Effects of Metformin, Health Intervention, and Activity Levels on Weight Loss in Postmenopausal Breast Cancer Survivors

Courtney Barbato, Devin Gregoire, Khang Tong

March 10, 2022

Abstract

Heart disease and cancer rank among the top two leading causes of death in women of all races and origins. Women in the post-menopausal stage in their life are at greater risk of weight gain, which can lead to other health complications. In this project, we aim to analyze the effects of some lifestyle variables that may help women lose weight as well as determine how successful Metformin in conjunction with a weight loss program is. We used data from the Transdisciplinary Research on Energetics and Cancer study at UC San Diego. We found that Metformin with a weightloss program was effective in helping post-menopausal breast cancer survivor women lose weight, that the walkability of one's neighborhood was negatively associated with success of weight loss, and that increasing the average amount of sitting time is associated with a decrease in odds of success.

Scientific Background

Menopause is defined to be the 12 months after a women receives their last period. Symptoms of menopause include decreases in metabolism rates and oestrogen levels. Oestrogen influences the fat distribution in the body and low levels can lead to increased levels of fatigue. All these symptoms can contribute to why women after menopause can begin to experience weight gain [1].

Cancer is one of the leading cause of death in women and breast cancer ranks among one of the most common cancers for women. Women who are overweight or obese (BMI > 25) are at higher risk of developing breast cancer. Fat cells make estrogen, which can allow breast cancer cells to develop and grow. Post-menopausal women are also at an increased risk of obtaining breast cancer. Unfortunately, women who gain weight after breast cancer diagnosis are more likely to have worse outcomes and are more likely to die from any comorbidity [2]. This concern is the reason for studying post-menopausal breast cancer survivors.

Metformin is a drug that is used to control blood sugar and taken by people with type II diabetes and is typically used in conjunction with a weight loss program. Our study aims to compare the effectiveness of taking Metformin with a weight loss program, versus using only one individually, using a placebo as a control.

The Physical Activity Neighborhood Environment Scale (PANES), measures how walkable a neighborhood is [5]. It makes intuitive sense that living in a more walkable neighborhood would encourage people to walk more and therefore have a positive effect on health. However, the literature is mixed and so are our results which we will expand upon in the discussion.

Study Background

Data was collected from 2011-2017 from 578 overweight women during the Transdisciplinary Research on Energetics and Cancer Study at UC San Diego [6]. Of these 578 women, 333 were part of the Reach for Health Trial which focused on postmenopausal early-stage breast cancer survivors and how metformin treatment and/or lifestyle intervention could impact weight loss and other measurements associated with cancer risk.

Of the 333 women, only 303 had followup data and one has missing data values for certain key variables. Our focus will be on the 302 overweight women part of the Reach for Health study.

Biomarker data was collected both before and after a 6 month period during clinical visits. Activity and sedentary behavior was recorded by an accelerometer. Average daily activity and sedentary minutes were then calculated by machine learning minutes. The remaining data on the subjects were obtained through self-report questionnaires.

In our current project, our primary outcome of interest will be whether a woman was able to lose $> 5\%$ of their body weight at followup.

Questions of Interest

1. Was the experiment treatment (Metformin + Weight loss program) successful in helping post-menopausal breast cancer survivors lose $> 5\%$ of their body weight within the experiment's timeframe?
2. Does one's environment, it's walkability, increase odds of success?
3. Is the amount of sitting time per week a significant predictor for odds of success?

Hypotheses

1. Yes, the treatment was successful.
2. One's environment does have an impact on one's lifestyle health choices, and so living in a neighborhood that scores higher on walkability, aesthetics, etc, metrics, should encourage more weight loss compared to those living in neighborhoods that score lower in those metrics.
3. Sitting kills. Given what we've been told growing up and the numerous health problems that have been associated with increased sitting time, the amount of sitting time should have an adverse effect on one's odds of losing $> 5\%$ of one's body weight.

