

Mark Qi

Health Analytics Project

Xqi7 - 902555498

Introduction

According to Silva 2019, lipomas are lumps under someone's skin due to an overgrowth of fat cells. They are classified as "benign tumors" and typically do not pose any issues to a person's health. The causes of the lipomas are generally unknown, but it affects a significant portion of the population. Reported in Charifa 2018, Lipoma affects 1% of the population, and come in various forms, primarily in the upper body, neck, and face regions.

Because of the unknown cause of lipoma, it's difficult to generalize what exact mechanism lipomas are formed but there have been some related risk factors, such as genetics and age. According to the Mayo Clinic website, lipoma most commonly occur in individuals between the ages of 40 and 60. Additionally, some minor risk factors include obesity, high cholesterol, diabetes, liver disease, and glucose intolerance.

Hypothesis: In my study, I decided to study the effects of obesity and general fitness of patients and its correlation with the occurrences of lipoma by way of regular exercise.

Methods

- How did you conduct your analysis?

In deciding on my topic of studying things that may affect the development of lipoma, I went fruitlessly through several topics. Those other topics include "Does having lipoma-removing surgery previously increases your chances of getting a lipoma again?" and "The effect of exercise on the risk of lipoma." Both of these ideas would be promising future areas of research but in the sample patients sets provided, there were not enough evidence to support either study.

- Concept Sets
 - [xqi7] Lipoma

- The concept set includes everything searching containing “lipoma” of the class “Clinical Finding” with a non-zero RC or DRC.
- Link:
 - <http://gt-health-analytics.us-east-1.elasticbeanstalk.com/#/conceptset/651/conceptset-expression>
- **Obesity + Overweightness**
 - Obesity and Overweight are both medically recognized terms for specific ranges of Body Mass Index (BMI). A normal BMI is between 18.5 - 25. An overweight person’s BMI is between 25-30, and an obese person is one who has a BMI of above 30.
 - Link:
 - <http://gt-health-analytics.us-east-1.elasticbeanstalk.com/#/conceptset/1931/conceptset-expression>
- Cohorts
 - Lipoma Patients
 - A lipoma patient is defined as someone who as at least 1 occurrence of what had been previously defined above as any valid SNOMED-approved condition containing the word “lipoma”.
 - Link:
 - <http://gt-health-analytics.us-east-1.elasticbeanstalk.com/#/cohortdefinition/898>
 - Sample Sizes:
 - CMSDESynPUF1k : 32
 - CMSDESynPUF100k : 674
 - CMSDESynPUF23m : 7452
 - [xqi7] Overweight or Obese Patients
 - An overweight or obese patients is someone who has had a BMI categorization of above 25.
 - Link:
 - <http://gt-health-analytics.us-east-1.elasticbeanstalk.com/#/cohortdefinition/895>
 - Sample Sizes:
 - CMSDESynPUF1k : 0
 - CMSDESynPUF100k : 6,624
 - CMSDESynPUF23m : 0
 - [xqi7] Not Overweight or Obese Patients
 - Someone who falls into this category has had any procedure that led them to be included in the SynPUF database but has never had any of the SNOMED clinical findings that indicated that their BMI was ever above 25.

- Link:
<http://gt-health-analytics.us-east-1.elasticbeanstalk.com/#/cohortdefinition/894>
- Sample Sizes:
 - CMSDESynPUF1k : 835
 - CMSDESynPUF100k : 78,139
 - CMSDESynPUF23m : 2,324,944

Study Parameters:

Estimation: [xqi7] Overweight Effects on Lipoma 2

Target: [xqi7] Overweight or Obese Patients

Comparator: [xqi7] Not Overweight or Obese Patients

Outcome: [xqi7] Lipoma Patients

For this estimation, I tested 2 populations of people who are pulled from the same population set. The only differentiating characteristic is that one set of the people has been cataloged as overweight or obese and the other set have not. Using this factor as an indicator and keeping everything else constant, we look for the singular effect of overweightness on the risk of detecting lipoma. In the scope of this study, all subjects that are in both the target and comparators (which there shouldn't be any) are removed from the testing and all other factors are kept as the standard from Lab 3.

Results

Overall, it doesn't look like there's a statistically significant indication that not being overweight or obese would decrease the odds of getting lipoma. In the estimation study, 124 Overweight and obese patients were compared against 124 non-overweight or obese patients and the result were that <5 target subjects had lipoma diagnoses and 7 comparator subjects had the diagnosis. The Hazard Ratio for this estimate was 0.33, indicating that there is a possibility that being overweight or obese may actually help the patients have a lower rate of lipoma occurrences. However, the Lower Bound of the 95% percentile was 0.05 and the Upper Bound 95% percentile was 1.45, indicating massive variation and no direct causation between being overweight and having a lipoma detection.

