### **Study Overview:**

This study examines vessel diameter and vascular permeability in mouse brain vasculature using intravital two-photon imaging (imaging over ~S1, down to ~0.3mm; T-Z stacks). The researchers compared two groups (control/sham and an experimental group/2XmTBI – imaged 3-6h post-injury/Sx) using various measurements of blood vessels.

### Methodology:

- 1. Imaging:
  - Used galvo-galvo Z-stacks with FITC dextran dye
  - 512x512 pixels (0.3 μm/pixel), Nikon 16X objective (NA 0.8) with 6X zoom
  - 800nm illumination, 50-60mW power output
  - Z-stacks: 5 μm steps, 200-300 μm depth
  - Time series: 10 seconds duration, 10 frames (1Hz) per Z-interval
- 2. Image Processing:
  - 3D median filter (2 pixels) to reduce PMT shot noise
  - Manual selection of line scan areas (14-50x6 μm), targeting smaller vessels (3-10 μm)
- 3. Data Analysis:
  - Custom Matlab and Python scripts
  - Calculated full-width half maximum (FWHM) for vessel diameter
- Measured Fv (interior vessel fluorescence), Fe (exterior fluorescence), and Fi (vessel wall fluorescence)
  - Calculated vessel wall permeability (Pv) as Fe/Fv
- 4. Statistical Analysis:
- Mann-Whitney U test (grpA v grpB):
- FWHM: U statistic = 7.0, p-value = 0.35238
- Fe/Fv: U statistic = 13.0, p-value = 0.91429

Both tests fail to reject the null hypothesis, indicating no significant differences between distributions.

- Kolmogorov-Smirnov test (grpA v grpB):
- FWHM: statistic = 0.0922, p-value = 0.6163
- Fe/Fv: statistic = 0.1918, p-value = 0.0167

The second test shows a significant difference (p < 0.05).

Sample sizes are provided: n=157 for one group (6 mice) and n=104 for another group (4 mice).

- Principal Component Analysis (PCA) on three variables: FeFv, Zmicrons, and mean FWHM ums
  - Mann-Whitney U tests and t-tests on PC scores between groups

### **Key Findings:**

- 1. PCA Results:
  - PC1 explains 52.94% of variance, PC2 explains 30.20%
  - PC1 strongly correlates with all three variables, especially FeFv and Zmicrons

### 2. Statistical Comparisons:

- No significant differences between groups for PC1 (p=0.422), PC2 (p=0.062), or PC3 (p=0.462)
- PC2 shows a trend towards significance (p=0.062)

### 3. Variable Relationships:

- Positive correlations between Zmicrons, mean\_FWHM\_ums, and FeFv
- FeFv and Zmicrons contribute most to data variation

### Page with "From 7-19-24" data:

This page presents statistical test results for two analyses:

Sample sizes are provided: n=157 for A group (6 mice) and n=104 for B group (4 mice).

- 1. Mann-Whitney U test:
  - FWHM: U statistic = 7.0, p-value = 0.35238
  - Fe/Fv: U statistic = 13.0, p-value = 0.91429

Both tests fail to reject the null hypothesis, indicating no significant differences between distributions.

### 2. Kolmogorov-Smirnov test:

- FWHM: statistic = 0.0922, p-value = 0.6163
- Fe/Fv: statistic = 0.1918, p-value = 0.0167

The second test shows a significant difference (p < 0.05).

- K-S test result:
  - $\circ$  Test 2 shows a statistic of 0.1918 with a p-value of 0.0167, which is significant at the  $\alpha$  = 0.05 level.
- Interpretation: The K-S test compares the cumulative distribution functions of two samples. A significant result suggests that the two groups (likely control and experimental) have different distributions for the variable being tested.
- Importance: This significant difference is particularly interesting because it contrasts with the non-significant results from the Mann-Whitney U tests and the PCA analysis. It suggests that there might be subtle differences between the groups that aren't captured by measures of central tendency or linear combinations of variables.
- Potential implications: a) Distribution shape: The K-S test is sensitive to differences in the shape of distributions, not just central tendency. This could indicate that while the groups may have similar means or medians, they differ in terms of spread, skewness, or other distributional characteristics. b) Subset of data: Given that only one K-S test was significant, this might relate to a specific aspect of the vascular measurements (e.g.,