Statistical Plan

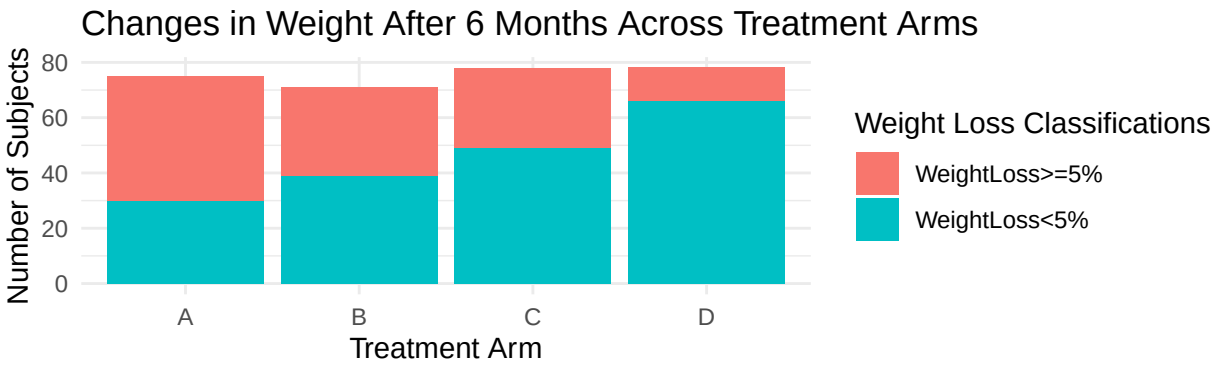
- Build a GLM model using success of weight loss as a primary outcome.
- Pare down the large list of confounders using heuristics and perform model selection on the remaining confounders using backwards selection with BIC as our penalty.
- Perform diagnostics using GOF tests, Hosmer-Lemeshow, and influence / outlier tests.

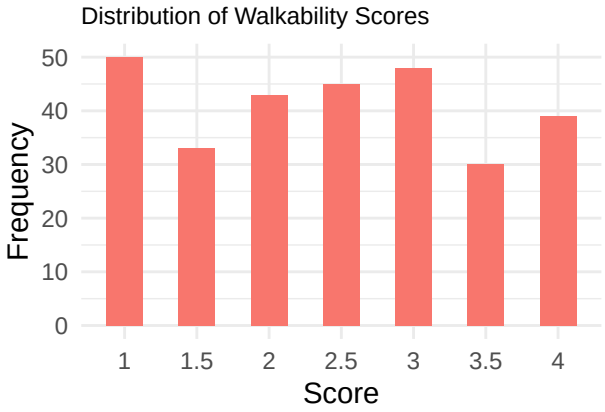
Data

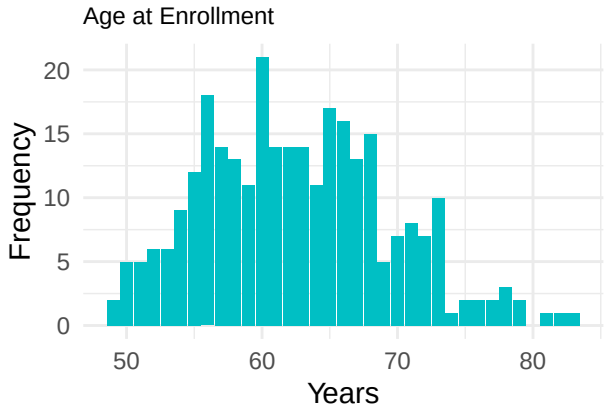
Table 1: A summary of our variables of interest and some demographic information.

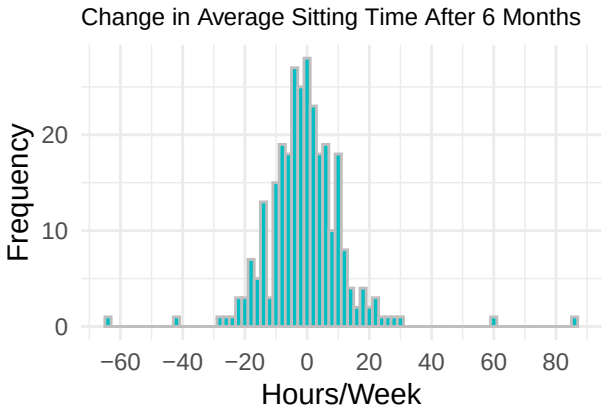
Variable	Type	Values/Units	Description
Weight Change	Binary	0,1	Our response variable and a measure of whether the participant lost more than 5% of their bodyweight.
Age at Enrollment	Continuous	49-83 years (mean: 63.1)	Age of participant at enrollment.

