# Group 5: Thyroid Eye Disease Study - Final Report


Authors

Yi Ai Chen

Niamh Evenhouse

Mrinalini Garg

Kitu Komya

Caitlin Shener

Mohini Vembusubramanian


Stats 141SL with Professor Esfandiari

March 23, 2018

## Abstract

Our project examined anatomical features related to thyroid eye disease (TED). Two serious outcomes of TED are strabismus and optic neuropathy (ON), which lead to impaired vision and blindness. We attempted to identify relationships between features, and identify predictors of optic neuropathy and strabismus. The study included 24 patients who exhibited TED and 21 control patients. We recoded these observations as 48 eyes with TED and 21 control eyes due on the basis that the right and left eyes of subjects with TED are independent.

In total, the analysis included 13 predictor variables and 2 outcome variables (strabismus and ON).

Our analysis began with data exploration. We constructed histograms of the quantitative variables to understand their distributions. In addition, we cross-plotted densities of numerical variables between experimental and control groups. Then, we explored the correlations between predictors without controlling for other factors to understand the underlying relationships before delving into model-building. To test our assumption of the independence between the right and left eyes in the experimental group, we leveraged the non-parametric Spearman's coefficient.

We used logistic regression to study two outcome variables, strabismus and ON. We fit two models (one including all variables, and one with variable selection) to each outcome variable. No predictors were statistically significant. To overcome the small sample bias, we also used a penalized likelihood method. Again, no predictors were significant. While a larger sample size may have led to significance, it is also possible that the features included in our data are not significantly associated with strabismus and ON.

This possibility was supported by scatterplots examining the relationships between three key predictors: muscle volume, orbit volume, and fat volume. While there were some indications that these variables were relevant to understanding TED, there were no clear clusters between the different groups (control, ON present, ON not present).

We also used Wilcoxon tests to test associations of numerical variables with our outcome variables. We found that patients with ON had a higher ratio of fat volume to muscle volume than other patients, significant at 95% confidence.

In conclusion, our findings suggest that strabismus and ON may be mediated by factors not included in this study, and that ratio of fat to muscle volume may be associated with ON. In future studies, medical professionals should consider using a larger experimental group, measuring additional factors (such as physiological factors in the eye, gender, duration of TED, and thyroid medication dosage) and further exploring the significance of the ratio of fat to muscle volume.

## Clear statement of the problems to be solved

Thyroid eye disease can lead to both strabismus (double vision), and optic neuropathy, which leads to blindness. Unfortunately, it is still unclear to medical professionals which symptoms of TED have a significant impact on the development of these conditions. Our study examines thirteen variables to test whether any of them are significant predictors of optic neuropathy or strabismus.

## Variables of the study and how they were measured

Our study examined 15 variables -- 2 outcomes and 12 predictors. The two outcome variables were strabismus and optic neuropathy. Strabismus is double vision and we represented it as 1 if it was present in the patient and 0 if it was not present in the patient. Optic neuropathy leads to a patient going blind. Again we represented it as 1 if it was present in the patient and 0 if it was not present in the patient.

Our predictor variables were CAS score, lagophthalmos, Hertel, OSDI score, medial wall bowing, decompression, fat volume over orbit volume, muscle volume over orbit volume, and medial rectus muscle volume. CAS score is a measure of inflammatory activity in the eye on a scale from 0 to 10, where 0 is no inflammation and 10 is high inflammation. Lagophthalmos is the millimeter distance between upper and lower eyelid that remains when a person tries to close their eyes. Hertel is a measure of how far forward the eye has moved inside the eye socket in millimeters. OSDI score is a graded questionnaire given to patients to rate the impact of dry eyes on their lives. We treated it as a numeric variable on a scale from 0 to 100, where 0 represented normal impact and 100 represented severe impact. Medial wall bowing is a binary variable representing the presence of indentation on the orbit's medial wall. Decompression is a binary variable representing whether or not the patient goes on to have decompression surgery (measurements are taken before decompression). Medial rectus muscle volume is the volume of a particular muscle that presses on the medial wall. Fat and muscle volume represent the total volume of fat and muscle, respectively, inside the orbit. Orbit volume is the total volume of the eye socket.

We explored multiple representations of fat, muscle, and orbit volume due to interrelation between these three variables. Medical literature suggests dividing fat and muscle volume by orbit volume to neutralize gender differences in orbit volume. In addition to fat, muscle, and orbit volume, we included these ratios in our explorations, and opted to use the ratios along with orbit volume in our regressions. We included an additional variable, fat volume divided by muscle volume, in our final analysis.

## Exploratory data analysis

Contingency Tables

To take a brief look at our dataset before running any models or tests, we decided to perform some simple exploratory data analysis on the experimental group (patients with TED). We were firstly interested in the presence of optic neuropathy in the right and left eyes.

| | ON (No) | ON (Yes) | Total |
|---|---|---|---|
| Right (OD) | 18 | 6 | 24 |
| Left (OS) | 21 | 3 | 24 |
| Total | 39 | 9 | 48 |

*Table 1: Contingency table of ON by right and left eye*

From the very beginning we can see that we have a very small sample size (only 6 with ON in the right eye, and 3 with ON in the left eye). This made us very cautious with the type of conclusions that we can draw from our test and models. The following contingency tables also show this:

| | ON (No) | ON (Yes) | Total |
|---|---|---|---|
| Strabismus (No) | 16 | 4 | 20 |
| Strabismus (Yes) | 23 | 5 | 28 |
| Total | 39 | 9 | 48 |

*Table 2: Contingency table of ON by strabismus*

| | ON (No) | ON (Yes) | Total |
|---|---|---|---|
| Decompression (No) | 21 | 1 | 22 |
| Decompression (Yes) | 14 | 4 | 28 |
| Total | 35 | 5 | 40 |