- vessel diameter, permeability, or depth) that shows group differences. c) Biological significance: In the context of brain trauma models, this could suggest subtle alterations in vascular properties that aren't uniform across all vessels or depths, but manifest as changes in the overall distribution of a particular measure.
- Considerations: a) Multiple comparisons: It's important to consider whether this p-value (0.0167) would remain significant after correcting for multiple comparisons, especially if many tests were performed. b) Effect size: While statistically significant, we should also consider the practical significance. The test statistic of 0.1918 suggests a moderate effect size.

### Page with "From 12-27-23 - 2D scatter plots":

- 1. Plot of Vessel diameter (FWHM, microns) vs. Normalized perivessel fluorescence intensity (mean Fe/Fv):
  - X-axis ranges from 10 to 60 (Fe/Fv)
  - Y-axis ranges from 2 to 10 microns (vessel diameter)
  - Two groups are shown: Group A (red) and Group B (blue)
  - Points are scattered without clear separation between groups
- 2. Plot of Z-depth (microns from surface) vs. Normalized perivessel fluorescence intensity (mean Fe/Fv):
  - X-axis ranges from 10 to 60 (Fe/Fv)
  - Y-axis ranges from 0 to 300 microns (Z-depth)
  - Same two groups are shown (A in red, B in blue)
  - Points are scattered across the plot, with some apparent clustering at certain depths

Both plots suggest overlapping distributions between the two groups, consistent with the statistical tests showing no significant differences in most comparisons.

### Page with PCA results:

- 1. A scatter plot of PC1 vs PC2:
  - Points are colored red and black, likely representing two groups
  - PC1 explains 52.94% of variance
  - PC2 explains 30.20% of variance
  - No clear separation between groups is visible
- 2. A bar plot showing correlations between PC1 and original variables:
  - Strong positive correlations for all three variables
  - FeFv and Zmicrons have the highest correlations (> 0.8)
  - mean\_FWHM\_ums has a moderate correlation (~ 0.5)
- 3. Statistical results comparing PC scores between groups:
  - All three PCs show non-significant differences (p > 0.05)
  - PC2 shows a trend towards significance (p = 0.062)

This analysis suggests that while there are strong relationships between the measured variables, there are no statistically significant differences between the two groups in terms of these principal components.

### Interpretation:

- 1. The lack of significant differences between groups suggests that the experimental condition did not produce large-scale changes in the measured vascular properties.
- 2. The strong correlations between variables might reflect underlying biological relationships in vascular structure and function. For example, vessel depth (Zmicrons) might be related to vessel size (FWHM) and blood flow characteristics (FeFv).
- 3. The trend towards significance in PC2 hints at possible subtle differences between groups, which might become significant with a larger sample size or more sensitive measurements.

#### **Limitations and Future Directions:**

- 1. The current analysis may not be sensitive enough to detect subtle changes in vascular properties following the experimental condition.
- 2. Including additional data such as X, Y coordinates and more detailed time series statistics could potentially increase the sensitivity of the analysis.
- 3. Using clustering algorithms (e.g., K-means or UMAP) on PC scores could reveal more complex patterns in the data.
- 4. Standardizing variables before PCA and carefully selecting the number of PCs to retain for clustering could improve the analysis.
- 5. Validating clustering results using domain knowledge and additional statistical tests would strengthen the findings.

**In conclusion,** while this study provides valuable insights into the relationships between vascular properties in the mouse brain, it does not demonstrate significant differences between the experimental groups.

Future work could focus on refining the analysis methods and increasing sample sizes to detect potentially subtle effects of the experimental condition on brain vasculature.

