Report for the analysis patients receiving weight loss treatments

*yz4184 Yunlin Zhou*

# 1. Overview

In this report, we will use SAS to clean and analyze a dataset called “weight\_loss”. The dataset contains information on patients receiving weight loss treatments over multiple visits. The information in the dataset is below.

* Columns A-M contain demographic and clinical characteristics of the patients examined
* N-P contain information about the patient’s initial visit
* Q-AP contain pairs of columns for each visit: weight at visit (kg) and meds taken at each visit, up to visit 14

We are interested in: what factors are associated with overall change in weight from initial visit to the final visit? Specifically, which variables are significant predictors of overall change in BMI and having lost more than 5% of initial body weight at the final visit? In this report, we will use linear and logistic regression to do the analysis, as well as other descriptive statistics to better understand the information.

# 2. Data Cleaning Notes

1. Values excluded or changed due to typos or implausible values

* Variables’ names change in excel table.
* Change the decimals from 2 to 0 in a unique row in excel table.
* Exclude the null rows in SAS.
* Reset the levels of Ethnicity in SAS.
* Add new indicator whether the patient was on metformin at initial visit.
* Add new indicator whether the patient was on metformin at any visits.
* Add new indicator whether a patient lost more than 5% of their initial weight by their final visit.
* Add new variable of BMI at every visit.
* Add new variable of overall change in BMI.
* Apply variable labels.
* Apply user-defined formats.

1. Levels of a categorical variable combined (for analysis purposes)

* Gender: Male, Female.
* Ethnicity: White, Black, Asian, Other and Decline.
* Demographic and clinical variables: 1=Yes, 0=No.

| **Number of Variable Levels** | | | | |
| --- | --- | --- | --- | --- |
| **Variable** | **Label** | **Levels** | **Missing Levels** | **Nonmissing Levels** |
| **Gender** | Gender | 2 | 0 | 2 |
| **Ethnicity** | Ethnicity | 5 | 0 | 5 |
| **HTN** | Hypertension diagnosis at initial visit | 2 | 0 | 2 |
| **Prediabetes** | Prediabetes diagnosis at initial visit | 3 | 1 | 2 |
| **T2DM** | Type II Diabetes diagnosis at initial visit | 2 | 0 | 2 |
| **HLD** | Hyperlipidemia diagnosis at initial visit | 2 | 0 | 2 |
| **CVD** | Cardiovascular disease diagnosis at initial visit | 2 | 0 | 2 |
| **NASH** | Non-Alcoholic Fatty Liver Disease/Non-Alcoholic Steatohepatitis diagnosis at initial visit | 2 | 0 | 2 |
| **Hypothyroidism** | Hypothyroidism diagnosis at initial visit | 2 | 0 | 2 |
| **OSA** | Obstructive sleep apnea diagnosis at initial visit | 2 | 0 | 2 |
| **Psych** | Psychiatric diagnosis at initial visit | 2 | 0 | 2 |
| **PCOS** | Polycystic Ovary Syndrome diagnosis at initial visit | 2 | 0 | 2 |

*Note: There is a patient having no prediabetes information. So the Prediabetes showed 3 levels in this table.*

# 3. Descriptive Statistics

**Frequency Table for Demographic and clinical variables (sample size 372)**

| **Gender** | | | | |
| --- | --- | --- | --- | --- |
| **Gender** | **Frequency** | **Percent** | **Cumulative Frequency** | **Cumulative Percent** |
| **F** | 271 | 72.85 | 271 | 72.85 |
| **M** | 101 | 27.15 | 372 | 100.00 |

| **Ethnicity** | | | | |
| --- | --- | --- | --- | --- |
| **Ethnicity** | **Frequency** | **Percent** | **Cumulative Frequency** | **Cumulative Percent** |
| **Asian** | 3 | 0.81 | 3 | 0.81 |
| **Black** | 16 | 4.30 | 19 | 5.11 |
| **Decline** | 204 | 54.84 | 223 | 59.95 |
| **Other** | 20 | 5.38 | 243 | 65.32 |
| **White** | 129 | 34.68 | 372 | 100.00 |

| **Hypertension diagnosis at initial visit** | | | | |
| --- | --- | --- | --- | --- |
| **HTN** | **Frequency** | **Percent** | **Cumulative Frequency** | **Cumulative Percent** |
| **No** | 252 | 67.74 | 252 | 67.74 |
| **Yes** | 120 | 32.26 | 372 | 100.00 |