*Table 3: Contingency table of ON by decompression*

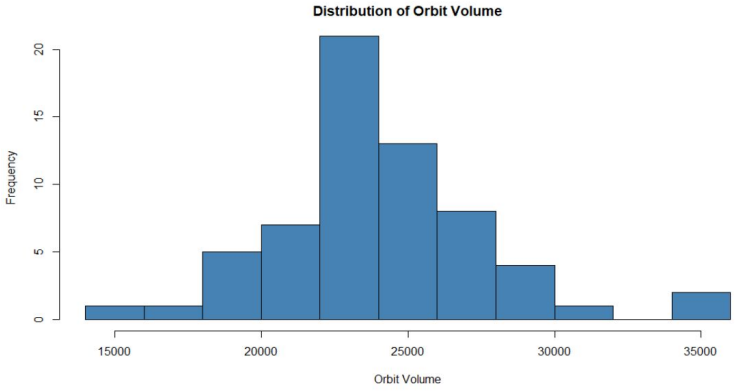
| | ON (No) | ON (Yes) | Total |
|---|---|---|---|
| Medial Bow (No) | 28 | 2 | 30 |
| Medial Bow (Yes) | 7 | 3 | 10 |
| Total | 35 | 5 | 40 |

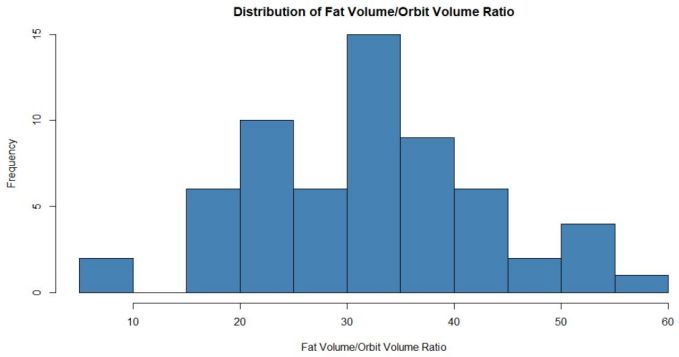
*Table 4: Contingency table of ON by medial wall bowing*

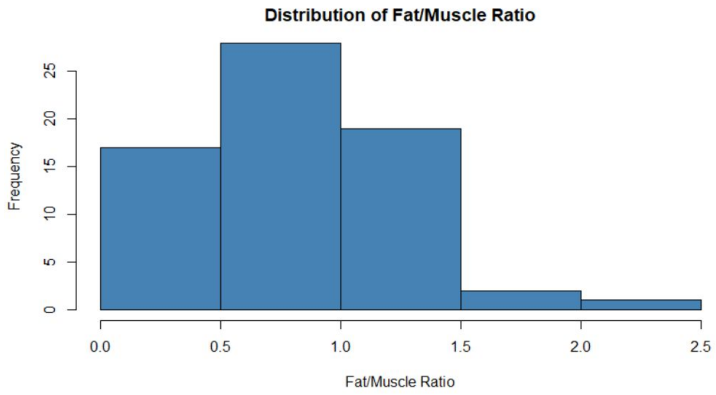
Because we have such a small sample size, we used Fisher's Exact Test to test independence between ON and decompression, ON and strabismus, and ON and medial wall bowing. The results showed that all p-values are insignificant; we cannot reject the null hypothesis of independence for all of these predictors and ON.

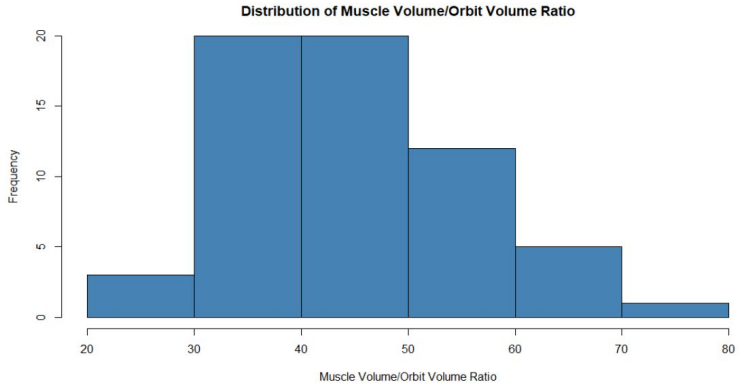
Histograms of Variables' Distributions

Now we move onto performing exploratory analysis on our numeric predictors: orbit volume, the fat volume/orbit volume ratio, the muscle volume/orbit volume ratio, and the fat volume/muscle volume ratio.




The orbit volume and fat volume/orbit volume are quite normally distributed, but the distributions of the muscle volume/orbit volume ratio and fat volume/muscle volume ratio are right-skewed.

Left and Right Eye Independence

While the left and right eyes of control patients are highly correlated, medical knowledge suggests that the right and left eyes of patients with TED can be considered independent, as each eye is affected differently by the disease. Making this assumption allows us to treat the left and right eyes as separate observations when conducting our statistical analysis, thereby doubling the experimental sample size. In order to test this assumption of independence we conducted some non-parametric tests of independence between measurements of TED patients' right and left eyes, calculating the Pearson correlation coefficient for the numerical measures like Hertel, lagophthalmos, muscle volume, and fat volume, and using the chi - squared test for categorical factors such as the presence of optical neuropathy and medial bowing.

| Variable | OS mean | OD Mean | OS SD | OD SD | Independent* |
|---|---|---|---|---|---|
| Lagophthalmos | 0.5 | 0.65 | 0.63 | 0.8 | ✓ |
| Hertel | 21.852 | 24.34 | 4.09 | 4.61 | ✓ |
| FV/OV | 34.95 | 30.76 | 12.84 | 11.38 | ✓ |
| MV/OV | 46.87 | 49.85 | 8.52 | 12.22 | ✓ |
| Medial Volume | 1904.09 | 2258.25 | 1080.69 | 1276.7 | ** |
| ON Proportion | 12.5% | 25% | 0.34 | 0.44 | ** |
| Medial Bowing | 20% | 30% | 0.41 | 0.47 | ** |