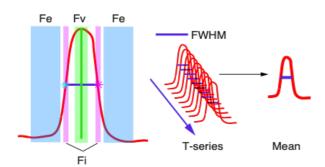
- → Classify by vessel type
- → Classify by ROI location
- → Use X,Y data
- → Future expts collect RBC velocity, neuro-vascular: combine w/ behavior (e.g., whisker-stim)

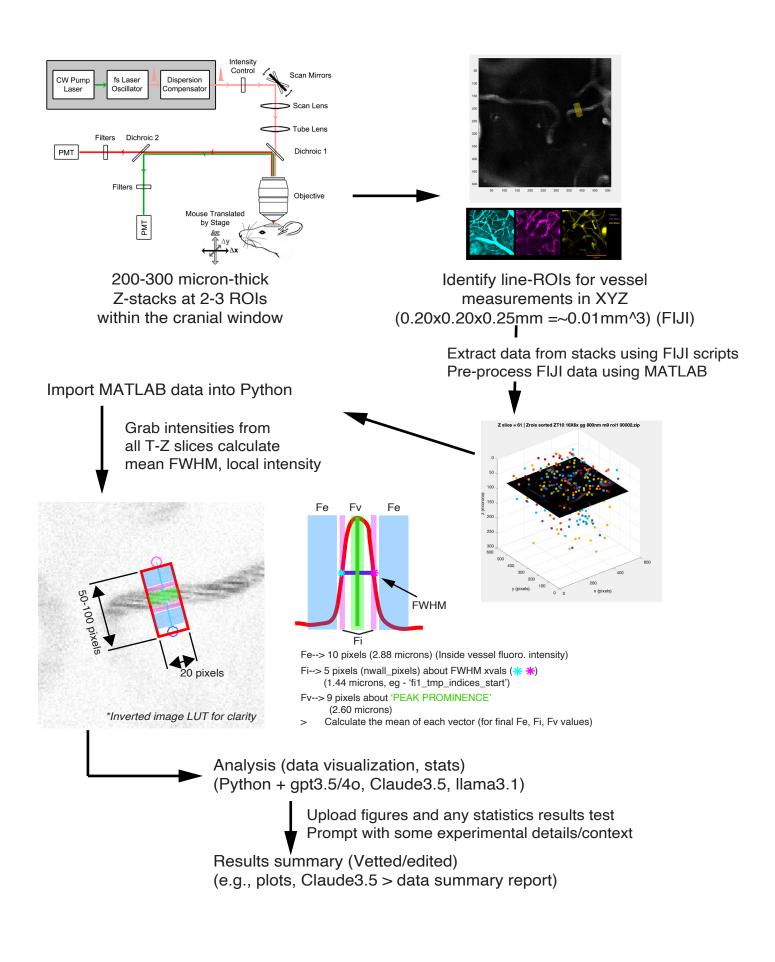
Measuring Vessel Diameter and Vascular Permeability (Pv) in Volumetric T-series

Galvo-galvo Z-stacks of the dextran dye (FITC) were acquired at 512x512 pixels (0.3 µm/pixel) using a Nikon 16X objective (NA 0.8) with 6X zoom, employing lambda=800nm illumination at a power output of 50-60mW. Z-stacks were captured at 5 µm step intervals, covering a depth ranging from approximately 200-300 µm. A time series (T) of 10 second duration was acquired at each Z-interval, comprising 10 frames (1Hz).

Prior to analysis, acquired image stacks (16-bit TIF files) underwent preprocessing to mitigate PMT shot noise using a 3D median filter (2 pixels, FIJI/ImageJ). Line scan areas measuring 14-50x6  $\mu$ m were manually selected throughout the Z-stack by an experimenter blinded to the experimental history, with a focus on targeting smaller diameter vessels (3-10  $\mu$ m).

