Obesity Exploration

Kushal Kumar Dey

Friday, December 05, 2014

In this report, we perform an exploratory analysis of Obesity and Abdominal Obesity with respect totwo genetic factors (namely the DGAT and the MC4R genes). Note that in the samples availabel, there were two levels of the MC4R gene (G/G and G/A) and the DGAT gene had three levels (TT, CT, CC).

## DGAT  
## Obesity\_index CC TC TT  
## 0 108 80 70  
## 1 7 41 66

|  |  |  |
| --- | --- | --- |
| Levels | Mean BMI | SD BMI |
| TT | 28.2027205882353 | 4.2854118548624 |
| TC | 26.5865289256198 | 5.53656507694434 |
| CC | 24.964347826087 | 3.68932198649847 |

The values of the pairwise comparison t-test between CC/TC, TT/TC and CC/TT are given by 0.0083983, 0.0101261 and 0.0101261. Also we perform a linear regression of BMI on the genetic effect of DGAT

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 24.9643 | 0.4262 | 58.5698 | 0.0000 |
| DGATTC | 1.6222 | 0.5953 | 2.7251 | 0.0067 |
| DGATTT | 3.2384 | 0.5790 | 5.5926 | 0.0000 |

A GLM model of Obesity on DGAT and the output is shown as below

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 24.9643 | 0.4262 | 58.5698 | 0.0000 |
| DGATTC | 1.6222 | 0.5953 | 2.7251 | 0.0067 |
| DGATTT | 3.2384 | 0.5790 | 5.5926 | 0.0000 |

Now we consider the MC4R gene with only two levels given in the sample represented by G/G and G/A. and we compute the summary statistics based on that

## MC4R  
## Obesity\_index G/A G/G  
## 0 181 77  
## 1 20 94

|  |  |  |
| --- | --- | --- |
| Levels | Mean BMI | SD BMI |
| G/G | 29.6926315789474 | 3.26087223606477 |
| G/A | 24.1094527363184 | 4.28744226771685 |

We use a linear model fit of BMI against the different levels of the MC4R gene and fit obesity against the MC4R gene. We try to see how the genetic factor influences the BMI. The odds atio for the table is given by 11.05, which is pretty large and shows a great deal of dependence of obesity on the MC4R levels. We first present a linear model fitting BMI against the MC4R gene

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 24.1095 | 0.2716 | 88.7838 | 0 |
| MC4RG/G | 5.5832 | 0.4005 | 13.9397 | 0 |

We also fit a generalized linear model fitting obesity against the different levels of the MC4R gene.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 0.0995 | 0.0285 | 3.4933 | 5e-04 |
| MC4RG/G | 0.4502 | 0.0420 | 10.7162 | 0e+00 |

Now we focus on the abdominal obesity, due to WHR and look at how this WHR depends on the levels of the DGAT and the MC4R genes.

## MC4R  
## WHO\_Ab\_obese\_indicator G/A G/G  
## 0 87 17  
## 1 114 154

|  |  |  |
| --- | --- | --- |
| Levels | Mean BMI | SD BMI |
| G/G | 0.923391812865497 | 0.0836390265133804 |
| G/A | 0.862537313432836 | 0.0770456348584803 |

The odds ratio of the Abdominal Obesity against MC4R is 6.924. This again suggests that MC4R is a significant genetic factor for abdominal obesity. We first fit a linear model of the Waist to Hip Ratio (WHR) against the levels of the MC4R gene.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 0.8625 | 0.0057 | 152.5856 | 0 |
| MC4RG/G | 0.0609 | 0.0083 | 7.2989 | 0 |

Next we consider fitting a enerlaized linear model with abdominal obesity as the response and the levels of the MC4R gene as the predictor.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 0.5672 | 0.0295 | 19.2359 | 0 |
| MC4RG/G | 0.3334 | 0.0435 | 7.6670 | 0 |

We observe how the abdominal obesity varies against the genotypic levels of the DGAT gene namely CC, TC and TT.