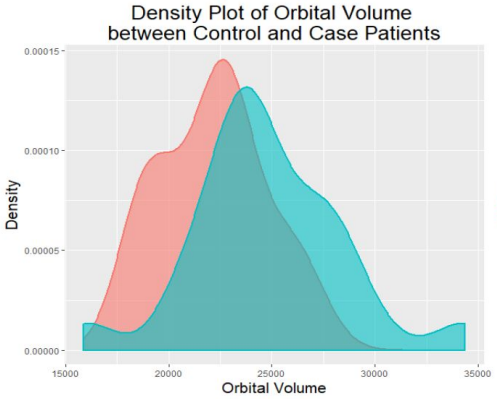
*Table 5: Independence test results between the left and right eyes for the experimental group.*
** Independence checked using non-parametric Spearman's coefficient of correlation for numerical variables.*

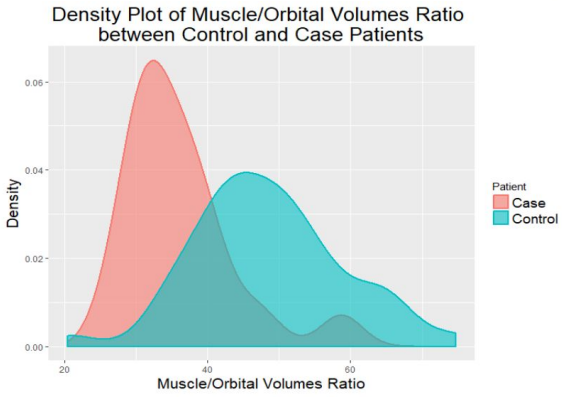
Per the results above we were able to verify the independence between the left and right eyes for the Lagophthalmos, Hertel, Fat Volume/ Orbital Volume and Muscle Volume/Orbital Volume variables. It was not possible to determine the independence of the variables marked with two stars as we did not have enough observations.

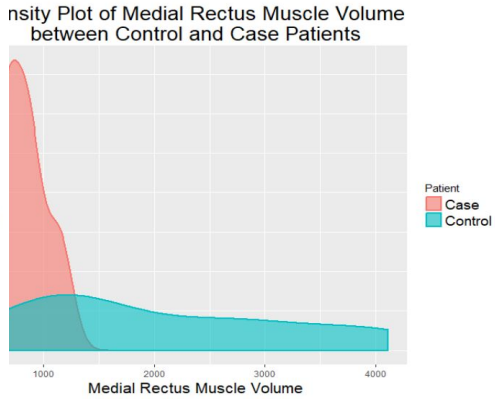
Control vs Experimental Groups

An important exploratory analysis was to compare the control group with the experimental group to see how these populations differed. Because the control group only had measurements for 5 variables, 5 comparisons were conducted. By running a simple t-test between these two groups to compare each of the 5 variables, the following results were obtained.

Statistically significantly variables different between control and case patients:

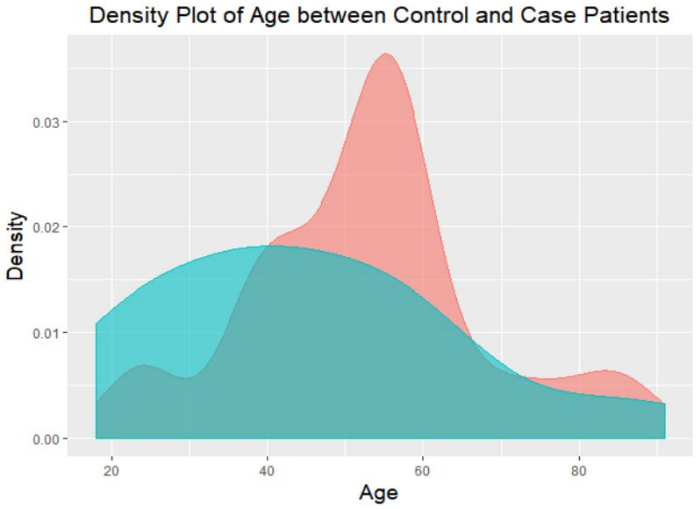


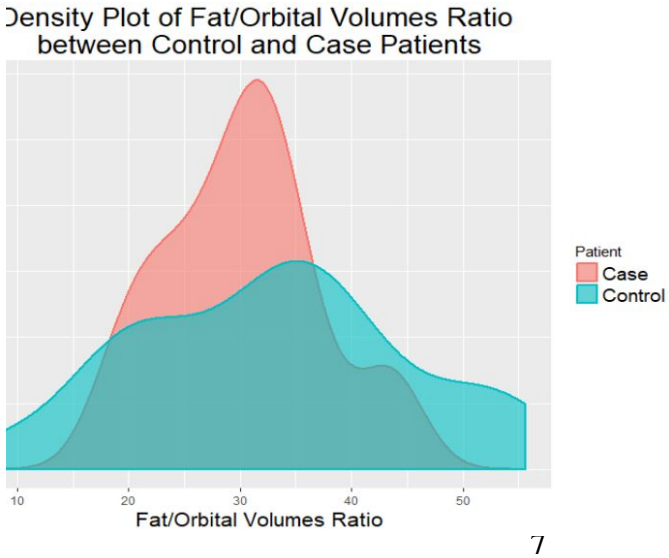




These three variables, orbital volume, muscle/orbital volumes ratio, and medial rectus muscle volume, were all statistically significantly different between the control (blue) patients and the experimental/case (red) patients. As seen in the figures above, it's very interesting to note that the spread/variance for the control patients is much larger than the experimental/case patients, who seem to have a more defined range for these variables.