| **Prediabetes diagnosis at initial visit** | | | | |
| --- | --- | --- | --- | --- |
| **Prediabetes** | **Frequency** | **Percent** | **Cumulative Frequency** | **Cumulative Percent** |
| **No** | 243 | 65.50 | 243 | 65.50 |
| **Yes** | 128 | 34.50 | 371 | 100.00 |
| **Frequency Missing = 1** | | | | |

| **Type II Diabetes diagnosis at initial visit** | | | | |
| --- | --- | --- | --- | --- |
| **T2DM** | **Frequency** | **Percent** | **Cumulative Frequency** | **Cumulative Percent** |
| **No** | 314 | 84.41 | 314 | 84.41 |
| **Yes** | 58 | 15.59 | 372 | 100.00 |

| **Hyperlipidemia diagnosis at initial visit** | | | | |
| --- | --- | --- | --- | --- |
| **HLD** | **Frequency** | **Percent** | **Cumulative Frequency** | **Cumulative Percent** |
| **No** | 256 | 68.82 | 256 | 68.82 |
| **Yes** | 116 | 31.18 | 372 | 100.00 |

| **Cardiovascular disease diagnosis at initial visit** | | | | |
| --- | --- | --- | --- | --- |
| **CVD** | **Frequency** | **Percent** | **Cumulative Frequency** | **Cumulative Percent** |
| **No** | 350 | 94.09 | 350 | 94.09 |
| **Yes** | 22 | 5.91 | 372 | 100.00 |

| **Non-Alcoholic Fatty Liver Disease/Non-Alcoholic Steatohepatitis diagnosis at initial visit** | | | | |
| --- | --- | --- | --- | --- |
| **NASH** | **Frequency** | **Percent** | **Cumulative Frequency** | **Cumulative Percent** |
| **No** | 349 | 93.82 | 349 | 93.82 |
| **Yes** | 23 | 6.18 | 372 | 100.00 |

| **Hypothyroidism diagnosis at initial visit** | | | | |
| --- | --- | --- | --- | --- |
| **Hypothyroidism** | **Frequency** | **Percent** | **Cumulative Frequency** | **Cumulative Percent** |
| **No** | 296 | 79.57 | 296 | 79.57 |
| **Yes** | 76 | 20.43 | 372 | 100.00 |

| **Obstructive sleep apnea diagnosis at initial visit** | | | | |
| --- | --- | --- | --- | --- |
| **OSA** | **Frequency** | **Percent** | **Cumulative Frequency** | **Cumulative Percent** |
| **No** | 301 | 80.91 | 301 | 80.91 |
| **Yes** | 71 | 19.09 | 372 | 100.00 |

| **Psychiatric diagnosis at initial visit** | | | | |
| --- | --- | --- | --- | --- |
| **Psych** | **Frequency** | **Percent** | **Cumulative Frequency** | **Cumulative Percent** |
| **No** | 246 | 66.13 | 246 | 66.13 |
| **Yes** | 126 | 33.87 | 372 | 100.00 |

| **Polycystic Ovary Syndrome diagnosis at initial visit** | | | | |
| --- | --- | --- | --- | --- |
| **PCOS** | **Frequency** | **Percent** | **Cumulative Frequency** | **Cumulative Percent** |
| **No** | 348 | 93.55 | 348 | 93.55 |
| **Yes** | 24 | 6.45 | 372 | 100.00 |

**Descriptive Statistics for BMI at initial visit (sample size 372)**



| **Analysis Variable : BMI1 Patient’s BMI at initial visit** | | |
| --- | --- | --- |
| **Median** | **25th Pctl** | **75th Pctl** |
| 34.05 | 30.49 | 37.87 |

*Note: Since the BMI at initial visit is highly skewed, we report median and interquartile range instead of Mean (SD) in this table.*

**Frequency Table for Metformin use at initial visit (sample size 372)**

| **Number of Variable Levels** | | |
| --- | --- | --- |
| **Variable** | **Label** | **Levels** |
| **Met1** | Indicater whether the patient was on metformin at initial visit | 2 |

| **Indicater whether the patient was on metformin at initial visit** | | | | |
| --- | --- | --- | --- | --- |
| **Met1** | **Frequency** | **Percent** | **Cumulative Frequency** | **Cumulative Percent** |
| **No** | 81 | 21.77 | 81 | 21.77 |
| **Yes** | 291 | 78.23 | 372 | 100.00 |