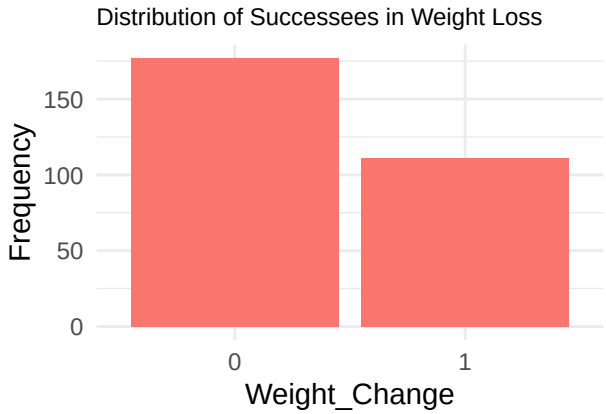
Variable	Type	Values/Units	Description
Treatment	Categorical	A = Metform + Weight Loss B = Weight Loss Only C = Metformin Only D = Placebo	Denotes which treatment arm the patient was in.
Change in Average Sitting Time	Continuous	-63.8–85.4 hours/week (mean: -0.76)	Change in weekly average sitting time after 6 months.
Walkability	Continuous	1-4	A measure of how walkable the patients' neighborhood is.











After examining our data, we realized that there were some changes we needed to make to make for a better analysis. We converted change in average sitting times from minutes per day to hours per week for more meaningful interpretation and combined race into three groups: White, Hispanic, and other. This is a reduction in the number of race categories we were given since there were so few non-white participants (239 white participants compared to 34 Hispanic and 29 other).

Upon initial visual examination of our variables of interest, we first notice that is that the Metformin and weight loss combination treat seemed to be the most successful in aiding weight loss, with the placebo being least effective. We will further quantify the significance and effect size of this difference in our model. Walkability scores almost had a uniform distribution, with change in average sitting time having an approximately normal distribution and age at enrollment being very slightly skewed to the right (older).

Confounding Variables

We proceeded with heuristics-based variable selection. We used intuition to remove variables that seemed redundant. For example, there were 16 variables, labelled p1, p2... p16. These all represented a type of physical activity, like vigorous biking, moderate biking, etc, so we combined all of these variables into one activity variable. Actually, we ended up not using this variable at all during variable selection since we realized that activity time is correlated with sitting time, one of our variables of interest.

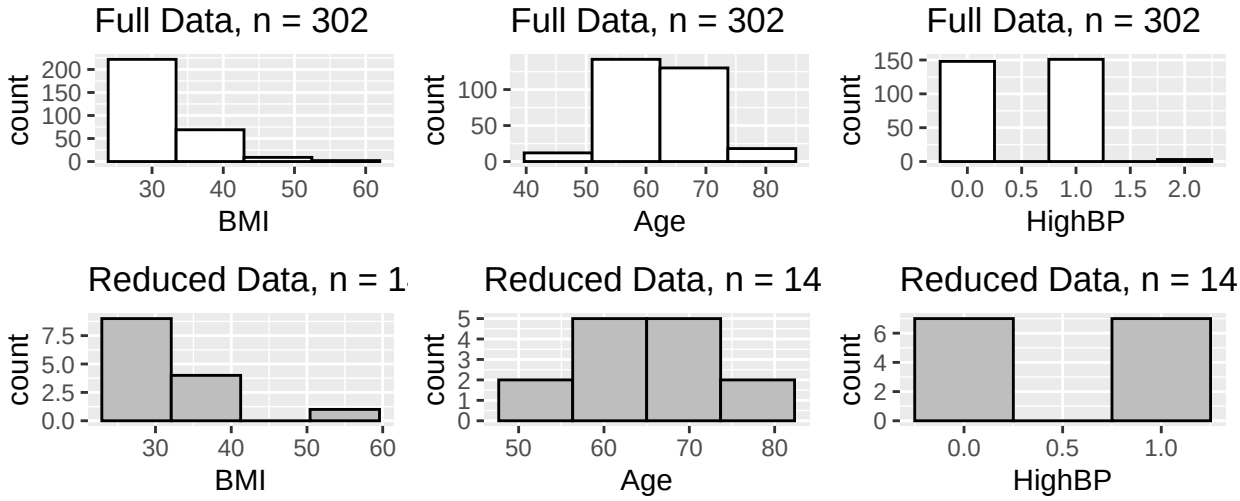
Another way that we removed variables is by throwing out variables that are downstream in causal chains. For example, we removed comorbidities such as heart attack, heart disease, and stroke since all are risk factors of high cholesterol or high triglycerides.