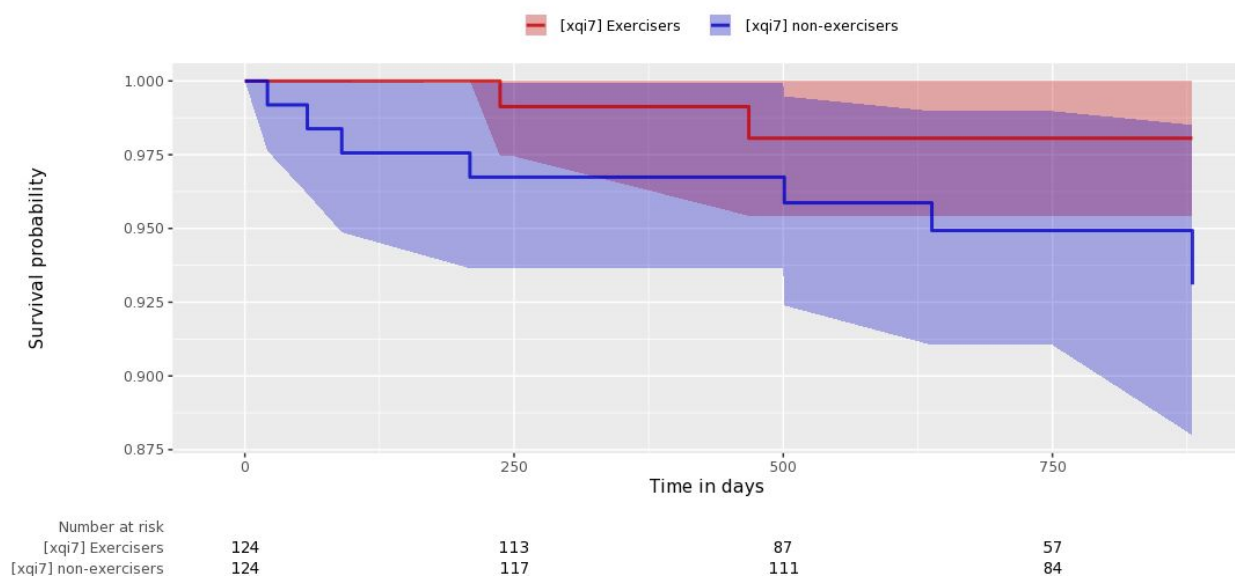


Figure 1. Kaplan Meier plot, showing survival as a function of time.

From the Kaplan Meier plot generated by the study, there were significant overlaps between the 2 curves and nothing conclusive can be said about the causality of the factors and symptoms discussed.

Characteristic	Before PS adjustment			After PS adjustment		
	Target	Comparator	Std. diff	Target	Comparator	Std. diff
	%	%		%	%	
Age group						
25-29	0.5	0.6	-0.01	<4.0	<4.0	-0.07
35-39	0.8	1.2	-0.04	<4.0	<4.0	-0.13
40-44	1.0	1.2	-0.02	<4.0	<4.0	0.07
45-49	1.8	2.6	-0.05	<4.0	4.8	-0.18
65-69	18.2	20.9	-0.07	20.2	21.0	-0.02
70-74	16.0	19.4	-0.09	17.7	19.4	-0.04
80-84	15.3	13.4	0.06	17.7	13.7	0.11
85-89	10.9	8.8	0.07	7.3	6.5	0.03
90-94	4.5	2.7	0.10	<4.0	<4.0	-0.06
Gender: female	56.6	56.9	-0.01	56.5	50.0	0.13
Race						
race = Black or African American	8.5	9.9	-0.05	10.5	12.1	-0.05
race = White	84.2	84.2	0.00	87.9	83.1	0.14
Ethnicity						
ethnicity = Hispanic or Latino	1.6	2.2	-0.04	<4.0	<4.0	-0.13
ethnicity = Not Hispanic or Latino	98.4	97.8	0.04	99.2	97.6	0.13

[xai7] Exercisers [xai7] non-exercisers

Table 1. Population characteristics comparison.

The population characteristics comparison of the target vs the comparison cohorts were almost identically matching each other, indicating that there's good isolation on that the targeted overweight and obesity traits between the two populations listed here.