Following selection and storage of line scan regions, TIF-image stacks were batch-processed, and intensity profiles were collected for each TIF, encompassing each group of 10 frames/every 5  $\mu$ m-Z-step. Custom-written Matlab and Python scripts were employed to determine the full-width half maximum (FWHM) for each line-profile, allowing estimation of vessel diameter (\*will work towards making code public access as appropriate via GitHub in the future). In addition, as shown in Figure 1, values for Fv, Fe, and Fi were calculated, where Fv is the mean fluorescence from the interior voxels, Fi is the total fluorescence from the vessel wall voxels, and Fe is the total fluorescence from all exterior points including those on the vessel wall (Fi). Vessel wall permeability (Pv) was calculated as the mean of Pv = Fe/Fv over 10sec (10 T-frames was subsequently calculated by dividing Fe (2x10 voxels) by Fv (9 voxels).





# Data = set1 - line ROIs by JEC (2023 v5 data)

Data from

/Users/jcoleman/Documents/--LARGE DATA--/#Pizzi/ .... TBI-sham leak analysis/process\_v4\_062723/process\_v5-2ii

```
ZT10_16X6x_gg_800nm_m5_roi2_00003.tif

ZT10_16X6x_gg_800nm_m6_roi1_00001.tif

ZT10_16X6x_gg_800nm_m7_roi1_00001.tif

ZT10_16X6x_gg_800nm_m9_roi1_00002.tif

ZT10_16X6x_gg_800nm_m10_roi1_00001.tif

ZT10_16X6x_gg_800nm_m11_roi1_00001.tif

ZT10_16X6x_gg_800nm_m12_roi1_00001.tif

ZT10_16X6x_gg_800nm_m13_roi1_00002.tif

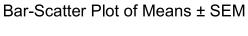
ZT10_16X6x_gg_800nm_m15_roi1_00001.tif

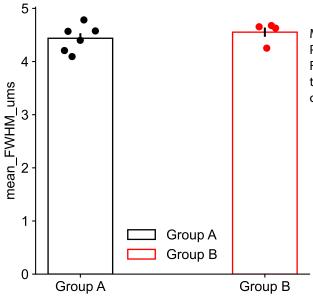
ZT10_16X6x_gg_800nm_m16_roi1_00001.tif
```

grpA = ['m5', 'm6', 'm9', 'm10', 'm11', 'm12'] grpB = ['m7', 'm13', 'm15', 'm16']

