Investigating Consumption & Exercise: A Study on the Effect of Carbohydrates on a Runner's Endurance

Daniel Jenkins, Genevieve Gonzales, Emmy Su, Saul Magallon-Soto

June 15th, 2024

Contents

1 Abstract

The overwhelming presence of carbohydrates in the American diet has become a focal point for observing the acquisition of metabolic and cardiovascular diseases. It has driven us to want to observe the impact carbs have on our daily physical performance. We hypothesize that the amount of carbohydrates one consumes daily will have a noticeable effect on their physical capabilities, specifically their endurance. In our study, we will use a Randomized Basic Factorial Design of Two-Factors and a One-Way ANOVA analysis to observe these effects on islanders in a simulation. Using proper random sampling techniques, we hope to witness the short-term effects carbohydrates can have on endurance through an appropriately represented population. We can then use this information to theorize on the possible detrimental and long-term effects of overconsuming carbs.

2 Introduction

Heart disease is the primary cause of death for United States citizens regardless of sex or racial identity. The most common type of heart disease is coronary artery disease, which roughly 5% of adults suffer from [1]. This is developed from a prolonged accumulation of plaque in the arteries, plaque that comes from the presence of too much low-density lipoprotein (LDL). LDL can be quickly gained in small amounts from eating carbohydrates, specifically simple sugars [2]. This plaque build up can become atherosclerosis, a process in which the arteries begin to clog and are unable to provide oxygen-rich blood to the heart and other organs of the body in a timely manner [3]. Then, this will typically lead to a heart attack, where 1 in every 10 people will die within a short period of time [4].

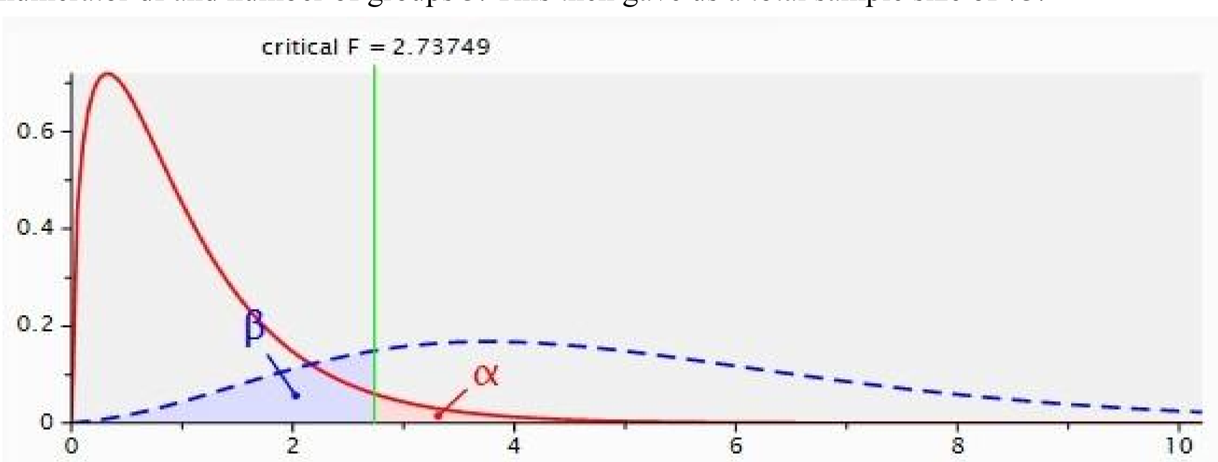
Simple sugars are a fast-acting type of carbohydrate, being broken down quickly by the body upon ingesting them. These simple sugars can come naturally from fruits, but are more often found in added sugars such as fructose and sucrose which is prevalent in many processed foods of the American diet [5]. These sugars are very important for providing our body and brains with the energy needed to function. For Americans, the Daily Recommended Intake of carbohydrates is less than 130 grams. However, the average American consumes over 250 grams of carbs a day [6]. So it is evident that there is a problem in how much we consume and how this issue can affect us over an extended length of time. To observe the immediate effects of this issue, we would need to test how the body uses these sugars entering into the bloodstream. One of the best ways to test this is through exercise, as physical activity requires a constant flow of oxygen to the cells of the body. This means through an activity like running, the ingested carbohydrates will be constantly running throughout the bloodstream to the muscles and other cells.