This left us with three different domains of confounders: lifestyle, comorbidities, and medication, which we will do variable selection for each independently.

Table 2: Our reduced list of confounding variables after removing variables by heuristics and before variable selection.

Category	Variables
Lifestyle/Demographic	Age, Race, SmokeStatus, Number of Vehicles, SleepHours, QOL (Physical), QOL (Mental), Drinks/Month, Education, Marital Status, Employment
Comorbidities	High Cholesterol, High Triglycerides, HighBP, Histology, Emphysema, Asthma, Diabetes, Epilepsy, HipFracture, Arthritis, Endocrine, Kidney, Depression
Medication	Statin, Aspirin, Acetaminophen, ACEInhib, BetaBlock, CalciumBlock, Cox2Inhib, Fosamax, Tranquilizer, H2Blocker, Antioxidant, Bcomplex, Calcium, VitD, VitC, VitE, FolicAcid Bcarot, Selenium, Glucosamine, Echinacea, GarlicSupp, Gingko

Afterwards, we looked at our full data set and noticed that 14 of our 302 observations had several missing values within the confounders we simplified our dataset to. We removed these 14 observations because they followed a similar distribution in comparison to the full dataset. The median BMI and age for the full and removed data were very similar, the proportion of having high blood pressure was the exact same, and we removed less than 5% of the data.



Data	Full	Removed
n	302	14
Median_BMI	30.09	27.72
Median_Age	62	65
Prop_HighBP	0.5	0.5

After removing 14 observations, we did BIC variable selection within each domain of confounding variables to have a stricter selection criteria and have fewer variables in the model. From lifestyle confounders, only age was significant. Acetaminophen, ACE inhibitors, Vitamin D, and Vitamin C were the only signficant medications. To our surprise, none of the comorbidities were selected to keep in the model. If we were able to talk to domain experts, perhaps we would be able to a priori say that certain comorbidities should remain in the model. However, we did not consult on this topic and so we will follow the results of our variable selection and not include comorbidities.

Once we completed variable selection, we combined all selected confounders and added back in our variables of interest.

Initial (and Final) Model

All of our variables of interest ended up being significant and so our initial model is our final model. We had no reason to suspect or test for an interaction term.

Table 4: Our final model with $> 5\%$ bodyweight loss as success at followup.

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-5.027	1.505	-3.340	0.001
Treatment A: Metformin + Weightloss	2.209	0.441	5.003	0.000
Treatment B: Weightloss Only	1.574	0.444	3.546	0.000
Treatment C: Metformin Only	1.340	0.436	3.074	0.002
Walkability	-0.282	0.144	-1.962	0.050
Change in avg sitting time	-0.038	0.013	-2.803	0.005
Age	0.053	0.022	2.473	0.013

	Estimate	Std. Error	z value	Pr(>|z|)
Acetaminophen	1.117	0.402	2.780	0.005
ACEInhib	-0.546	0.222	-2.454	0.014
VitC	-0.896	0.330	-2.716	0.007
VitD	0.927	0.383	2.418	0.016