Not statistically significantly variables different between control and case patients:

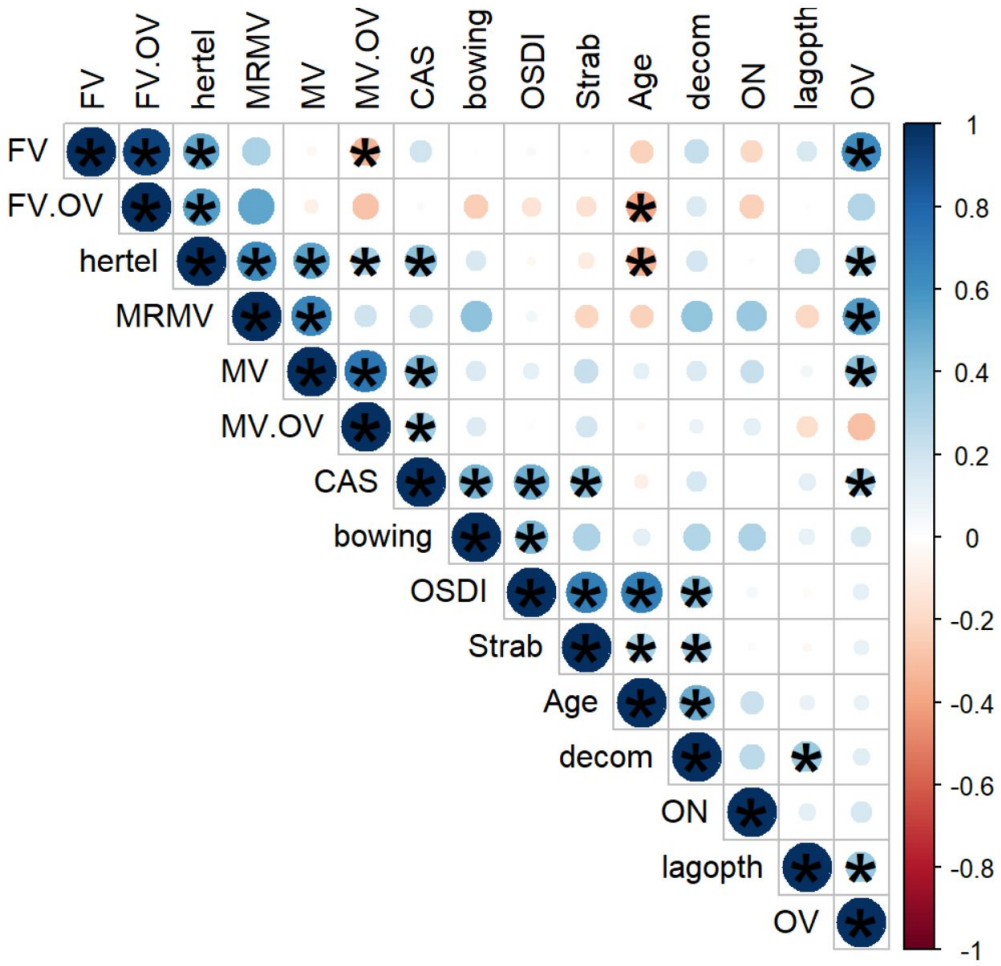


These two variables, age and fat/orbital volumes ratio, were not statistically significantly different between the control patients (blue) and the experimental/case (red) patients. It's fascinating to note that age is not correlated with the experimental group. Moreover, both of these distributions look somewhat normally distributed, possibly suggesting that these variables are not related to those with TED.

Covariance Matrix

Before constructing our models, we used a correlation matrix to explore the relations between our variables. The correlations of numeric-numeric and numeric-categorical variable pairs are represented by the pearson coefficient, while the correlations of categorical-categorical pairs are represented by phi. In the below graphic, the size and darkness of a dot represents the magnitude of correlation. Red dots represent negative correlations, and blue positive. Significance at 95% confidence level is denoted by an asterisk.



The coefficient and p-value matrices are included below for reference:

*Correlation coefficients:*

| | Age | CAS | OSDI | Strab | ON | lago | hertel | bowing | decom | FV | MV | OV | FV.OV | MV.OV | MRMV |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Age | 1.000 | -0.086 | 0.681 | 0.317 | 0.219 | 0.098 | -0.340 | 0.1189 | 0.494 | -0.2239 | 0.103 | 0.090 | -0.374 | -0.024 | -0.222 |
| CAS | -0.086 | 1.000 | 0.489 | 0.428 | 0.000 | 0.114 | 0.408 | 0.4660 | 0.170 | 0.1910 | 0.454 | 0.309 | 0.030 | 0.346 | 0.209 |
| OSDI | 0.681 | 0.489 | 1.000 | 0.681 | 0.049 | -0.024 | -0.031 | 0.4533 | 0.423 | 0.0348 | 0.106 | 0.110 | -0.141 | 0.016 | 0.052 |
| Strab | 0.317 | 0.428 | 0.681 | 1.000 | 0.027 | -0.035 | -0.108 | 0.3026 | 0.348 | 0.0125 | 0.229 | 0.093 | -0.156 | 0.184 | -0.220 |
| ON | 0.219 | 0.000 | 0.049 | 0.027 | 1.000 | 0.116 | 0.013 | 0.3055 | 0.266 | -0.2010 | 0.226 | 0.175 | -0.237 | 0.117 | 0.370 |
| lago | 0.098 | 0.114 | -0.024 | -0.035 | 0.116 | 1.000 | 0.252 | 0.0995 | 0.366 | 0.1643 | 0.055 | 0.379 | -0.012 | -0.177 | -0.208 |
| hertel | -0.340 | 0.408 | -0.031 | -0.108 | 0.013 | 0.252 | 1.000 | 0.1654 | 0.186 | 0.5218 | 0.534 | 0.369 | 0.554 | 0.321 | 0.631 |
| bowing | 0.119 | 0.466 | 0.453 | 0.303 | 0.306 | 0.100 | 0.165 | 1.0000 | 0.290 | -0.0083 | 0.160 | 0.162 | -0.245 | 0.149 | 0.404 |
| decom | 0.494 | 0.170 | 0.423 | 0.348 | 0.266 | 0.366 | 0.186 | 0.2901 | 1.000 | 0.2370 | 0.159 | 0.123 | 0.158 | 0.081 | 0.398 |
| FV | -0.224 | 0.191 | 0.035 | 0.012 | -0.201 | 0.164 | 0.522 | -0.0083 | 0.237 | 1.0000 | -0.037 | 0.621 | 0.923 | -0.332 | 0.314 |
| MV | 0.103 | 0.454 | 0.106 | 0.229 | 0.226 | 0.055 | 0.534 | 0.1598 | 0.159 | -0.0374 | 1.000 | 0.415 | -0.076 | 0.728 | 0.655 |
| OV | 0.090 | 0.309 | 0.110 | 0.093 | 0.175 | 0.379 | 0.369 | 0.1618 | 0.123 | 0.6214 | 0.415 | 1.000 | 0.296 | -0.299 | 0.557 |
| FV.OV | -0.374 | 0.030 | -0.141 | -0.156 | -0.237 | -0.012 | 0.554 | -0.2446 | 0.158 | 0.9230 | -0.076 | 0.296 | 1.000 | -0.285 | 0.526 |
| MV.OV | -0.024 | 0.346 | 0.016 | 0.184 | 0.117 | -0.177 | 0.321 | 0.1493 | 0.081 | -0.3320 | 0.728 | -0.299 | -0.285 | 1.000 | 0.200 |
| MRMV | -0.222 | 0.209 | 0.052 | -0.220 | 0.370 | -0.208 | 0.631 | 0.4043 | 0.398 | 0.3137 | 0.655 | 0.557 | 0.526 | 0.200 | 1.000 |

*Associated p-values:*

| | Age | CAS | OSDI | Strab | ON | lago | hertel | bowing | decom | FV | MV | OV | FVOV | MVOV | MRMV |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Age | 0.0e+00 | 0.5612 | 1.3e-06 | 2.8e-02 | 1.3e-01 | 0.549 | 1.8e-02 | 4.6e-01 | 1.2e-03 | 1.3e-01 | 4.9e-01 | 5.6e-01 | 1.5e-02 | 8.8e-01 | 4.1e-01 |
| CAS | 5.6e-01 | 0.0000 | 1.4e-03 | 2.4e-03 | 1.0e+00 | 0.485 | 4.0e-03 | 2.4e-03 | 2.9e-01 | 1.9e-01 | 1.2e-03 | 4.2e-02 | 8.5e-01 | 2.5e-02 | 4.4e-01 |
| OSDI | 1.3e-06 | 0.0014 | 0.0e+00 | 1.3e-06 | 7.6e-01 | 0.882 | 8.5e-01 | 3.3e-03 | 6.5e-03 | 8.3e-01 | 5.1e-01 | 5.2e-01 | 4.1e-01 | 9.3e-01 | 8.7e-01 |
| Strab | 2.8e-02 | 0.0024 | 1.3e-06 | 4.3e-12 | 8.5e-01 | 0.831 | 4.7e-01 | 5.6e-02 | 2.8e-02 | 9.3e-01 | 1.2e-01 | 5.5e-01 | 3.2e-01 | 2.4e-01 | 4.1e-01 |
| ON | 1.3e-01 | 1.0000 | 7.6e-01 | 8.5e-01 | 4.3e-12 | 0.475 | 9.3e-01 | 5.3e-02 | 9.3e-02 | 1.7e-01 | 1.2e-01 | 2.5e-01 | 1.3e-01 | 4.6e-01 | 1.6e-01 |
| lago | 5.5e-01 | 0.4852 | 8.8e-01 | 8.3e-01 | 4.8e-01 | 0.000 | 1.2e-01 | 5.4e-01 | 2.0e-02 | 3.1e-01 | 7.3e-01 | 2.2e-02 | 9.4e-01 | 3.0e-01 | 5.2e-01 |
| hert | 1.8e-02 | 0.0040 | 8.5e-01 | 4.7e-01 | 9.3e-01 | 0.116 | 0.0e+00 | 3.1e-01 | 2.5e-01 | 1.4e-04 | 9.2e-05 | 1.4e-02 | 1.4e-04 | 3.8e-02 | 8.8e-03 |
| bow | 4.6e-01 | 0.0024 | 3.3e-03 | 5.6e-02 | 5.3e-02 | 0.541 | 3.1e-01 | 2.5e-10 | 6.7e-02 | 9.6e-01 | 3.2e-01 | 3.5e-01 | 1.5e-01 | 3.8e-01 | 1.9e-01 |
| decom | 1.2e-03 | 0.2940 | 6.5e-03 | 2.8e-02 | 9.3e-02 | 0.020 | 2.5e-01 | 6.7e-02 | 2.5e-10 | 1.4e-01 | 3.3e-01 | 4.7e-01 | 3.6e-01 | 6.4e-01 | 2.0e-01 |
| FV | 1.3e-01 | 0.1935 | 8.3e-01 | 9.3e-01 | 1.7e-01 | 0.311 | 1.4e-04 | 9.6e-01 | 1.4e-01 | 0.0e+00 | 8.0e-01 | 6.8e-06 | 3.5e-18 | 3.2e-02 | 2.4e-01 |
| MV | 4.9e-01 | 0.0012 | 5.1e-01 | 1.2e-01 | 1.2e-01 | 0.735 | 9.2e-05 | 3.2e-01 | 3.3e-01 | 8.0e-01 | 0.0e+00 | 5.1e-03 | 6.3e-01 | 4.7e-08 | 5.9e-03 |
| OV | 5.6e-01 | 0.0415 | 5.2e-01 | 5.5e-01 | 2.5e-01 | 0.022 | 1.4e-02 | 3.5e-01 | 4.7e-01 | 6.8e-06 | 5.1e-03 | 0.0e+00 | 5.7e-02 | 5.5e-02 | 3.8e-02 |
| FVOV | 1.5e-02 | 0.8528 | 4.1e-01 | 3.2e-01 | 1.3e-01 | 0.943 | 1.4e-04 | 1.5e-01 | 3.6e-01 | 3.5e-18 | 6.3e-01 | 5.7e-02 | 1.1e-308 | 6.7e-02 | 5.3e-02 |
| MVOV | 8.8e-01 | 0.0249 | 9.3e-01 | 2.4e-01 | 4.6e-01 | 0.301 | 3.8e-02 | 3.8e-01 | 6.4e-01 | 3.2e-02 | 4.7e-08 | 5.5e-02 | 6.7e-02 | 0.0e+00 | 4.9e-01 |
| MRMV | 4.1e-01 | 0.4370 | 8.7e-01 | 4.1e-01 | 1.6e-01 | 0.517 | 8.8e-03 | 1.9e-01 | 2.0e-01 | 2.4e-01 | 5.9e-03 | 3.8e-02 | 5.3e-02 | 4.9e-01 | 7.1e-109 |