|  |  |  |
| --- | --- | --- |
| Levels | Mean BMI | SD BMI |
| TT | 0.903970588235294 | 0.0809963822856004 |
| TC | 0.883388429752066 | 0.0950662376013716 |
| CC | 0.882086956521739 | 0.0788902819285193 |

Next we present a linear model depicting the variation of the Waist to Hip Ratio (WHR) depending on the different levels of the DGAT gene

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 0.8821 | 0.0079 | 111.0026 | 0.0000 |
| DGATTC | 0.0013 | 0.0111 | 0.1173 | 0.9067 |
| DGATTT | 0.0219 | 0.0108 | 2.0271 | 0.0434 |

Also of interest is how the indicator of obesity changes with th different levels of the DGAT gene (A logistic linear model is fitted).

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | t value | Pr(>|t|) |
| (Intercept) | 0.7043 | 0.0410 | 17.1775 | 0.0000 |
| DGATTC | -0.1010 | 0.0573 | -1.7645 | 0.0785 |
| DGATTT | 0.1339 | 0.0557 | 2.4035 | 0.0167 |

Now we present the logistic regression of the Obesity, Overweight data and the Abdominal Obesity data against the genetic factors DGAT gene, MC4R gene, Age and Gender and we see, how these different factors contribute to explaining the probability that a person is obese/ overweight/ abdominally obese. First we look at the abdominally obese individuals.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | z value | Pr(>|z|) |
| (Intercept) | -4.7556661 | 0.7880079 | -6.035048 | 0.0000000 |
| DGATTC | -1.0584622 | 0.3771941 | -2.806147 | 0.0050138 |
| DGATTT | 0.5098612 | 0.3938686 | 1.294496 | 0.1954943 |
| MC4RG/G | 2.2491777 | 0.3634724 | 6.188028 | 0.0000000 |
| gender | 3.1523227 | 0.4047759 | 7.787822 | 0.0000000 |
| Age | 0.0305290 | 0.0189546 | 1.610638 | 0.1072587 |

The AIC for this model is given by 279.21. We consider now the overweight data (overweight implies BMI>25). We consider an indicator variable that is 1 if a person is overweight (BMI>25) and 0 otherwise.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | z value | Pr(>|z|) |
| (Intercept) | -2.2132171 | 0.6537582 | -3.385376 | 0.0007108 |
| DGATTC | -0.5213699 | 0.3273194 | -1.592847 | 0.1111944 |
| DGATTT | 0.6736472 | 0.3337316 | 2.018530 | 0.0435361 |
| MC4RG/G | 2.9237494 | 0.3508340 | 8.333712 | 0.0000000 |
| gender | 0.6854998 | 0.2764627 | 2.479538 | 0.0131553 |
| Age | 0.0243992 | 0.0170006 | 1.435195 | 0.1512316 |

The AIC of the model is 346.66. We consider now the overweight data (obese implies BMI>30). We consider an indicator variable that is 1 if a person is overweight (BMI>30) and 0 otherwise.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Estimate | Std. Error | z value | Pr(>|z|) |
| (Intercept) | -7.5378820 | 0.9718011 | -7.756610 | 0.0000000 |
| DGATTC | 2.1871560 | 0.4805071 | 4.551767 | 0.0000053 |
| DGATTT | 2.7041768 | 0.4740022 | 5.704988 | 0.0000000 |
| MC4RG/G | 2.1737559 | 0.3111868 | 6.985372 | 0.0000000 |
| gender | 0.5601825 | 0.2968925 | 1.886819 | 0.0591846 |
| Age | 0.0760152 | 0.0200353 | 3.794061 | 0.0001482 |

The AIC of the model is given by 305.81. Note from the above analysis that we did, it is very clear that the MC4R gene has a very substantial effect on both abdominal and regular obesity, nd even for overweight data. The DGAT gene is significant too for obesity, the TC genotype is significant for the abdominal obesity which is slightly strange, however TT and CC levels are not found to be significant for abdominal obesity. The DGAT gene fails to be significant fr the overweight data. Age is found to be a significant impacting factor for obesity but not for abdominal obesity and gender is found to be highly signifiant for abdominal obesity and only midly significant for the general obesity.