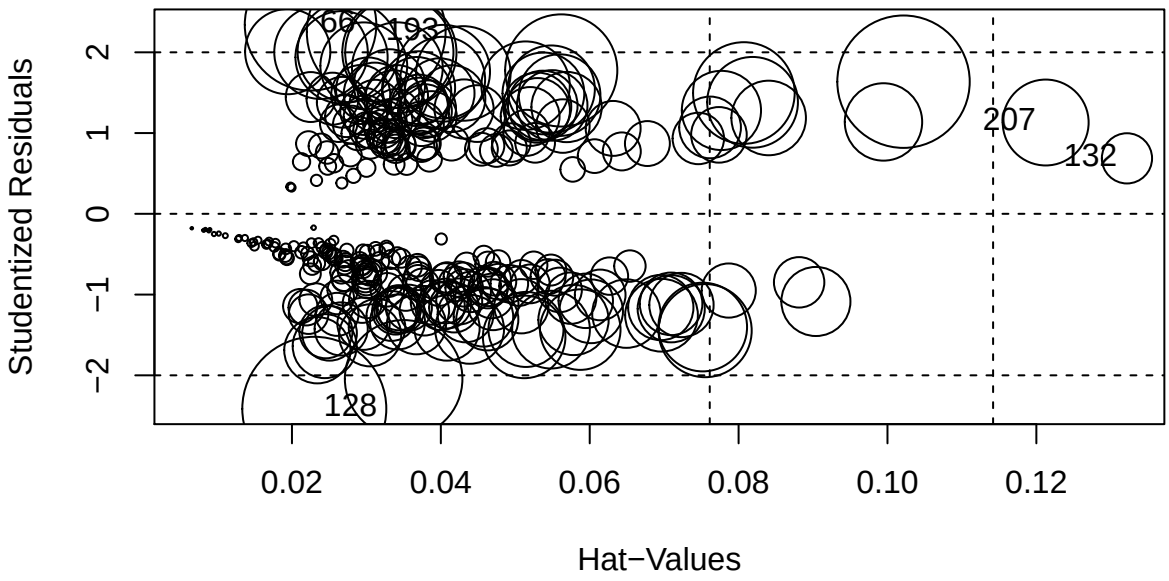
We also suspected that perhaps 5% bodyweight loss is a bit of an ambitious goal for a 6 month timespan. So we also created a model with the same variables but with $> 3\%$ bodyweight loss as the measure of a successful outcome.

Table 5: Our final model with $> 3\%$ bodyweight loss as success at followup.

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-3.372	1.500	-2.249	0.025
Treatment A: Metformin + Weightloss	2.726	0.445	6.123	0.000
Treatment B: Weightloss Only	1.699	0.421	4.033	0.000
Treatment C: Metformin Only	1.336	0.403	3.313	0.001
Walkability	-0.542	0.148	-3.653	0.000
Change in avg sitting time	-0.048	0.014	-3.330	0.001
Age	0.054	0.022	2.448	0.014
Acetaminophen	0.947	0.440	2.150	0.032
ACEInhib	-0.635	0.219	-2.905	0.004
VitC	-0.826	0.333	-2.479	0.013
VitD	0.267	0.363	0.735	0.462

Model Diagnostics

```
##
##  Hosmer and Lemeshow test (binary model)
##
## data:  df$Weight_Change, m.fivepercent.loss$fitted.values
## X-squared = 9.6811, df = 8, p-value = 0.2881
```



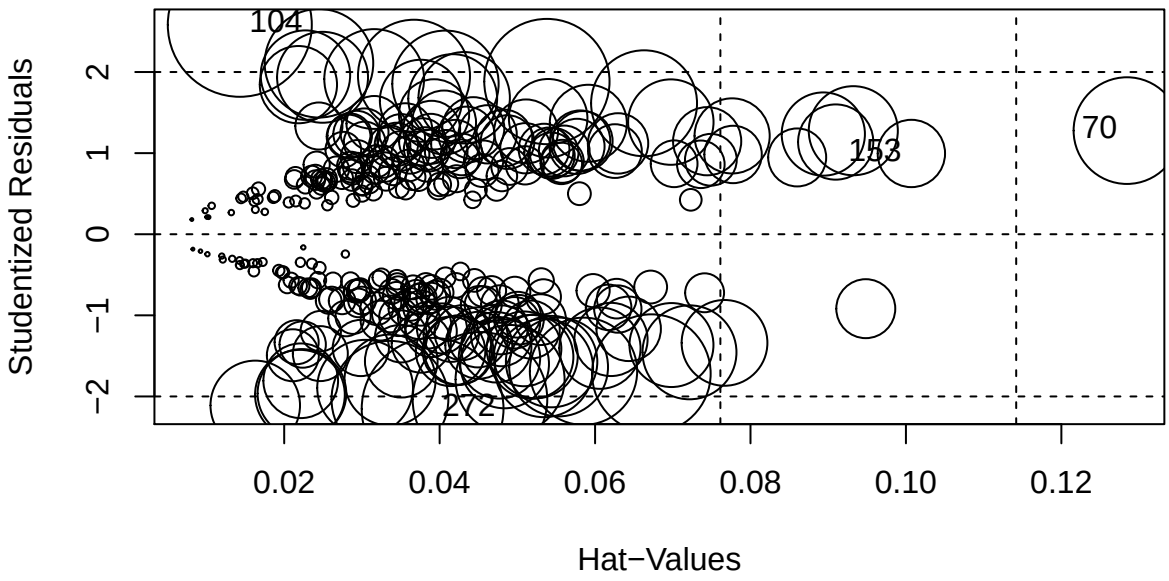

```
##         StudRes        Hat       CookD
## 66   2.3407851 0.02249869 0.026584874
## 128 -2.4145402 0.02300198 0.031865133
## 132  0.6868588 0.13216836 0.003874154
## 193  2.2506449 0.03131878 0.029588872
## 207  1.1307626 0.12125167 0.011336030
```