## Statistical analysis used to answer the research questions

To understand if any of the 12 variables measured in the study were significant predictors of Optic Neuropathy or Strabismus we used logistic regression and penalized logistic regression as models. An overview of the variables used to build the two models are present in the tables below:

| Mean of | CAS.score | OSDI.score | orbit_volume | muscle_vol | fat_vol | medial_rectus_muscle_vol |
|---|---|---|---|---|---|---|
| ON: Yes | 2.66 | 47.5 | 24464.25 | 11223.36 | 8718.93 | 1759.774 |
| ON: No | 2.66 | 51.6 | 26108.66 | 12933.60 | 6835.62 | 2616.837 |

*Table 6: Logistic Model and Penalized Logistic Model for Predicting Optic Neuropathy*

| | |
|---|---|
| CAS Score | type= numeric |
| OSDI Score | type= numeric |
| Lagophthalmos (left and right ) | type= numeric |
| Decompression (left and right) | type= categorical |
| Hertel (left and right) | type= numeric |
| Fat Vol/ Orbit Vol (left and right) | type= numeric |
| Muscle Vol/ Orbit Vol (left and right) | type= numeric |
| Medial Wall Bowing (left and right) | type= categorical |
| Medial Rectus Muscle Vol (left and right) | type= numeric |

*Table 7: Logistic Model and Penalized Logistic Model for Predicting Strabismus:*

As for comparing the relationships between the different variable for our different response variables including the presence of ON, strabismus or the decision to get decompression we made use of non-parametric Wilcoxon tests that compare the distributions of the numerical and chi-squared tests for the categorical variables. The use of non-parametric tests allowed us to compensate for the non-normal distribution of the variables and the small sample size. In most cases owing possibly to the small sample size we were unable to detect statistically significant relationships, especially for smaller effect sizes.

## Summary of results

Comparing numerical variables individually across groups of interest

| Variable | ON = 1 | ON = 0 | Decompression | No Decompression | Strabismus | No Strabismus |
|---|---|---|---|---|---|---|
| Lagophthalmos | 0.79 | 0.54 | 0.85*** | 0.3409*** | 0.55 | 0.607 |
| Hertel | 22.22 | 22.07 | 23.05* | 21.45* | 21.71 | 22.65 |
| FV/OV | 25.86* | 34.02* | 35.62 | 31.61 | 31.38 | 35.25 |
| MV/OV | 51.35* | 47.86* | 49.29 | 47.47 | 49.86 | 45.92 |
| MV/FV | 2.42*** | 1.601*** | 1.79 | 1.61 | 1.95* | 1.49* |

*Table 8: Comparing numerical variables across the categorical variables of interest, this table shows the mean values of the variables of interest for patients who satisfy the criteria in the columns. The comparison was done using non-parametric Wilcoxon tests. '*' indicates a difference at confidence level= 0.8 and *** indicates a significant difference at confidence level= 0.95.*

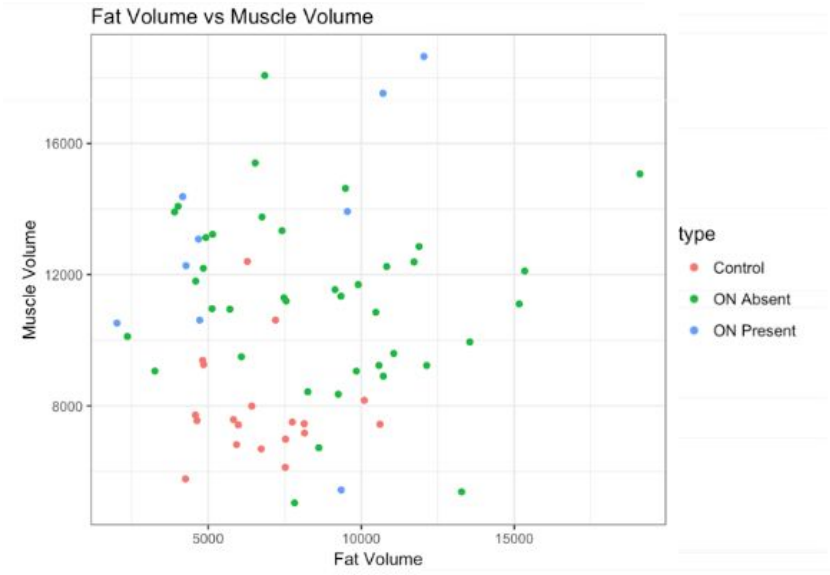
Per the table above, it can be seen that the ratio of Muscle Volume to Fat Volume is significantly different for the patients who have ON (mean MV/FV = 2.42) when compared to patients who do not have ON (mean MV/FV = 1.60) and for patients who have strabismus (mean = 1.95) versus patients who don't (mean = 1.49). As for decompression, lagopthalmos is seen as having a significant difference for patients who opt for decompression vs. patients who do not.