*Note: There are no variables with VIF>10 in this model.*

# 4. Analysis

**Using Linear Regression to analyze which variables are significant predictors of overall change in BMI.**

1. Check for collinearity of predictor variables (except for the categorical variables)

|  |  |
| --- | --- |
| **Number of Observations Read** | 372 |
| **Number of Observations Used** | 371 |
| **Number of Observations with Missing Values** | 1 |

| **Analysis of Variance** | | | | | |
| --- | --- | --- | --- | --- | --- |
| **Source** | **DF** | **Sum of Squares** | **Mean Square** | **F Value** | **Pr > F** |
| **Model** | 11 | 351.97099 | 31.99736 | 2.55 | 0.0041 |
| **Error** | 359 | 4504.61616 | 12.54768 |  |  |
| **Corrected Total** | 370 | 4856.58715 |  |  |  |

|  |  |  |  |
| --- | --- | --- | --- |
| **Root MSE** | 3.54227 | **R-Square** | 0.0725 |
| **Dependent Mean** | 3.52246 | **Adj R-Sq** | 0.0441 |
| **Coeff Var** | 100.56232 |  |  |

| **Parameter Estimates** | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Variable** | **Label** | **DF** | **Parameter Estimate** | **Standard Error** | **t Value** | **Pr > |t|** | **Variance Inflation** |
| **Intercept** | Intercept | 1 | 2.82828 | 0.78249 | 3.61 | 0.0003 | 0 |
| **HTN** | Hypertension diagnosis at initial visit | 1 | 0.56398 | 0.43037 | 1.31 | 0.1909 | 1.19838 |
| **Prediabetes** | Prediabetes diagnosis at initial visit | 1 | 0.06840 | 0.42308 | 0.16 | 0.8717 | 1.19598 |
| **T2DM** | Type II Diabetes diagnosis at initial visit | 1 | -0.39763 | 0.58040 | -0.69 | 0.4937 | 1.31370 |
| **HLD** | Hyperlipidemia diagnosis at initial visit | 1 | -0.33838 | 0.41713 | -0.81 | 0.4178 | 1.10562 |
| **CVD** | Cardiovascular disease diagnosis at initial visit | 1 | -0.73145 | 0.81349 | -0.90 | 0.3692 | 1.09148 |
| **NASH** | Non-Alcoholic Fatty Liver Disease/Non-Alcoholic Steatohepatitis diagnosis at initial visit | 1 | 1.62473 | 0.78087 | 2.08 | 0.0382 | 1.04839 |
| **Hypothyroidism** | Hypothyroidism diagnosis at initial visit | 1 | 0.22414 | 0.46214 | 0.48 | 0.6280 | 1.02861 |
| **OSA** | Obstructive sleep apnea diagnosis at initial visit | 1 | 1.43886 | 0.49234 | 2.92 | 0.0037 | 1.10909 |
| **Psych** | Psychiatric diagnosis at initial visit | 1 | 0.51590 | 0.39403 | 1.31 | 0.1913 | 1.02955 |
| **PCOS** | Polycystic Ovary Syndrome diagnosis at initial visit | 1 | -1.62945 | 0.76501 | -2.13 | 0.0339 | 1.04699 |
| **Metformin** | Indicator whether the patient was on metformin at any visits | 1 | 0.22207 | 0.76280 | 0.29 | 0.7711 | 1.04095 |







*Note: There are no variables with VIF>10 in this model.*

1. Automatic Model Selection Procedure

***The selected model is the model at the last step (Step 1).***

|  |  |
| --- | --- |
| **Effects:** | Intercept OSA |

| **Analysis of Variance** | | | | |
| --- | --- | --- | --- | --- |
| **Source** | **DF** | **Sum of Squares** | **Mean Square** | **F Value** |
| **Model** | 1 | 174.90382 | 174.90382 | 13.79 |
| **Error** | 369 | 4681.68333 | 12.68749 |  |
| **Corrected Total** | 370 | 4856.58715 |  |  |

|  |  |
| --- | --- |
| **Root MSE** | 3.56195 |
| **Dependent Mean** | 3.52246 |
| **R-Square** | 0.0360 |
| **Adj R-Sq** | 0.0334 |
| **AIC** | 1317.56326 |
| **AICC** | 1317.62866 |
| **SBC** | 952.39567 |

| **Parameter Estimates** | | | | |
| --- | --- | --- | --- | --- |
| **Parameter** | **DF** | **Estimate** | **Standard Error** | **t Value** |
| **Intercept** | 1 | 3.188436 | 0.205649 | 15.50 |
| **OSA** | 1 | 1.745407 | 0.470094 | 3.71 |

The final model should include Gender, Ethnicity and OSA as predictors.