Various diagnostic plots were attempted to assess our model, but because we have continuous predictors, the model could not be aggregated. Goodness-of-fit was assessed using a Hosmer-Lemeshow test with the null hypothesis being that the model is a good fit for our data. Given that the p-value is 0.29, we fail to reject the null hypothesis that the model is a good fit.

Next, the model could be assessed for outliers and influential points using an influence plot. At first glance, these points are not overly concerning since the residuals aren't egregiously high or low, nor are the influential points extremely distanced from others. Observations such as 132 and 207 may be influential since they achieve the 5% weight loss goal while also having the two largest negative changes in sitting time over the course of the trial.

A sensitivity analysis was conducted to assess whether these outliers and influential points drastically change the model. The only notable change after removing these points is that the predictor walkability changes from significant to marginally significant ($p = 0.051$). Overall, the impact of these outliers and influential points is limited.

```
##
##  Hosmer and Lemeshow test (binary model)
##
## data:  df$Weight_Change3, m.threepercent.loss$fitted.values
## X-squared = 9.2183, df = 8, p-value = 0.3242
```

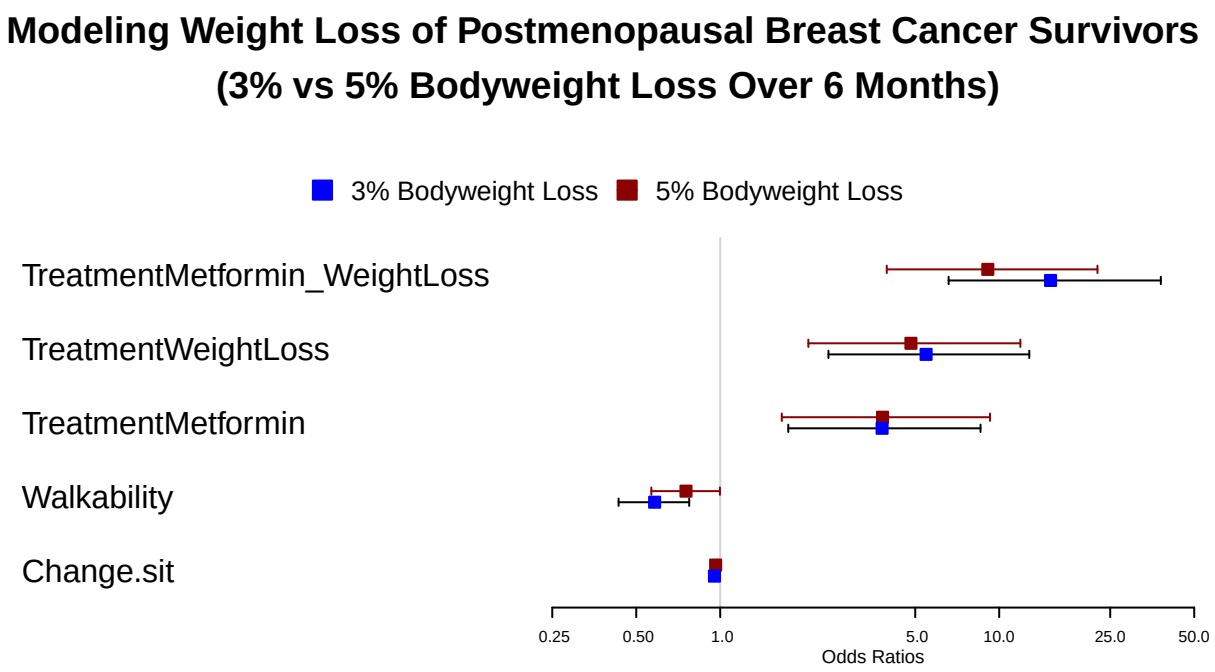



```
##        StudRes        Hat       CookD
## 70   1.2775982 0.12843761 0.016716864
## 104  2.5823479 0.01428729 0.030460397
## 153  0.9957535 0.10069920 0.006661858
## 272 -2.1522806 0.03909617 0.029531300
```