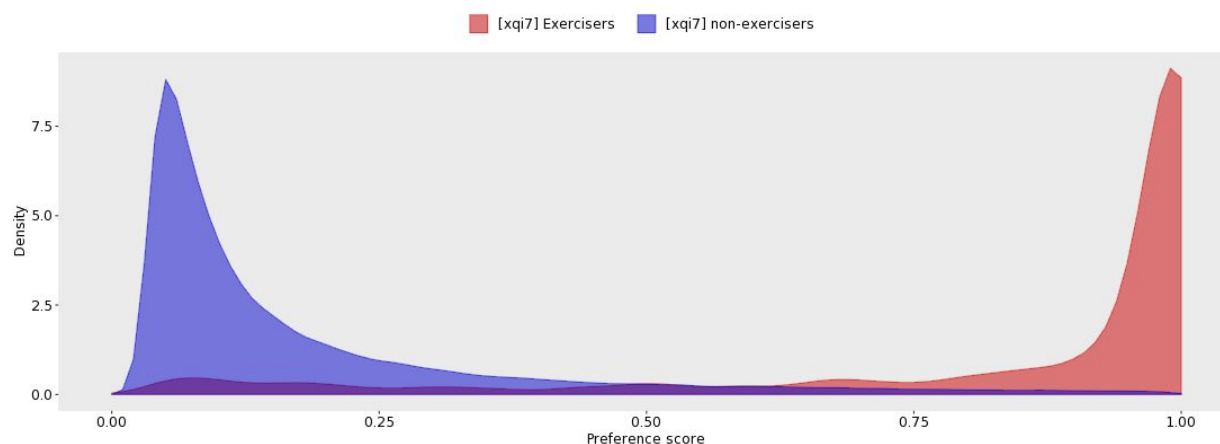


Figure 2. Preference score distribution. The preference score is a transformation of the propensity score that adjusts for differences in the sizes of the two treatment groups. A higher overlap indicates subjects in the two groups were more similar in terms of their predicted probability of receiving one treatment over the other.

According to the Propensity scores, there is very little overlap between the exerciser and non-exerciser populations.

Discussion

In this estimation study, it was found that a person's BMI value has no major correlation with the causality of lipoma in patients.

It was not surprising to me that there were no major ties between the 2, but what did surprise me was that fewer patients with overweight and obesity characteristics were even detected with lipoma when compared to the patients that were of BMI value of 25 and lower. From all of the websites that I had read, unhealthy lifestyles and overweightness had small contributions and may increase the chance of lipomas. To see a complete reversal in the expected numbers completed took me by surprise.

However, I do believe there were many limitations to the study. First and foremost, this study was conducted on a test dataset of CMSDESynPUF23m, which is a generated test set containing mocked user data. It is entirely possible that there are missing pieces of data that aren't mocked or other oddities that skews this information when compared to a real patient dataset. For example, when I was looking at the patient population in

different cohorts, the CMSDESynPUF data set for 100k actually had more hits for people with obesity than the CMSDESynPUF23m dataset.

Available CDM Sources

	Source Name	Generation Status	People	Records	Generated	Generation Duration	
▶ Generate	CMSDESynPUF100k	COMPLETE	6,624	6,624	10/05/2019 3:37 PM	00:00:10	View Reports
▶ Generate	CMSDESynPUF1k	COMPLETE	0	0	10/05/2019 3:37 PM	00:00:11	View Reports
▶ Generate	CMSDESynPUF23m	COMPLETE	0	0	10/05/2019 3:37 PM	00:00:12	View Reports

Figure 2. Overweight or Obese population cohort population counts.

Because of the oddities that were found for the CMSDESynPUF dataset, it may be prudent to retry the estimation experiment with a different dataset.

Additionally, the population that the study was performed over was widely skewed. The population distribution for the target and control cohorts were very similar to each other in terms of age and race distribution, but that does not exactly match the real world. For example, 95% of the population in the test dataset are over 65 years old and 84.2% of the test population is listed their race as White. This represents significant skew and is not representative of the America of today.

Conclusion

In this estimation study, it was found that a person's BMI value has no major correlation with the causality of lipoma in patients. However, there are several limitations to the study and the population used in the SynPUF does not resemble the age and ethnical makeup of the country we live in today, so there is room for additional iterations on this topic.

References

- Silva, Joana Cavaco. "Lipoma: Causes, Symptoms, Diagnosis, and Removal." Medical News Today, MediLexicon International, <https://www.medicalnewstoday.com/articles/322551.php>.
- Charifa, Ahmad. "Lipomas, Pathology." StatPearls [Internet]., U.S. National Library of Medicine, 27 Oct. 2018, <https://www.ncbi.nlm.nih.gov/books/NBK482343/>.
- "Lipoma." Mayo Clinic, Mayo Foundation for Medical Education and Research, 3 Apr. 2018, <https://www.mayoclinic.org/diseases-conditions/lipoma/symptoms-causes/syc-20374470>.