This led us to our study's design of testing the effects of foods with different carbohydrate contents on physical performance. We believe that the amount of carbs ingested before exercising will impact how well one can perform that physical activity. In order to manipulate the carb content that our participants consume, we will have them eat either Cream Cheese, White Bread, or Milk Chocolate, which contain a low, moderate, and high content of carbohydrates respectively. Then, we wanted to choose an activity that would effectively get much of the body's blood flow moving quickly so that the carbohydrates could be well distributed to the cells. This is why we choose outdoor running, an exercise that is highly efficient at utilizing carbohydrates as a source of energy [7]. We can have our participants then run different lengths, of 1 km, 5 km, and 10 km, to see how short, medium, and long periods of exercise are affected differently. We propose that the amount of carbohydrates consumed will have an impact on the time it takes to complete a distance run.

3 Methodology

3.1 Sampling

In order to test our hypothesis, we needed to obtain a fair distribution of participants that would accurately represent an entire population. The first step was to determine the size of the sample needed so that our tests may be significant. This was done using the software GPower, where we used the statistical test of ANCOVA: Fixed effects, main effects and interactions in the F tests test family. We then used a large effect size ($f = 0.4$) as we were limited on time and resources. For our significance level, we used $\alpha = 0.05$, and for our Power, we used $1 - \beta = 0.80$, the minimum level required for significance in most statistical tests akin to this one. Finally, since we have 3 levels of our independent and dependent variable, we made our numerator df and number of groups 3. This then gave us a total sample size of 73.



For the purpose of our study, we wanted to have an equal number of participants for each of our test groups. Since we would have 9 groups, the minimum number of people needed would be 81, so that each group could have 9 members. So we rounded our sample size up from 73 to 81. Then, we needed to do our best to obtain these samples randomly from the island population. To accomplish this, we compiled a list of all of the city names of the islands and the corresponding number of houses each city has. Starting by setting the seed, we used the R sample function to sample 900 cities, their individual house number, and their overall house number from across the entire simulation, without replacement. We decided to choose 900 samples based on the notion that with 810 samples, we would obtain a participant approximately 1 in every 10 houses, and the extra 90 acted as a cushion for extra participants in case we needed to sample more.

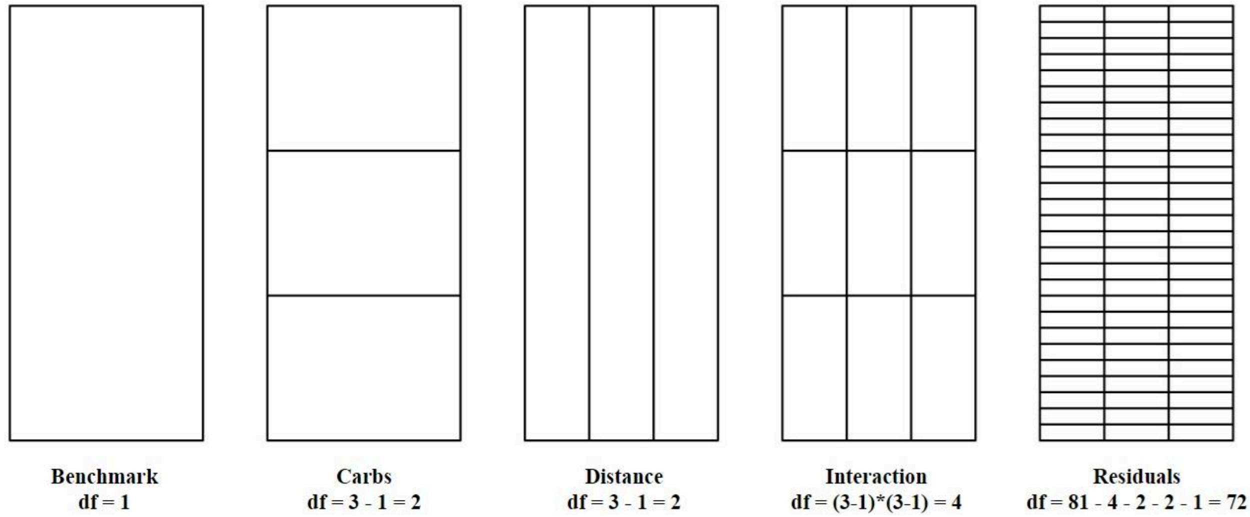
3.2 Design

Our design is a Randomized Basic Factorial Design of Two-Factors, with the two factors being carb content and running distance. Their levels are as follows:

Carb Content (Food Type)	Running Distance
2 grams (Cream Cheese)	1 km
25 grams (White Bread)	5 km
100 grams (Milk Chocolate)	10 km

Our response variable will be the time it takes the participants to complete their run in minutes. The nuisance factors we attempted to control in this study were sex and age. We decided to only sample females between the ages of 18 and 32 years old since this is the age range for peak physical performance [8].