On our newly constructed 3% bodyweight loss model, the Hosmer-Lemeshow test returns a p-value of 0.32. This again indicates that we fail to reject the null hypothesis that the model is a good fit.

The influence plot marks four points of interest. Like the influence plot of the 5% weight loss model, the outlier and influential points aren't extremely distanced from the bulk of the data. Points like 70 and 153 may be influential as they achieve the 3% weight loss threshold while having one of the highest increase in sitting time and being one of the oldest patients in the study, respectively. Observation 272 may have a negative residual since this person failed to meet the 3% weight loss threshold despite being in the most effective treatment and being a relatively younger individual.

Like as before, a sensitivity test was conducted to ascertain the effect of these observations on the overall model. We found that none of our predictors of interest varied much in their estimates, and they all remained significant. This shows that removing these points does not change the interpretations or relationships in the 3% weight loss model.

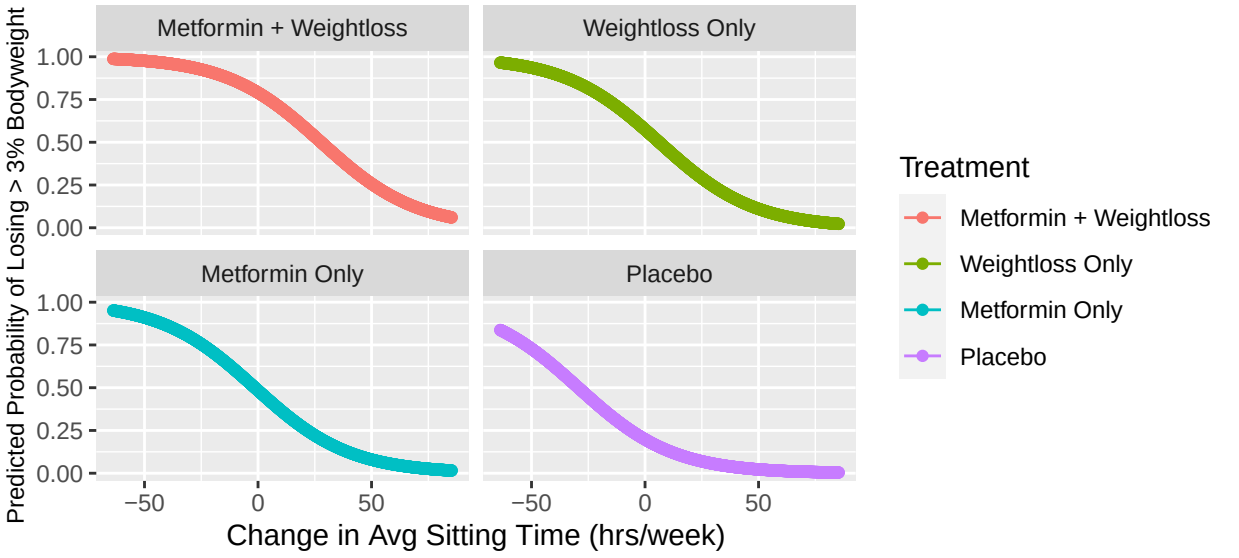


By using a forest plot we can visualize and compare the odds ratios and their respective confidence intervals between the two models. The 3% model had higher odds of weight loss success for the first two treatments - relative to the placebo - compared to the 5% weight loss model. The third treatment has a similar estimate for both models, but the 3% model has a narrower confidence interval. Contrary to our hypothesis, an increase in the rating of neighborhood walkability seems to decrease the probability of achieving 3% and 5% weight loss thresholds. The change in sit time appears to have a small effect size but confirms our original hypothesis (increased sitting time leads to decreased probability of weight loss).

