 577~5,240; 1,660 [1,195 - 2,303] |
| **y1** | 1,879 (1,020); 595~5,310; 1,656 [1,171 - 2,362] |
| **y2** | 1,848 (1,057); 573~5,251; 1,585 [1,116 - 2,333] |
| **y3** | 1,776 (1,085); 0~5,275; 1,494 [1,120 - 2,265] |
| **y4** | 1,722 (1,076); 0~5,182; 1,467 [1,010 - 2,209] |
| **AAC\_general** | 2,145 (2,520); -1,947~9,864; 1,797 [370 - 3,329] |
| **Eye** |  |
| OD | 24 (50%) |
| OS | 24 (50%) |
| *1*Mean (SD); Minimum~Maximum; Median [IQR]; n (%) | |

# Calculate Correlation and Summary Statistics(Mean and standard deviation) for Secondary Outcome(change\_MAIA,change\_VA)

[1] “change\_MAIA Correlation” Correlation Coefficient: 0.8285188 95% Confidence Interval - Lower Bound: 0.6496618 95% Confidence Interval - Upper Bound: 0.9204668

| correlation | lower\_bound | upper\_bound |
| --- | --- | --- |
| 0.8285188 | 0.6496618 | 0.9204668 |

[1] “MAIA Baseline Correlation” Correlation Coefficient: 0.9876187 95% Confidence Interval - Lower Bound: 0.9721802 95% Confidence Interval - Upper Bound: 0.9945135

| correlation | lower\_bound | upper\_bound |
| --- | --- | --- |
| 0.9876187 | 0.9721802 | 0.9945135 |

[1] “MAIA Month 24 Correlation” Correlation Coefficient: 0.9891309 95% Confidence Interval - Lower Bound: 0.9755547 95% Confidence Interval - Upper Bound: 0.9951856

| correlation | lower\_bound | upper\_bound |
| --- | --- | --- |
| 0.9891309 | 0.9755547 | 0.9951856 |

[1] ” ” [1] “change\_VA Correlation” Correlation Coefficient: 0.3242686 95% Confidence Interval - Lower Bound: -0.05552564 95% Confidence Interval - Upper Bound: 0.6220872

| correlation | lower\_bound | upper\_bound |
| --- | --- | --- |
| 0.3242686 | -0.0555256 | 0.6220872 |

[1] “VA Baseline Correlation” Correlation Coefficient: 0.851132 95% Confidence Interval - Lower Bound: 0.7004858 95% Confidence Interval - Upper Bound: 0.9291643

| correlation | lower\_bound | upper\_bound |
| --- | --- | --- |
| 0.851132 | 0.7004858 | 0.9291643 |

[1] “VA Month 24 Correlation” Correlation Coefficient: 0.8406198 95% Confidence Interval - Lower Bound: 0.6811679 95% Confidence Interval - Upper Bound: 0.9239358

| correlation | lower\_bound | upper\_bound |
| --- | --- | --- |
| 0.8406198 | 0.6811679 | 0.9239358 |

[1] “Summary statistics for MAIA”

| **Variable** | **N = 58***1* |
| --- | --- |
| **Eye** |  |
| OD | 29 (50%) |
| OS | 29 (50%) |
| **MAIA\_Month24** | 9 (8); 0~26; 8 [2 - 14] |
| Unknown | 6 |
| **MAIA\_Baseline** | 9 (8); 0~26; 8 [2 - 16] |
| Unknown | 6 |
| **change\_MAIA** | -0.53 (1.87); -8.60~3.30; -0.30 [-0.90 - 0.35] |
| Unknown | 6 |
| *1*n (%); Mean (SD); Minimum~Maximum; Median [IQR] | |

[1] “Summary statistics for VA”

| **Variable** | **N = 58***1* |
| --- | --- |
| **Eye** |  |
| OD | 29 (50%) |
| OS | 29 (50%) |
| **VA\_Month24** | 75 (12); 28~89; 78 [72 - 83] |
| Unknown | 1 |
| **VA\_Baseline** | 77 (9); 42~89; 79 [74 - 83] |
| Unknown | 1 |
| **change\_VA** | -1.4 (5.4); -30.0~11.0; 0.0 [-2.0 - 1.0] |
| Unknown | 1 |
| *1*n (%); Mean (SD); Minimum~Maximum; Median [IQR] | |

# Linear mixed-effects model summary for primary outcome EZ\_width

[1] “Summary of mixed-effects model for EZ\_width” Linear mixed model fit by REML. t-tests use Satterthwaite’s method [ lmerModLmerTest] Formula: EZ\_width ~ Time + (1 | patient\_id) Data: data

REML criterion at convergence: 7366.4

Scaled residuals: Min 1Q Median 3Q Max -3.3740 -0.4045 0.0239 0.3882 3.7423

Random effects: Groups Name Variance Std.Dev. patient\_id (Intercept) 1192427 1092.0  
Residual 56536 237.8  
Number of obs: 524, groups: patient\_id, 27

Fixed effects: Estimate Std. Error df t value Pr(>|t|)  
(Intercept) 1784.1731 210.9121 26.2577 8.459 5.65e-09  ***Time -7.3730 0.7697 496.1530 -9.580 < 2e-16***  — Signif. codes: 0 ‘***’ 0.001 ’****’ 0.01 ’*’ 0.05 ‘.’ 0.1 ’ ’ 1

Correlation of Fixed Effects: (Intr) Time -0.066

# GEE Model for secondary outcomes

**[1] “Summary of GEE model for MAIA change”**

Call: geeglm(formula = change\_MAIA ~ 1, family = gaussian, data = MAIA, id = patient\_id, corstr = “exchangeable”)

Coefficients: (Intercept) -0.5269231

Degrees of Freedom: 52 Total (i.e. Null); 51 Residual

Scale Link: identity Estimated Scale Parameters: [1] 3.425814

Correlation: Structure = exchangeable Link = identity Estimated Correlation Parameters: alpha 0.8181229

Number of clusters: 26 Maximum cluster size: 2

Call: geeglm(formula = change\_MAIA ~ 1, family = gaussian, data = MAIA, id = patient\_id, corstr = “exchangeable”)

Coefficients: Estimate Std.err Wald Pr(>|W|) (Intercept) -0.5269 0.3461 2.318 0.128

Correlation structure = exchangeable Estimated Scale Parameters:

Estimate Std.err

(Intercept) 3.426 2.074 Link = identity

Estimated Correlation Parameters: Estimate Std.err alpha 0.8181 0.1126 Number of clusters: 26 Maximum cluster size: 2

**[1] “Summary of GEE model for VA change”**

Call: geeglm(formula = change\_VA ~ 1, family = gaussian, data = VA, id = patient\_id, corstr = “exchangeable”)

Coefficients: (Intercept) -1.328

Degrees of Freedom: 57 Total (i.e. Null); 56 Residual

Scale Link: identity Estimated Scale Parameters: [1] 29

Correlation: Structure = exchangeable Link = identity Estimated Correlation Parameters: alpha 0.3019

Number of clusters: 29 Maximum cluster size: 2

Call: geeglm(formula = change\_VA ~ 1, family = gaussian, data = VA, id = patient\_id, corstr = “exchangeable”)

Coefficients: Estimate Std.err Wald Pr(>|W|) (Intercept) -1.33 0.81 2.69 0.1

Correlation structure = exchangeable Estimated Scale Parameters:

Estimate Std.err

(Intercept) 29 14.4 Link = identity

Estimated Correlation Parameters: Estimate Std.err alpha 0.302 0.122 Number of clusters: 29 Maximum cluster size: 2