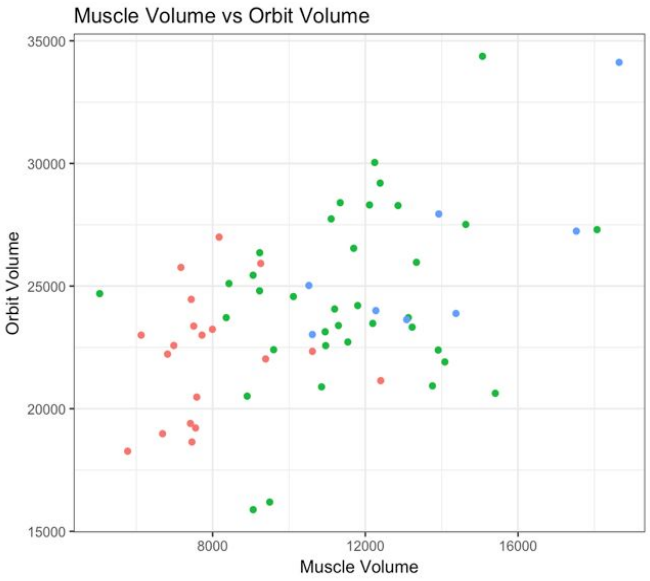
## Interpretation of results

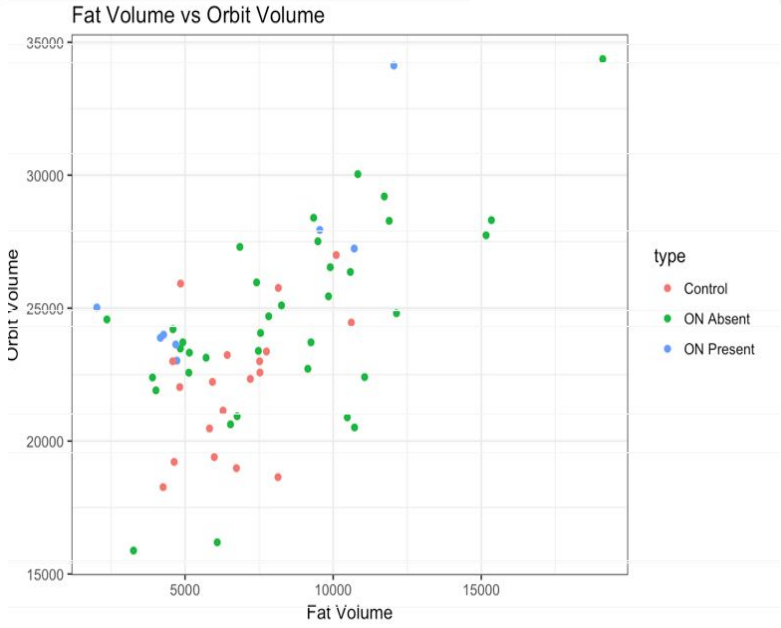
Our analysis provided evidence for the claims that there exist differences between the control and experimental patients, as well as between experimental patients with and without ON. However, there were significant predictors in our regression models predicting strabismus and optic neuropathy. Our only significant result with regard to these conditions was the higher fat-to-muscle volume ratio among patients with ON and patients with strabismus. The insignificance of other predictors may come from such a small sample size, so working with more data may help to clarify these results. However, after consultation with the doctors, it is very possible that there are external confounding variables present in TED that were not explored through this study. Thus, further analyses on other features are necessary to better understand the disease.