1. Run the full and the final model and interpret

**Full model**

| **Class Level Information** | | |
| --- | --- | --- |
| **Class** | **Levels** | **Values** |
| **Gender** | 2 | F M |
| **Ethnicity** | 5 | Asian Black Decline Other White |

|  |  |
| --- | --- |
| **Number of Observations Read** | 372 |
| **Number of Observations Used** | 371 |

| **Source** | **DF** | **Sum of Squares** | **Mean Square** | **F Value** | **Pr > F** |
| --- | --- | --- | --- | --- | --- |
| **Model** | 11 | 351.970987 | 31.997362 | 2.55 | 0.0041 |
| **Error** | 359 | 4504.616161 | 12.547677 |  |  |
| **Corrected Total** | 370 | 4856.587148 |  |  |  |

| **R-Square** | **Coeff Var** | **Root MSE** | **bmi\_change Mean** |
| --- | --- | --- | --- |
| 0.072473 | 100.5623 | 3.542270 | 3.522462 |

| **Source** | **DF** | **Type I SS** | **Mean Square** | **F Value** | **Pr > F** |
| --- | --- | --- | --- | --- | --- |
| **HTN** | 1 | 37.1227594 | 37.1227594 | 2.96 | 0.0863 |
| **Prediabetes** | 1 | 14.1079653 | 14.1079653 | 1.12 | 0.2897 |
| **T2DM** | 1 | 2.2672054 | 2.2672054 | 0.18 | 0.6710 |
| **HLD** | 1 | 3.3881237 | 3.3881237 | 0.27 | 0.6036 |
| **CVD** | 1 | 9.5151654 | 9.5151654 | 0.76 | 0.3844 |
| **NASH** | 1 | 71.1545194 | 71.1545194 | 5.67 | 0.0178 |
| **Hypothyroidism** | 1 | 2.1949096 | 2.1949096 | 0.17 | 0.6760 |
| **OSA** | 1 | 133.6171748 | 133.6171748 | 10.65 | 0.0012 |
| **Psych** | 1 | 20.2794690 | 20.2794690 | 1.62 | 0.2044 |
| **PCOS** | 1 | 57.2602486 | 57.2602486 | 4.56 | 0.0333 |
| **Metformin** | 1 | 1.0634462 | 1.0634462 | 0.08 | 0.7711 |

| **Source** | | **DF** | | **Type III SS** | | **Mean Square** | | **F Value** | | **Pr > F** | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **HTN** | | 1 | | 21.5480583 | | 21.5480583 | | 1.72 | | 0.1909 | |
| **Prediabetes** | | 1 | | 0.3279519 | | 0.3279519 | | 0.03 | | 0.8717 | |
| **T2DM** | | 1 | | 5.8893681 | | 5.8893681 | | 0.47 | | 0.4937 | |
| **HLD** | | 1 | | 8.2569648 | | 8.2569648 | | 0.66 | | 0.4178 | |
| **CVD** | | 1 | | 10.1443087 | | 10.1443087 | | 0.81 | | 0.3692 | |
| **NASH** | | 1 | | 54.3216650 | | 54.3216650 | | 4.33 | | 0.0382 | |
| **Hypothyroidism** | | 1 | | 2.9515121 | | 2.9515121 | | 0.24 | | 0.6280 | |
| **OSA** | | 1 | | 107.1710042 | | 107.1710042 | | 8.54 | | 0.0037 | |
| **Psych** | | 1 | | 21.5099098 | | 21.5099098 | | 1.71 | | 0.1913 | |
| **PCOS** | | 1 | | 56.9253813 | | 56.9253813 | | 4.54 | | 0.0339 | |
| **Metformin** | | 1 | | 1.0634462 | | 1.0634462 | | 0.08 | | 0.7711 | |