The factor diagram for this design is:



3.3 Procedure

Now that we had a list of randomly sampled houses from all the cities on all of the islands, we needed to go to each house and ask the first eligible participant from each household for their consent in our study. If the house we went to had no males or females between 18 and 32 years old, the next sample was used. If the eligible participant declined to participate in our study, the next sample was used. This process was repeated until we obtained 81 different participants, which had us use roughly 600 of our 900 samples. These were then randomly assigned into the 9 groups so that each one had 9 members using R's sample function once again. The 9 different groups are summarized in the following table:

2 g Carbs + 1 km (Cream Cheese 50 g)	**25 g Carbs + 1 km (White Bread 50 g)**	**100 g Carbs + 1 km (Milk Chocolate 50 g)**
2 g Carbs + 5 km	**25 g Carbs + 5 km**	**100 g Carbs + 5 km**
2 g Carbs + 10 km	**25 g Carbs + 10 km**	**100 g Carbs + 10 km**

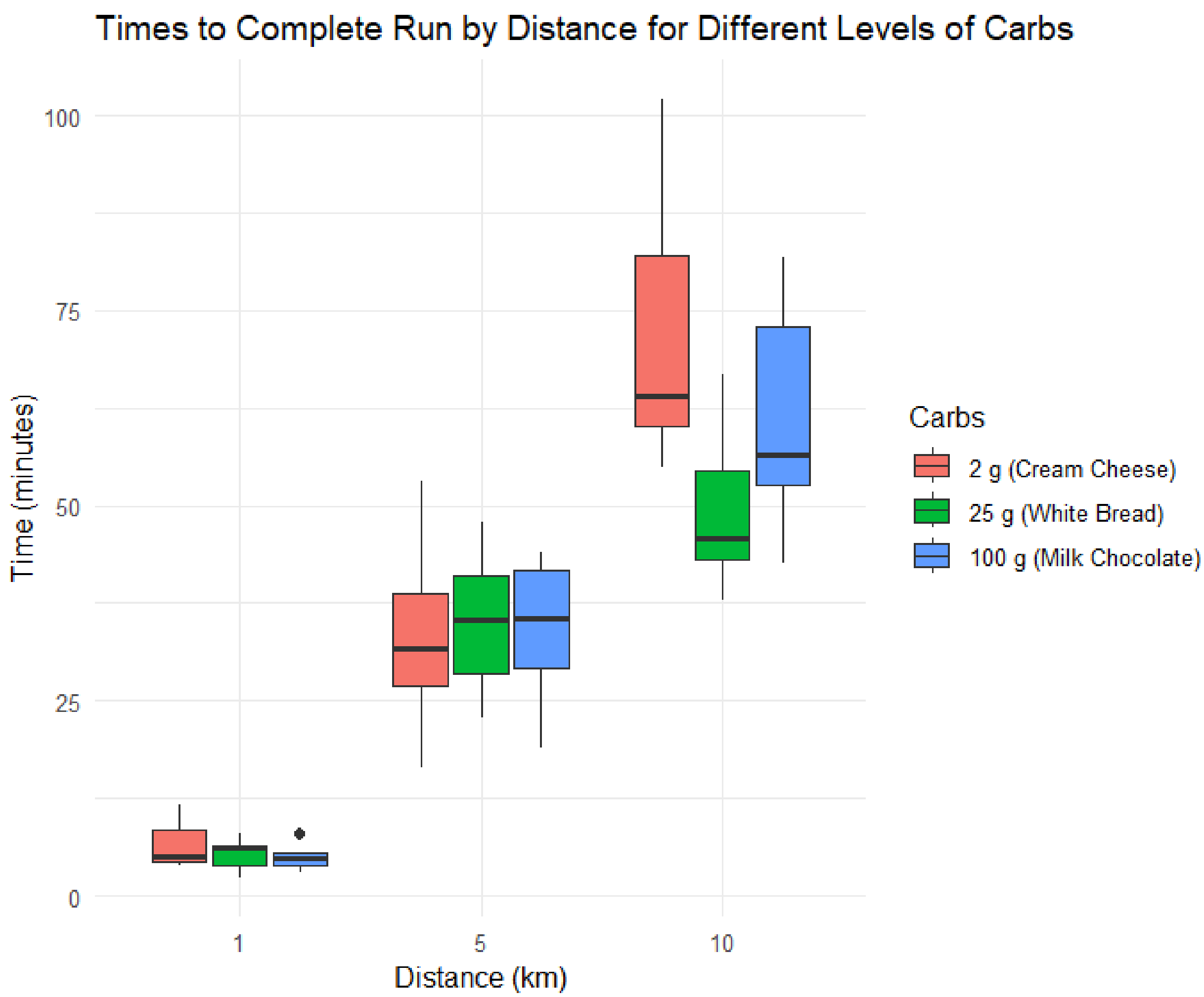
Once the groups were formed, we ran the following procedure for each group:

1) Feed the participant their group's food item.
2) Wait 1 minute after they finish eating.
3) Make them run their distance run.
    a) For 10 km, we had the participants run two 5 km runs back to back.
4) Record the time it took them to complete their run.

4 Analysis

4.1 Box Plot

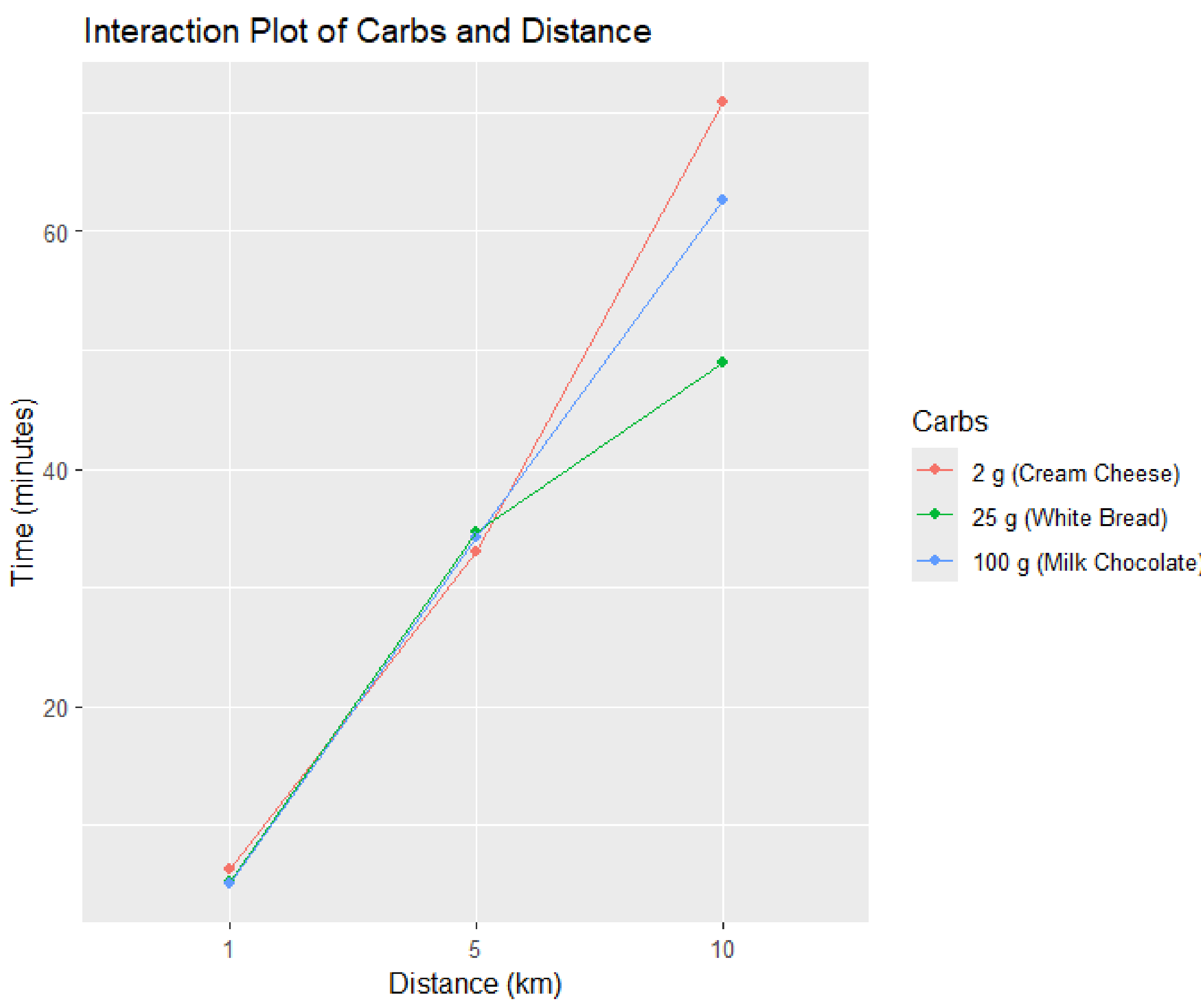
We ran a One-Way ANOVA on our collected data using R's *lm* and *anova* functions, using "Time" as our response variable, and the product of "Carbs" and "Distance" as our predictor variables. The relationship of these variables can be seen in the following boxplot:



From this plot, it is already evident that there is likely an effect between the relationship of the consumed carb content and the time it takes to complete the run. The medians of the box plots for the 1 km and 5 km runs, represented by the horizontal black bar, are very close together, but for the 10 km run, there seems to be a great amount of diversity. As the distance of the run increases, the height of the boxes increases, meaning that the deviation of the data is increasing, being most noticeable at the 10 km run. The whiskers of the boxes also get longer as the distance increases, meaning that more of the data falls outside of the Interquartile Range, extending further past the median value. Additionally, on the 1 km run, there is a black dot present that indicates the presence of an outlier.

4.2 Interaction Plot

To further investigate the relationship of our predictor variables, especially how distance and carb content interact with each other, the following interaction plot was constructed:



This plot once again shows that at the short and medium distance runs, the carb content consumed does not seem to have much of an effect on the time it takes to complete the run. However, at the long distance of 10 km, there is clearly a great difference between the mean times of each carb group. The low carb content of 2 g took the longest, the high carb content of

100 g took less time, and the moderate carb content of 25 g took the least amount of time. This would indicate that there is a type of interaction between carb content and distance ran.

4.3 ANOVA Results

To concretely see the effects of the relationship of the consumption of carbohydrates and running distance on endurance, we need to analyze the results of our ANOVA test. The results of this test are summarized in the following ANOVA table:

Factor	Df	Sum Sq	Mean Sq	F-value	P-value
Carbs	2	685	342.3	3.5815	0.0329
Distance	2	41311	20655.4	216.1356	< 2.2e-16
Interaction	4	1535	383.7	4.0154	0.0054
Residuals	72	6881	95.6		