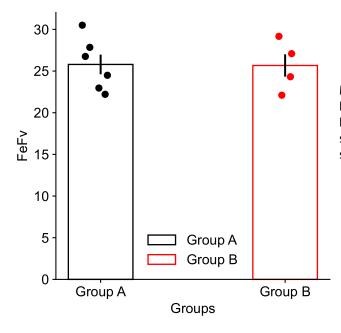
☐ Zrois_sorted_ZT10_16X6x_gg_800nm_m5_roi2_00003.zip	Jul 30, 2023 at 1:55 PM	10 KB	ZIP archive
Trois_sorted_ZT10_16X6x_gg_800nm_m5_roi2_00003roilNFO.CSV	Jul 31, 2023 at 1:47 PM	4 KB	CSV Document
Trois_sorted_ZT10_16X6x_gg_800nm_m6_roi1_00001.zip	Apr 26, 2023 at 9:00 AM	6 KB	ZIP archive
Zrois_sorted_ZT10_16X6x_gg_800nm_m6_roi1_00001roilNFO.CSV	Jul 31, 2023 at 1:48 PM	2 KB	CSV Document
Trois_sorted_ZT10_16X6x_gg_800nm_m7_roi1_00001.zip	Jul 29, 2023 at 6:38 PM	4 KB	ZIP archive
Zrois_sorted_ZT10_16X6x_gg_800nm_m7_roi1_00001roilNFO.CSV	Jul 31, 2023 at 1:49 PM	2 KB	CSV Document
Trois_sorted_ZT10_16X6x_gg_800nm_m9_roi1_00002.zip	Jul 29, 2023 at 6:53 PM	8 KB	ZIP archive
Zrois_sorted_ZT10_16X6x_gg_800nm_m9_roi1_00002roilNFO.CSV	Jul 31, 2023 at 1:50 PM	3 KB	CSV Document
Trois_sorted_ZT10_16X6x_gg_800nm_m10_roi1_00001.zip	Jul 30, 2023 at 1:22 PM	15 KB	ZIP archive
Zrois_sorted_ZT10_16X6x_gg_800nm_m10_roi1_00001roilNFO.CSV	Jul 31, 2023 at 1:38 PM	6 KB	CSV Document
Trois_sorted_ZT10_16X6x_gg_800nm_m11_roi1_00001.zip	Jul 29, 2023 at 7:11 PM	7 KB	ZIP archive
Zrois_sorted_ZT10_16X6x_gg_800nm_m11_roi1_00001roiINFO.CSV	Jul 31, 2023 at 1:40 PM	3 KB	CSV Document
Trois_sorted_ZT10_16X6x_gg_800nm_m12_roi1_00001.zip	Jul 29, 2023 at 7:26 PM	15 KB	ZIP archive
Zrois_sorted_ZT10_16X6x_gg_800nm_m12_roi1_00001roiINFO.CSV	Jul 31, 2023 at 1:41 PM	6 KB	CSV Document
Trois_sorted_ZT10_16X6x_gg_800nm_m13_roi1_00002.zip	Jul 30, 2023 at 10:42 AM	12 KB	ZIP archive
Zrois_sorted_ZT10_16X6x_gg_800nm_m13_roi1_00002roilNFO.CSV	Jul 31, 2023 at 1:43 PM	4 KB	CSV Document
Trois_sorted_ZT10_16X6x_gg_800nm_m15_roi1_00001.zip	Jul 30, 2023 at 11:34 AM	14 KB	ZIP archive
Zrois_sorted_ZT10_16X6x_gg_800nm_m15_roi1_00001roilNFO.CSV	Jul 31, 2023 at 1:44 PM	5 KB	CSV Document
Trois_sorted_ZT10_16X6x_gg_800nm_m16_roi1_00001.zip	Jul 30, 2023 at 2:04 PM	16 KB	ZIP archive
Zrois_sorted_ZT10_16X6x_gg_800nm_m16_roi1_00001roilNFO.CSV	Jul 31, 2023 at 1:46 PM	6 KB	CSV Document

means - take mean of all ROIs from all Z for each subject (n=1 mouse); CDF plots show the mean of each T-series for all ROIs from each subject (n=X rois from group1, etc)

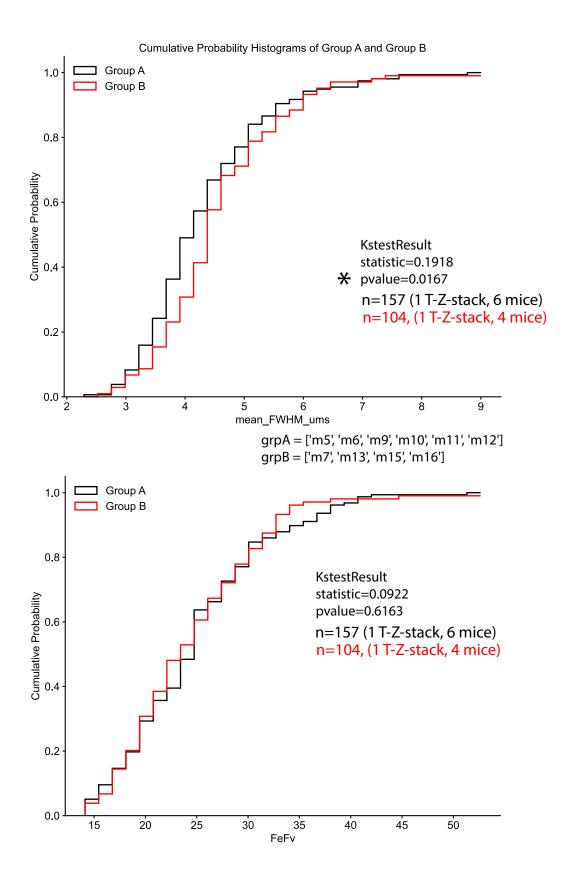




Mann-Whitney U statistic: 7.0 P-value: 0.35238095238095235 Fail to reject the null hypothesis: the distributions are not significantly different