| **Parameter** | **Estimate** | **Standard Error** | **t Value** | **Pr > |t|** |
| --- | --- | --- | --- | --- |
| **Intercept** | 2.828276908 | 0.78249466 | 3.61 | 0.0003 |
| **HTN** | 0.563976779 | 0.43036686 | 1.31 | 0.1909 |
| **Prediabetes** | 0.068398311 | 0.42307947 | 0.16 | 0.8717 |
| **T2DM** | -0.397633023 | 0.58040316 | -0.69 | 0.4937 |
| **HLD** | -0.338377320 | 0.41713122 | -0.81 | 0.4178 |
| **CVD** | -0.731446037 | 0.81349097 | -0.90 | 0.3692 |
| **NASH** | 1.624732604 | 0.78086693 | 2.08 | 0.0382 |
| **Hypothyroidism** | 0.224138439 | 0.46214231 | 0.48 | 0.6280 |
| **OSA** | 1.438863888 | 0.49233729 | 2.92 | 0.0037 |
| **Psych** | 0.515895633 | 0.39402541 | 1.31 | 0.1913 |
| **PCOS** | -1.629449245 | 0.76501427 | -2.13 | 0.0339 |
| **Metformin** | 0.222069667 | 0.76280432 | 0.29 | 0.7711 |

From the full model we can see that only NASH OSA PCOS are signifficant predictors. OSA is a very significant predictor.

**Final model**

| **Class Level Information** | | |
| --- | --- | --- |
| **Class** | **Levels** | **Values** |
| **Gender** | 2 | F M |
| **Ethnicity** | 5 | Asian Black Decline Other White |

|  |  |
| --- | --- |
| **Number of Observations Read** | 372 |
| **Number of Observations Used** | 372 |

| **Source** | **DF** | **Sum of Squares** | **Mean Square** | **F Value** | **Pr > F** |
| --- | --- | --- | --- | --- | --- |
| **Model** | 1 | 175.649877 | 175.649877 | 13.88 | 0.0002 |
| **Error** | 370 | 4682.586440 | 12.655639 |  |  |
| **Corrected Total** | 371 | 4858.236317 |  |  |  |

| **R-Square** | **Coeff Var** | **Root MSE** | **bmi\_change Mean** |
| --- | --- | --- | --- |
| 0.036155 | 101.0932 | 3.557476 | 3.519006 |

| **Source** | **DF** | **Type I SS** | **Mean Square** | **F Value** | **Pr > F** |
| --- | --- | --- | --- | --- | --- |
| **OSA** | 1 | 175.6498774 | 175.6498774 | 13.88 | 0.0002 |

| **Source** | **DF** | **Type III SS** | **Mean Square** | **F Value** | **Pr > F** |
| --- | --- | --- | --- | --- | --- |
| **OSA** | 1 | 175.6498774 | 175.6498774 | 13.88 | 0.0002 |

| **Parameter** | **Estimate** | **Standard Error** | **t Value** | **Pr > |t|** |
| --- | --- | --- | --- | --- |
| **Intercept** | 3.185273279 | 0.20504954 | 15.53 | <.0001 |
| **OSA** | 1.748569800 | 0.46935461 | 3.73 | 0.0002 |



1. Residual Check



**Conclusion:**

The residuals are not obviously skewed. We therefore proceed the normality assumption.

The F-test p-value is significant, indicating that at least one X-variable is a significant predictor of the outcome.

The R-squared values are not close to 1 in both full model and the final model, indicating that this model explains a low degree of variation in the outcome, BMI change.

OSA is a significant predictor of BMI change at 5% significant level.

The 95% confidence interval for OSA is (0.8286, 2.6685).

For people with Obstructive sleep apnea diagnosis at initial visit, BMI change increases by 1.75 units, on average.

OSA has a negative effect on decrease BMI.

**Using Logistic Regression to analyze which variables are significant predictors of having lost more than 5% of initial body weight at the final visit.**

1. Automatic Model Selection Procedure

| **Model Information** | | |
| --- | --- | --- |
| **Data Set** | WORK.WEIGHT\_LOSS |  |
| **Response Variable** | loss5 | Indicator whether the patient lost 5% weight |
| **Number of Response Levels** | 2 |  |
| **Model** | binary logit |  |
| **Optimization Technique** | Fisher's scoring |  |

|  |  |
| --- | --- |
| **Number of Observations Read** | 372 |
| **Number of Observations Used** | 371 |

| **Response Profile** | | |
| --- | --- | --- |
| **Ordered Value** | **loss5** | **Total Frequency** |
| **1** | No | 103 |
| **2** | Yes | 268 |