Conclusion

Table 6: Exponentiated coefficients and their 95% confidence intervals from the $> 3\%$ bodyweight loss model.

	Estimate	95% CI
Treatment A: (Metformin + Weightloss)	15.277	$(6.589, 37.969)$
Treatment B: (WeightLoss only)	5.470	$(2.443, 12.818)$
Treatment C: (Metformin only)	3.804	$(1.754, 8.572)$
Walkability	0.582	$(0.432, 0.774)$
Change in avg sitting time	0.953	$(0.926, 0.979)$



The experimental treatment was a success; the combination of Metformin and guided weightloss program was the most effective. Those in treatment A had 15.28 times greater odds (95% CI: 6.59, 37.97) of successfully losing >3% bodyweight compared to the placebo group.

Walkability score of one's neighborhood was associated negatively with one's odds of success. For each unit increase in neighborhood walkability, the expected odds of successfully losing >3% bodyweight decreased by 0.58 (95% CI: 0.43, 0.77) times.

Change in average sitting time was associated negatively with one's odds of success. For each additional hour/week spent sitting on average, odds of successfully losing >3% bodyweight expected to decrease by 0.95 (95% CI: 0.93, 0.98).

Looking at the plot of our model, we notice that interestingly, in the Metformin and weightloss program treatment, if a participant did not change their average sitting time from baseline to followup (Δsit = 0), then they had an expected probability of success of about 75% whereas those in the placebo group who did not change their sitting habits have an expected probability of success of only around 25%! Metformin and the weightloss program each on their own correspond to 50% success.

Discussion

Our first hypothesis was correct in that the treatment in the study was successful. Our conjecture about walkability having a positive effect on weight loss was not the case in our model. Another study conducted in Taiwan with old people using the PANE Score also found a negative association between walkability and weight loss. The authors of that study also mention that their findings contradict a study from Canada where they found the opposite result [4]. Perhaps there are cultural or economic differences that we did not account for in the model. For example, perhaps more walkable neighborhoods are more likely to be expensive, and so only rich people can afford to live there, but perhaps richer people tend to use cars more and walk less. This should be investigated further in a future study.

The conclusion for sitting time matched our hypothesis. It makes sense that being less sedentary, and thus more active, contributes positively to health outcomes.

A possible recommendation from this project is that post-menopausal women, and especially those who are breast cancer survivors, should reduce the amount of time that they sit throughout the day, and that they should consider talking to their doctor about taking Metformin and enrolling in a weight loss program if they have type II diabetes and are currently overweight.

References

[1] Davis SR, Castelo-Branco C, Chedraui P et al. 2012, 'Understanding weight gain at menopause', Climacteric, vol. 15, no. 5, pp. 419–429

[2] Depolo Jamie. "Weight Gain After Breast Cancer Diagnosis May Be a Bigger Issue Than Thought in Australia." 13 Mar. 2020, https://www.breastcancer.org/research-news/gaining-weight-after-dx-big-problem-in-australia

[3] Donnelly JE, Blair SN, Jakicic JM, Manore MM;Rankin JW;Smith BK. "American College of Sports Medicine Position Stand. Appropriate Physical Activity Intervention Strategies for Weight Loss and Prevention of Weight Regain for Adults." Medicine and Science in Sports and Exercise, U.S. National Library of Medicine, https://pubmed.ncbi.nlm.nih.gov/19127177/.

[4] Liao, Yung et al. "Walk Score® and Its Associations with Older Adults' Health Behaviors and Outcomes." International journal of environmental research and public health vol. 16,4 622. 20 Feb. 2019, doi:10.3390/ijerph16040622

[5] Sallis, James F. et al. "Evaluating a Brief Self-Report Measure of Neighborhood Environments for Physical Activity Research and Surveillance: Physical Activity Neighborhood Environment Scale (PANES)." Journal of physical activity and health, 2010, vol. 7, 533-540. 2010.

[6] Xu, Selene Yue, et al. "Modeling Temporal Variation in Physical Activity Using Functional Principal Components Analysis." Statistics in Biosciences, vol. 11, no. 2, 2019, pp. 403–421., https://doi.org/10.1007/s12561-019-09237-3.