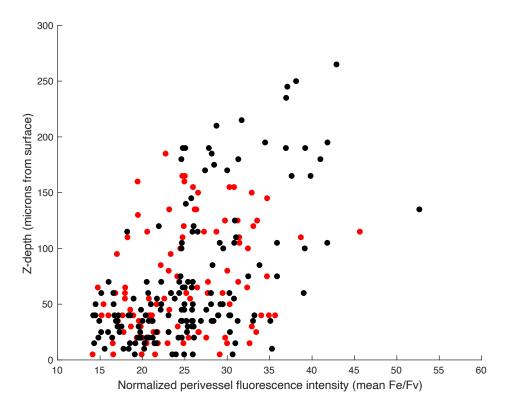


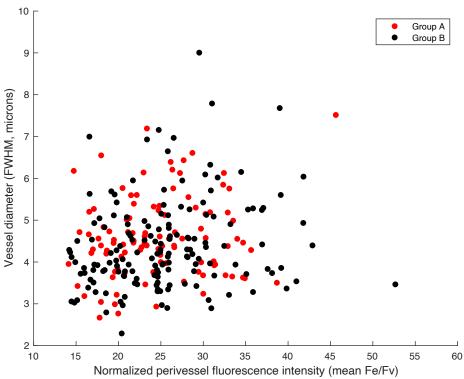
Mann-Whitney U statistic: 13.0 P-value: 0.9142857142857143 Fail to reject the null hypothesis: the distributions are not significantly different



From 12-27-23 - 2D scatter plots
Data = set1 - line ROIs by JEC (2023 v5 data)

- \*\* See APPENDIX 3D-Scatter Figures for
  - 1) location of ROIs in 3D (ie full XYZ data by subject/Zstack)
  - 2) Associated FWHM or Fe/Fv values with each ROI in 3D

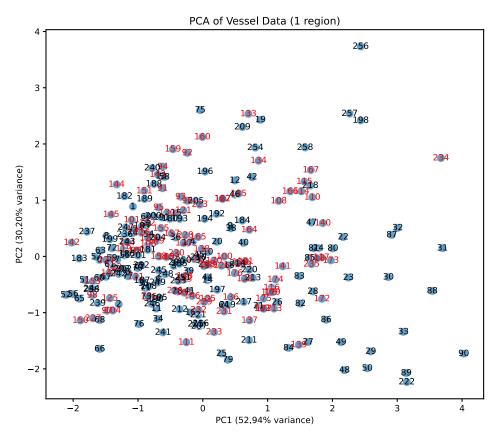




### PCA (w/ scaled data, 3 components, blind to group)

# Variables: FeFv | Zmicrons | mean FWHM ums

- \*\* See "APPENDIX Results 1" for more explanation and stats
- \*\*\* In the future, analyze w/ X, Y and trial with Tdata (over 10s; eg, mean, std, min, max, Fano factor)
- \*\*\*\* Could employ additional/alternative ways to visualize and 'segment' data (eg next step k-means clustering, UMAN)



Results for PC1:

Test: Mann-Whitney U test

Statistic: 7534.0

p-value: 0.4224659152628756

Results for PC2:

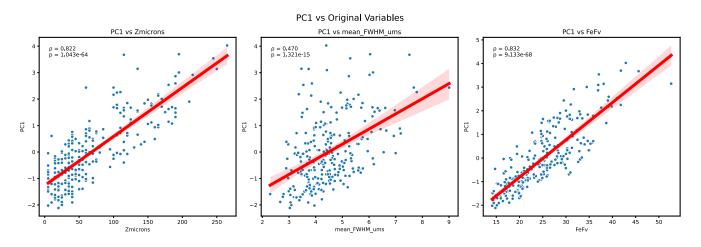
Test: Mann-Whitney U test

Statistic: 6909.0

p-value: 0.062434681758270716

Results for PC3: Test: t-test

Statistic: 0.7367184225588419 p-value: 0.4619655774546293



#### **PROMPT**

<Can you interpret these plots? And provide some synthesis of the results? For context, the data analyzed are from measurements in mouse brain vasculature imaged using intravital 2-photon imaging and are derived from ROIs in 3D image stacks acquired using galvo-galvo scanner and 6X zoom with a Nikon 16x lens. There are two groups represented (control-red and experimental-black) but the data were analyzed blind to group (only represented in the PCA plot by text color of ROI#). Here are stattisal results when comparing group PC scores:</p>