***Probability modeled is loss5='No'.***

*Note: 1 observation was deleted due to missing values for the response or explanatory variables.*

|  |
| --- |
| ***Stepwise Selection Procedure*** |

|  |
| --- |
| ***Step 0. Intercept entered:*** |

| **Model Convergence Status** |
| --- |
| Convergence criterion (GCONV=1E-8) satisfied. |

|  |  |  |
| --- | --- | --- |
| **-2 Log L** | = | 438.299 |

| **Residual Chi-Square Test** | | |
| --- | --- | --- |
| **Chi-Square** | **DF** | **Pr > ChiSq** |
| 12.9715 | 11 | 0.2952 |

|  |
| --- |
| ***Step 1. Effect OSA entered:*** |

| **Model Convergence Status** |
| --- |
| Convergence criterion (GCONV=1E-8) satisfied. |

| **Model Fit Statistics** | | |
| --- | --- | --- |
| **Criterion** | **Intercept Only** | **Intercept and Covariates** |
| **AIC** | 440.299 | 435.070 |
| **SC** | 444.215 | 442.902 |
| **-2 Log L** | 438.299 | 431.070 |

| **Testing Global Null Hypothesis: BETA=0** | | | |
| --- | --- | --- | --- |
| **Test** | **Chi-Square** | **DF** | **Pr > ChiSq** |
| **Likelihood Ratio** | 7.2287 | 1 | 0.0072 |
| **Score** | 6.5912 | 1 | 0.0102 |
| **Wald** | 6.2929 | 1 | 0.0121 |

| **Residual Chi-Square Test** | | |
| --- | --- | --- |
| **Chi-Square** | **DF** | **Pr > ChiSq** |
| 6.2349 | 10 | 0.7952 |

|  |  |
| --- | --- |
| **Note:** | No effects for the model in Step 1 are removed. |

|  |  |
| --- | --- |
| **Note:** | No (additional) effects met the 0.05 significance level for entry into the model. |

| **Summary of Stepwise Selection** | | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Step** | **Effect** | | **DF** | **Number In** | **Score Chi-Square** | **Wald Chi-Square** | **Pr > ChiSq** | **Variable Label** |
| **Entered** | **Removed** |
| **1** | **OSA** |  | 1 | 1 | 6.5912 |  | 0.0102 | Obstructive sleep apnea diagnosis at initial visit |

| **Analysis of Maximum Likelihood Estimates** | | | | | |
| --- | --- | --- | --- | --- | --- |
| **Parameter** | **DF** | **Estimate** | **Standard Error** | **Wald Chi-Square** | **Pr > ChiSq** |
| **Intercept** | 1 | -0.8157 | 0.1252 | 42.4467 | <.0001 |
| **OSA** | 1 | -0.8807 | 0.3511 | 6.2929 | 0.0121 |

| **Odds Ratio Estimates** | | | |
| --- | --- | --- | --- |
| **Effect** | **Point Estimate** | **95% Wald Confidence Limits** | |
| **OSA** | 0.414 | 0.208 | 0.825 |

| **Association of Predicted Probabilities and Observed Responses** | | | |
| --- | --- | --- | --- |
| **Percent Concordant** | 20.0 | **Somers' D** | 0.117 |
| **Percent Discordant** | 8.3 | **Gamma** | 0.414 |
| **Percent Tied** | 71.7 | **Tau-a** | 0.047 |
| **Pairs** | 27604 | **c** | 0.559 |







|  |
| --- |
| The model selected by stepwise function only includes OSA. |

1. Run the full and the final model and interpret

**Full mod**

| **Testing Global Null Hypothesis: BETA=0** | | | |
| --- | --- | --- | --- |
| **Test** | **Chi-Square** | **DF** | **Pr > ChiSq** |
| **Likelihood Ratio** | 13.3595 | 11 | 0.2705 |
| **Score** | 12.9715 | 11 | 0.2952 |
| **Wald** | 12.3125 | 11 | 0.3406 |