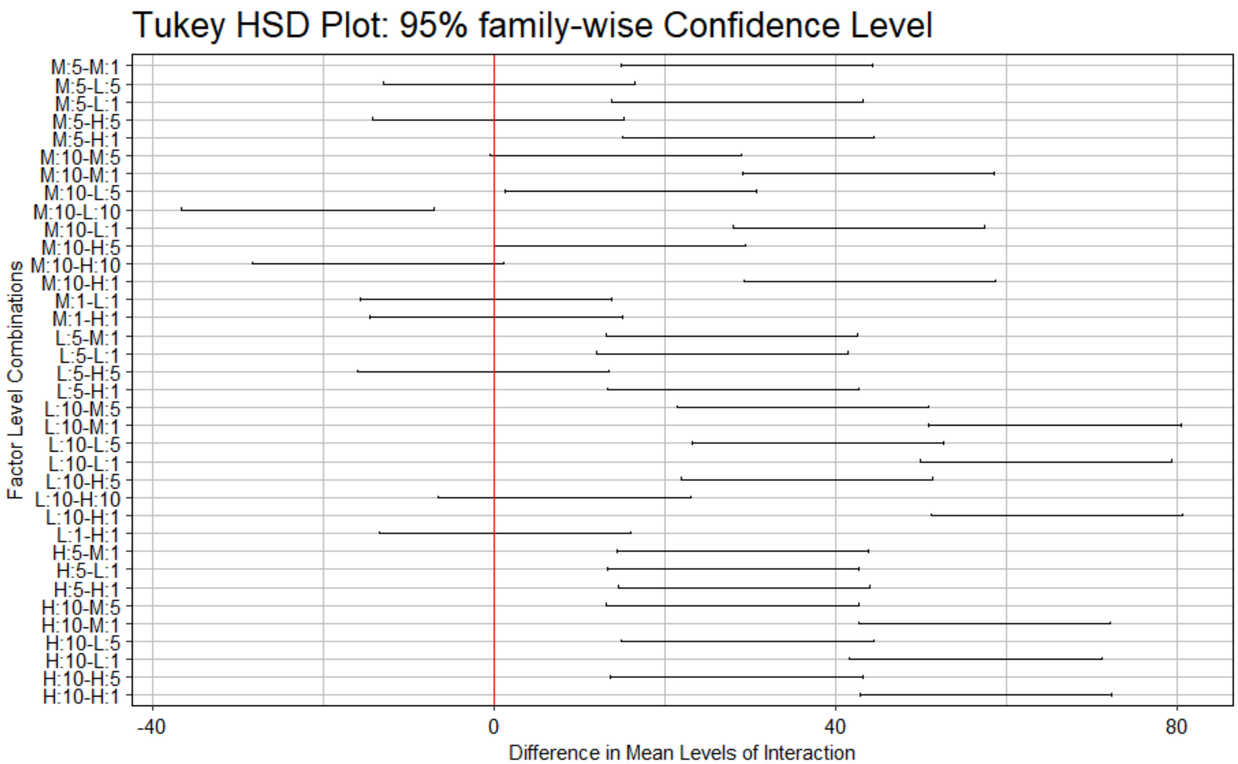
Looking at the "P-value" column of the table, we can compare the values to our significance level $\alpha = 0.05$ and determine if they are significant. Starting with Carbs, with a value of 0.0329, which is less than α, we can conclude that Carbs is a significant variable. This means that the amount of carbohydrates consumed is a significant predictor for the time it takes to complete a run. Then the variable of Distance, obviously being significant with a value close to 0. Then we look at our interaction factor, with a value of 0.0054, again less than α, so the interaction of carbs and distance is also significant in determining one's endurance.

The Multiple R-squared value of our analysis is 0.8635. This means that 86.35% of the variance in our data can be explained by our model.

5 Post Hoc Analysis

5.1 Tukey HSD

Since the results of our ANOVA test were valid, it is appropriate to run a post hoc analysis to determine the differences between our treatment groups. One way we can do that is to run a Tukey Honestly Significant Difference (HSD) test to test the significance of the differences between means of each group. The test is summarized in the following graph. For simplicity, the different carb contents will be abbreviated to L (2 g), M (25 g), and H (100 g):