Results for PC1:

Test: Mann-Whitney U test

Statistic: 7534.0, p-value: 0.4224659152628756

Results for PC2:

Test: Mann-Whitney U test

Statistic: 6909.0, p-value: 0.062434681758270716

Results for PC3: Test: t-test

Statistic: 0.7367184225588419, p-value: 0.4619655774546293>

#### **RESPONSE**

Based on the provided plots and statistical results, here's an interpretation and synthesis of the data:

- 1. PCA Analysis (Figure 1):
  - PC1 explains 52.94% of the variance in the data.
  - PC2 explains 30.20% of the variance.
  - Together, PC1 and PC2 account for 83.14% of the total variance.
- The plot shows a spread of data points without clear clustering between the control (red) and experimental (black) groups.
- 2. PC1 vs Original Variables (Figure 2):
  - PC1 shows strong positive correlations with all three original variables:
  - a. Zmicrons ( $\rho = 0.822$ , p = 1.043e-64)
  - b. mean\_FWHM\_ums ( $\rho = 0.470$ , p = 1.321e-15)
  - c. FeFv ( $\rho = 0.832$ , p = 9.133e-68)
  - FeFv has the strongest correlation with PC1, followed closely by Zmicrons.
- 3. Statistical Comparisons:
- No significant differences were found between control and experimental groups for any of the principal components:
  - PC1: p = 0.422 (Mann-Whitney U test)
  - PC2: p = 0.062 (Mann-Whitney U test)
  - PC3: p = 0.462 (t-test)

### Synthesis:

- 1. Variable Relationships: The PCA reveals that the three measured variables (Zmicrons, mean\_FWHM\_ums, and FeFv) are positively correlated. This suggests that these aspects of mouse brain vasculature tend to vary together.
- 2. Dominant Features: FeFv and Zmicrons are most strongly associated with PC1, indicating they contribute most to the primary axis of variation in the data. This could suggest that differences in vessel depth (Zmicrons) and a measure related to blood flow or vessel function (FeFv) are the most distinguishing features among the samples.
- 3. Lack of Group Separation: The PCA plot and statistical tests indicate no significant differences between control and experimental groups across the principal components. This suggests that the experimental condition did not produce large-scale, consistent changes in the measured vascular properties.
- 4. Potential Subtle Effects: The near-significant p-value for PC2 (0.062) hints at possible subtle differences between groups, which might become significant with a larger sample size or more sensitive measurements.
- 5. Data Structure: The PCA effectively reduced the dimensionality of the data, with the first two components explaining over 80% of the variance. This suggests that much of the variation in the data can be captured by these two composite measures.
- 6. Biological Interpretation: The strong correlations between variables might reflect underlying biological relationships in vascular structure and function. For example, vessel depth (Zmicrons) might be related to vessel size (FWHM) and blood flow characteristics (FeFv).

In conclusion, while the analysis reveals interesting relationships between the measured vascular properties, it does not provide evidence for significant differences between the control and experimental groups in this mouse brain vasculature study. The results suggest that any effects of the experimental condition on these vascular properties, if present, are subtle and not easily distinguished using this PCA approach.

#### \*\*\*\*\*

Based on the information provided:

### 1. PC2 composition:

PC2 is more likely to be influenced by mean\_FWHM\_ums (vessel width) rather than Zmicrons or FeFv. This is because:

- Zmicrons and FeFv have very strong correlations with PC1 (0.822 and 0.832 respectively).
- The correlation between mean\_FWHM\_ums and PC1 is notably lower (0.470).
- PC2 often captures the next most important source of variation not explained by PC1.