## Interpretation of plots

Scatterplots of Variables

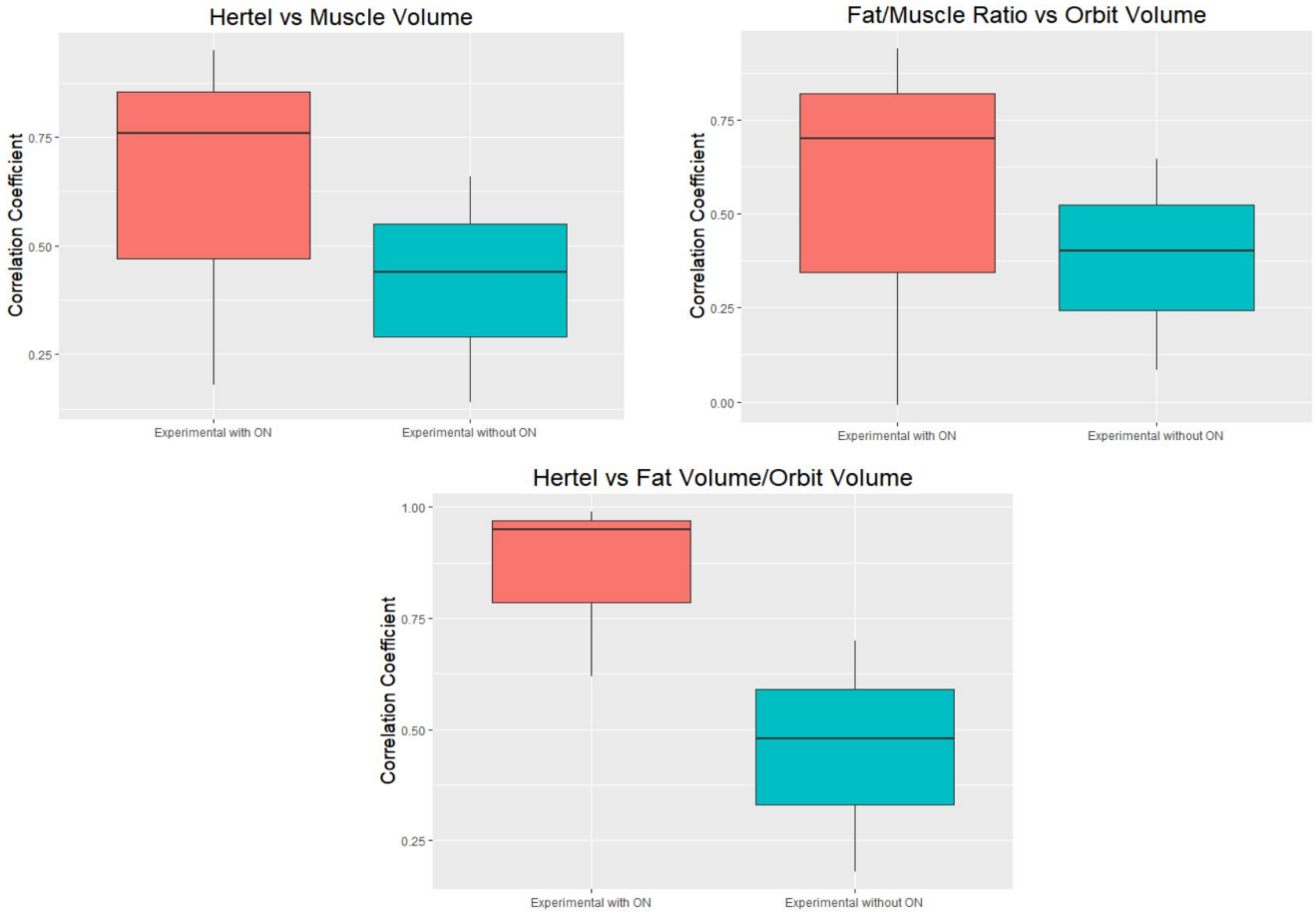



These scatterplots plot variables to variable by patient type. The color represents whether ON is present or not (if a case/experimental patient), or if the patient was a control patient. Due to the fact that the green and blue points are intermixed together and there is no clear clustering, this plot reveals that muscle volume and fat volume are not good indicators of whether optic neuropathy will be present in a patient or not.

Correlation Coefficient of Significant Variables in Experimental ON vs Experimental non-ON




The boxplots above show the correlation coefficient based from the scatterplots from the prior section. These boxplots are comparing the correlation coefficient of experimental group with ON (red) to experimental group without ON (blue). These variable to variable comparisons were found to be statistically significant between the two populations. In essence, what these boxplots are measuring is the correlation coefficient between two variables. The higher the correlation coefficient, the more tight the fit is/the more clustered the model is/less random the relationship/a better relationship of the variables.

The middle line shows the estimated correlation coefficient, while the end of the tails (not the end of the box) show the 95% confidence interval of the correlation coefficient. In all of these three cases, although there is some overlap in the intervals between the two populations, the correlation coefficient has a higher range for the experimental group with ON. This finding makes sense because it says that there is a difference between experimental with ON and experimental without ON groups, or at least with these variables, showing that those with ON have an actual relationship with the variables, while those without ON have a more random relationship.

## Overall conclusions

In conclusion, our findings demonstrated that external confounding factors were impacting TED, since all the models developed yielded no significance. For the future, medical professionals should consider measuring physiological factors in the eye along with anatomical ones in addition to factors such as gender, time with disease, and thyroid medication dosages.

In addition, to overcome some of the challenges we faced, physicians should consider implementing a larger study while there are logistical and financial challenges that come along with gathering clinical data at a large volume.

With this additional information, we can understand TED even better and hope to prevent blindness and double vision for future generations of thyroid patients.

## Challenges of the study

1. **Small Sample Size**

The experimental group consisted of 24 people who were infected by thyroid of which only 8 showed optic neuropathy, which meant only 9 eyes with optic neuropathy out of a total of 48 eyes. As a result the independence tests and models may have a low power, allowing for large effects to go undetected. While the groups for patients having and not having strabismus were split in a fairly balanced manner (14 patients with strabismus and 10 patients with no strabismus) the sample size was still suboptimal for drawing causal correlations.

2. **Missing values**

In addition to the small sample size, missing values in the experimental patients such as decompression and medial wall bowing (8 missing values for both of these variables) further limited the power of our tests and models.

## Recommendations for the future

The biggest challenge we faced in analyzing our data was the sample size. Significant tests showed no significance because of how small the sample size is. Although collecting such data is tricky, for more innovative results, waiting to analyze once there is more data is highly suggested. Otherwise, the non-significance of many of the statistical tests are confounded by small sample size, and it becomes hard to determine whether variables were not significant due to small sample size or due to the variable itself.

Furthermore, although the control patients do not have TED, it may be beneficial to do more analyses between the control and diseased patients. The control patients had many missing values, so collecting measurements for this population as well would add to the study.

This study is a first in analyzing thyroid data, but based on these initial findings, collecting other measurements of patients is advised. Consultation with doctors suggested that features such as pressure may play a role, so using advice from doctors as well as literature may help in choosing more features to collect that may play a role in TED.