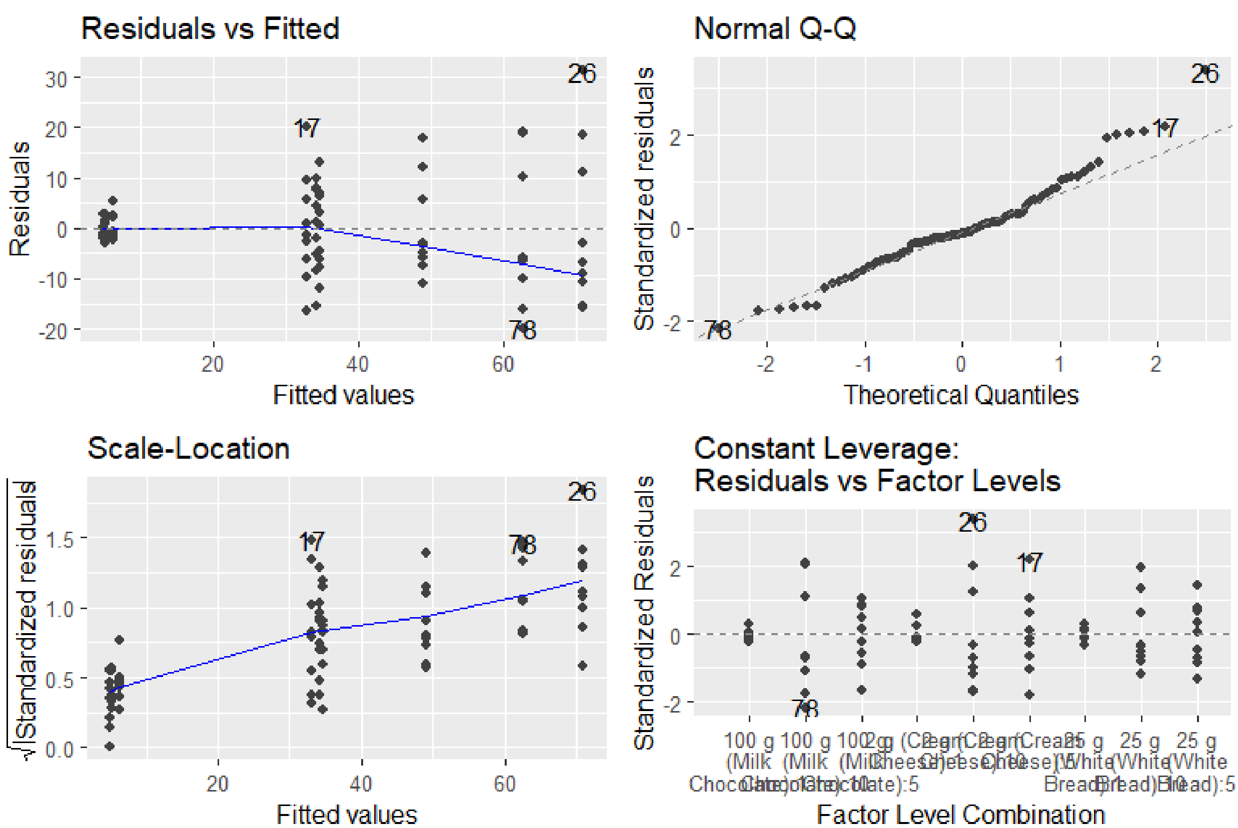
Most intervals in the Tukey HSD analysis do not contain the value 0, meaning they are significantly distinguishable between the other interaction groups. The only ones that do contain 0 are summarized in the following table:

Factor	**Difference**	**Lower**	**Upper**	**P-adjusted**
$L{:}1 - H{:}1$	-1.244	-13.493	15.982	1.000
$M{:}1 - H{:}1$	0.200	-14.538	14.938	1.000
$M{:}1 - L{:}1$	-1.044	-15.782	13.693	1.000
$L{:}5 - H{:}5$	-1.267	-16.004	13.471	1.000
$M{:}5 - H{:}5$	0.467	-14.271	15.204	1.000
$M{:}10 - H{:}5$	14.706	-0.032	29.443	0.051
$M{:}5 - L{:}5$	1.733	-13.004	16.471	1.000
$M{:}10 - M{:}5$	14.239	-0.499	28.977	0.066
$L{:}10 - H{:}10$	8.222	-6.516	22.960	0.692
$M{:}10 - H{:}10$	-13.650	-28.388	1.088	0.091

Most of these values make sense when compared to the previous boxplot, showing that at short and medium length runs, the interaction between carb content and distance is insignificant. The two that raise attention are the last ones of the interaction between L:10-H:10 and M:10-H:10. These would somewhat suggest that the difference between the low and high carb content at a long distance run, as well as the difference between the medium and the high carb content at a long distance run are insignificant. However, the difference between the low and medium carb content at a long distance run is still significant, showing that the interaction of carb content and distance is still an important factor.

5.2 Diagnostic Plots

The next way we can check the validity of our data is to check the diagnostic plots of our model, which can be seen in the following plots:



Residuals vs Fitted: This plot should show us an approximately horizontal line with errors randomly distributed around the line. It is clear near the right tail of the plot there exists leverage values that are manipulating the fitted line to skew.

Normal Q-Q: This plot identifies the normality of the errors, where you would want to see the data points mostly follow straight along the dotted line. Again, near the right tail of this graph, the points deviate slightly from the line, which makes us need to question the normality of our errors.

Scale-Location: This plot shows the homoskedasticity of our data. We would want this line to be horizontal, with the errors randomly distributed around the line. In this case, the line skews up as you move positively along the x-axis, which violates our assumption of equal variance.

Constant Leverage: This plot helps identify any possible leverage values in the data that may be influencing the results of the model. The values that jump above or fall below a standardized residual level of 2 or -2 respectively may be seen as high leverages that impact our data.

6 Conclusions

Our goal was to determine if the level of carbohydrates a person consumes would impact their endurance through the time it would take them to finish different lengths of running. The results we obtained attempt to visualize the short-term effects of the overconsumption of carbohydrates in the daily diet on physical performance.