### 2. Including additional data:

Yes, including X, Y coordinates and additional time series statistics could potentially increase the sensitivity and power of your analysis:

- Spatial information (X, Y) could reveal location-dependent patterns in the vasculature.
- Time series statistics (mean, stdev, min, max) for each metric at each Z-step could capture dynamic aspects of vessel behavior.
- This additional information might reveal subtle differences between groups that aren't apparent in the current analysis.
- However, be cautious of overfitting with too many variables. You may need to use feature selection techniques or regularization methods.

### 3. Using PC scores for clustering:

Yes, you can definitely use the PC scores as input for clustering algorithms like K-means or UMAP:

- This is a common approach in high-dimensional data analysis.
- PCA reduces dimensionality while preserving most of the variance, making it easier for clustering algorithms to find patterns.
  - K-means could be useful if you expect distinct, spherical clusters.
- UMAP might be better for capturing non-linear relationships and more complex cluster shapes.
  - Consider trying both and comparing results.

### Additional suggestions:

- 1. Standardize your variables before PCA if you haven't already, especially if including new variables with different scales.
- 2. Consider using scree plots or explained variance ratios to determine how many PCs to retain for clustering.
- 3. If using K-means, try different numbers of clusters and use methods like the elbow method or silhouette score to determine the optimal number.
- 4. For UMAP, experiment with different parameter settings (n\_neighbors, min\_dist) to find the most informative representation.
- 5. Validate your clustering results using domain knowledge and additional statistical tests.

Remember, while these techniques can reveal patterns in your data, interpreting the biological significance of these patterns requires careful consideration of the underlying physiology and experimental context.

#### \*\*\*\*\*

Including the grouping variable in the PCA analysis would not typically enhance the power or change the results in the way you might expect. Here's why:

#### 1. Nature of PCA:

PCA is an unsupervised method that doesn't consider group labels when determining the principal components. It focuses on maximizing variance across all data points, regardless of their group membership.

### 2. Potential Impact:

- If you include the grouping variable as a binary feature (e.g., 0 for control, 1 for experimental), it would be treated as just another variable in the PCA.
- This binary variable would likely have minimal impact on the overall variance structure, especially if the other variables have much larger ranges of values.
- It might appear as a small contribution to one of the later principal components, but it's unlikely to significantly alter the main PCs.

### 3. Appropriate Use of Group Information:

Instead of including the group variable in PCA, it's more common and usually more informative to:

- Perform PCA without the group variable
- Then analyze how the resulting PC scores differ between groups (as you've already done with your statistical tests)

### 4. Alternatives for Group-Aware Analysis:

If you want to incorporate group information more directly, consider:

- Linear Discriminant Analysis (LDA): This supervised method explicitly tries to find linear combinations of features that best separate the groups.
- Partial Least Squares Discriminant Analysis (PLS-DA): Similar to PCA but takes group information into account.
- ANOVA-simultaneous component analysis (ASCA): Combines ANOVA with PCA to analyze effects of experimental factors.

### 5. Enhancing Power:

To potentially enhance the power of your analysis:

- Ensure you have adequate sample sizes for each group
- Consider using more sophisticated statistical methods that can handle high-dimensional data and group comparisons simultaneously (e.g., MANOVA, permutation tests)
- Use dimension reduction techniques like PCA as a preprocessing step, then apply group comparison methods to the reduced data

#### 6. Visualization:

You can still use group information to enhance your visualization of PCA results, such as color-coding points by group in your scatter plots.

In summary, while including the grouping variable in PCA wouldn't directly enhance its power, there are other ways to leverage this information in your analysis. The approach you've taken - performing PCA and then comparing group differences in PC scores - is a valid and commonly

## APPENDIX – Results 1

used method. If you want to more directly incorporate group information into your dimension reduction, consider the alternative methods mentioned above.

# APPENDIX – 3D-Scatter Figures

Data plotted using 3D scatter script in Python. Each plot is one animal, ROI location data from 1 full TZ-stack. Specified variables are encoded via the heat plot (e.g., FWHM in microns)