| **Analysis of Maximum Likelihood Estimates** | | | | | |
| --- | --- | --- | --- | --- | --- |
| **Parameter** | **DF** | **Estimate** | **Standard Error** | **Wald Chi-Square** | **Pr > ChiSq** |
| **Intercept** | 1 | -0.6817 | 0.4820 | 2.0000 | 0.1573 |
| **HTN** | 1 | -0.1030 | 0.2828 | 0.1328 | 0.7156 |
| **Prediabetes** | 1 | -0.2034 | 0.2764 | 0.5417 | 0.4617 |
| **T2DM** | 1 | 0.0988 | 0.3690 | 0.0717 | 0.7889 |
| **HLD** | 1 | -0.2810 | 0.2780 | 1.0218 | 0.3121 |
| **CVD** | 1 | 0.3508 | 0.5064 | 0.4800 | 0.4884 |
| **NASH** | 1 | 0.1722 | 0.4925 | 0.1223 | 0.7266 |
| **Hypothyroidism** | 1 | 0.0271 | 0.2965 | 0.0084 | 0.9271 |
| **OSA** | 1 | -0.7419 | 0.3658 | 4.1127 | 0.0426 |
| **Psych** | 1 | 0.1875 | 0.2508 | 0.5586 | 0.4548 |
| **PCOS** | 1 | 0.6940 | 0.4398 | 2.4895 | 0.1146 |
| **Metformin** | 1 | -0.1559 | 0.4680 | 0.1110 | 0.7390 |

| **Odds Ratio Estimates** | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Effect** | | **Point Estimate** | | **95% Wald Confidence Limits** | | | |
| **HTN** | | 0.902 | | 0.518 | | 1.570 | |
| **Prediabetes** | | 0.816 | | 0.475 | | 1.402 | |
| **T2DM** | | 1.104 | | 0.536 | | 2.275 | |
| **HLD** | | 0.755 | | 0.438 | | 1.302 | |
| **CVD** | | 1.420 | | 0.526 | | 3.832 | |
| **NASH** | | 1.188 | | 0.452 | | 3.119 | |
| **Hypothyroidism** | | 1.028 | | 0.575 | | 1.837 | |
| **OSA** | | 0.476 | | 0.232 | | 0.975 | |
| **Psych** | | 1.206 | | 0.738 | | 1.972 | |
| **PCOS** | | 2.002 | | 0.845 | | 4.740 | |
| **Metformin** | | 0.856 | | 0.342 | | 2.141 | |

| **Association of Predicted Probabilities and Observed Responses** | | | |
| --- | --- | --- | --- |
| **Percent Concordant** | 61.0 | **Somers' D** | 0.246 |
| **Percent Discordant** | 36.4 | **Gamma** | 0.253 |
| **Percent Tied** | 2.7 | **Tau-a** | 0.099 |
| **Pairs** | 27604 | **c** | 0.623 |



| **Partition for the Hosmer and Lemeshow Test** | | | | | |
| --- | --- | --- | --- | --- | --- |
| **Group** | **Total** | **loss5 = No** | | **loss5 = Yes** | |
| **Observed** | **Expected** | **Observed** | **Expected** |
| **1** | 37 | 6 | 4.97 | 31 | 32.03 |
| **2** | 36 | 5 | 6.39 | 31 | 29.61 |
| **3** | 37 | 8 | 8.34 | 29 | 28.66 |
| **4** | 45 | 12 | 11.38 | 33 | 33.62 |
| **5** | 37 | 10 | 10.27 | 27 | 26.73 |
| **6** | 57 | 15 | 17.16 | 42 | 39.84 |
| **7** | 40 | 14 | 12.64 | 26 | 27.36 |
| **8** | 37 | 13 | 12.72 | 24 | 24.28 |
| **9** | 45 | 20 | 19.12 | 25 | 25.88 |

| **Hosmer and Lemeshow Goodness-of-Fit Test** | | |
| --- | --- | --- |
| **Chi-Square** | **DF** | **Pr > ChiSq** |
| 1.3690 | 7 | 0.9865 |



From the full model we can see that only OSA is the significant predictor.