By testing our design on our sampled islanders and running an ANOVA test on their results, we can see the impact of carbohydrate content and distance on running times. Both of our factors of carbs and distance were significant in this study, as well as the interaction between the two factors. Further analyzing our findings, we discovered that the content of carbs consumed affects longer distance running far more than short or medium distance running. Additionally, we came to find out that a moderate carbohydrate consumption led to faster times than a high and a low carbohydrate amount.

We acknowledge that the measures used for this experiment were not perfect, and other factors may have come in to influence our results. When it comes to the different food types that represent the different carb levels, there are ingredients in these foods that may impact the body in different ways. Thus, fats prevalent in Cream Cheese and Milk Chocolate may have played a role in the slower times the participants received. In addition, we recognize the presence of outliers in our data, which limit the validity of our results based on a breach of error normality.

When it comes to reevaluating and replicating our study, we would want to use a smaller effect size so our data could be more accurate and our outliers would have a lessened impact on our data. Some ideas for different studies we would want to try to examine more topics relevant to this one would include a change of our variables. For our predictor variables, trying to isolate certain molecules like carbohydrates could establish a purer relationship between consumption and performance. Also, it would be interesting to study other molecules like lipids and proteins to see their effect on endurance. We could possibly change our response variable from endurance to other physical activities such as strength or speed. We could add a blocking factor of sex to observe the differences that males and females experience in this same study. We may also want to block on age groups, as 18 to 32 years old may have a great variability, or perhaps older or younger participants would exhibit the effects differently. Our design could also be changed to a repeated measures design so that the person-to-person variability could be studied.

Knowing what we know now, this study serves to inform on the effects carbohydrates have on our physical performance. This can help us understand how this may translate to further developing effects over a longer period of time. Hopefully, this study can serve to better our knowledge of consuming in moderation and maintaining healthy habits and lifestyles.

7 References

“Heart Disease Facts.” *Centers for Disease Control and Prevention*, Centers for Disease Control and Prevention, 15 May 2024, www.cdc.gov/heart-disease/data-research/facts-stats/index.html.

Siri, Patty W, and Ronald M Krauss. “Influence of Dietary Carbohydrate and Fat on LDL and HDL Particle Distributions.” *Current Atherosclerosis Reports*, U.S. National Library of Medicine, 7 Nov. 2005, pubmed.ncbi.nlm.nih.gov/16256003/#:~:text=Increased%20dietary%20carbohydrates%2C%20particularly%20simple,involve%20increasing%20plasma%20triglyceride%20concentrations.

“What Is Atherosclerosis?” *National Heart Lung and Blood Institute*, U.S. Department of Health and Human Services, 24 Mar. 2022, www.nhlbi.nih.gov/health/atherosclerosis.

“Heart Attack (Myocardial Infarction).” *Cedars Sinai Health Library*, The StayWell Company, LLC, www.cedars-sinai.org/health-library/diseases-and-conditions/h/heart-attack-myocardial-infarction.html#:~:text=Chances%20of%20surviving%20a%20heart,well%20as%20have%20other%20disorders.

“Carbohydrates: Getting the Most out of Fiber, Starches & Sugars.” *Cleveland Clinic*, Cleveland Clinic, 8 Mar. 2024, my.clevelandclinic.org/health/articles/15416-carbohydrates.

McKinney, Christine. “Carbohydrate Goals.” *The Johns Hopkins Patient Guide to Diabetes*, This website is a publication of The Johns Hopkins University and copyright © 2016-2024, Division of Endocrinology, Diabetes, & Metabolism, The Johns Hopkins University., 22 Mar. 2024, hopkinsdiabetesinfo.org/carbohydrate-goals/.

Cantwell, J D. “Cardiovascular Aspects of Running.” *Clinics in Sports Medicine*, U.S. National Library of Medicine, pubmed.ncbi.nlm.nih.gov/3902253/.

Nikolaidis, Pantelis T, et al. “The Age of Peak Performance in Women and Men Duathletes - the Paradigm of Short and Long Versions in ‘Powerman Zofingen.’” *Open Access Journal of Sports Medicine*, U.S. National Library of Medicine, 18 July 2018, www.ncbi.nlm.nih.gov/pmc/articles/PMC6054758/.