**Final model**

| **Model Information** | | |
| --- | --- | --- |
| **Data Set** | WORK.WEIGHT\_LOSS |  |
| **Response Variable** | loss5 | Indicator whether the patient lost 5% weight |
| **Number of Response Levels** | 2 |  |
| **Model** | binary logit |  |
| **Optimization Technique** | Fisher's scoring |  |

|  |  |
| --- | --- |
| **Number of Observations Read** | 372 |
| **Number of Observations Used** | 372 |

| **Response Profile** | | |
| --- | --- | --- |
| **Ordered Value** | **loss5** | **Total Frequency** |
| **1** | No | 103 |
| **2** | Yes | 269 |

|  |
| --- |
| ***Probability modeled is loss5='No'.*** |

| **Model Convergence Status** |
| --- |
| Convergence criterion (GCONV=1E-8) satisfied. |

| **Model Fit Statistics** | | |
| --- | --- | --- |
| **Criterion** | **Intercept Only** | **Intercept and Covariates** |
| **AIC** | 440.948 | 435.801 |
| **SC** | 444.867 | 443.639 |
| **-2 Log L** | 438.948 | 431.801 |

| **Testing Global Null Hypothesis: BETA=0** | | | | | | | |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Test** | | | **Chi-Square** | | **DF** | **Pr > ChiSq** | |
| **Likelihood Ratio** | | | 7.1471 | | 1 | 0.0075 | |
| **Score** | | | 6.5179 | | 1 | 0.0107 | |
| **Wald** | | | 6.2257 | | 1 | 0.0126 | |
| **Analysis of Maximum Likelihood Estimates** | | | | | | | | | | |
| **Parameter** | | **DF** | **Estimate** | | **Standard Error** | | | **Wald Chi-Square** | | **Pr > ChiSq** |
| **Intercept** | | 1 | -0.8205 | | 0.1251 | | | 43.0104 | | <.0001 |
| **OSA** | | 1 | -0.8759 | | 0.3510 | | | 6.2257 | | 0.0126 |

| **Odds Ratio Estimates** | | | |
| --- | --- | --- | --- |
| **Effect** | **Point Estimate** | **95% Wald Confidence Limits** | |
| **OSA** | 0.416 | 0.209 | 0.829 |

| **Association of Predicted Probabilities and Observed Responses** | | | |
| --- | --- | --- | --- |
| **Percent Concordant** | 19.9 | **Somers' D** | 0.116 |
| **Percent Discordant** | 8.3 | **Gamma** | 0.412 |
| **Percent Tied** | 71.8 | **Tau-a** | 0.047 |
| **Pairs** | 27707 | **c** | 0.558 |





**Conclusion:**

Since there is only 1 predictor in the final model, the hosmer and lemeshow goodness-of-fit test cannot be shown in our output.

The p-value is smaller than 0.05, indicating that OSA is a significant predictor of losing weight over 5% at 5% significant level.

The 95% confidence interval for OSA is (0.209, 0.829).

For the patients who had Obstructive sleep apnea diagnosis at initial visit has 0.416 times the odds of a positive test as compared to the patients who had not.

The area under ROC curve is 0.5581 in the final model, which indicating that this model is not a good fit. Also, the final model has a poorer fit than the full model.

OSA has a negative effect on losing weight.

# 5. Figures

**The trend of BMI from initial visit to final visit**



| **Variable** | **Label** | **Mean** | **Median** |
| --- | --- | --- | --- |
| BMI1 BMI2 BMI3 BMI4 BMI5 BMI6 BMI7 BMI8 BMI9 BMI10 BMI11 BMI12 BMI13 BMI14 | Patient’s BMI at initial visit Patient’s BMI at visit 2 Patient’s BMI at visit 3 Patient’s BMI at visit 4 Patient’s BMI at visit 5 Patient’s BMI at visit 6 Patient’s BMI at visit 7 Patient’s BMI at visit 8 Patient’s BMI at visit 9 Patient’s BMI at visit 10 Patient’s BMI at visit 11 Patient’s BMI at visit 12 Patient’s BMI at visit 13 Patient’s BMI at visit 14 | 35.4941839 34.7641709 34.0032138 34.2414583 32.9110489 33.1502538 32.3980636 32.2351728 32.1427492 32.0849206 31.9677441 32.0134045 32.1138245 32.0816630 | 34.0497432 33.2368878 32.5136875 31.6205213 31.5773435 31.1480626 30.8938248 30.6255831 30.5371393 30.2734375 30.0365898 30.2274862 30.3669380 30.5093386 |

**The frequency of the top 5 more common comorbidities in order of most common to least common**



| **Obs** | **como** | **COUNT** | **PERCENT** |
| --- | --- | --- | --- |
| **1** | Prediabetes | 128 | 16.7539 |
| **2** | Psych | 126 | 16.4921 |
| **3** | HTN | 120 | 15.7068 |
| **4** | HLD | 116 | 15.1832 |
| **5** | Hypothyroidism | 76 | 9.9476 |