

Data Analysis Plan – Capstone Project

Hypothesis: Adults with self-reported food insecurity are at a higher risk for metabolic syndrome.

Secondary Hypothesis: Adults with self-reported food insecurity are more likely to have a normal BMI in conjunction with metabolic syndrome and less likely to be overweight/obese without metabolic syndrome.

Exposure:

Food Insecurity as measured by the Adult Food Security Category on a scale 1-4.

Outcomes:

Primary: Metabolic Syndrome – meets 3 or more criteria

Waist Circumference: ≥ 102 cm for men and ≥ 88 cm for women

or ≥ 94 cm in men and ≥ 80 cm in women

Blood Pressure: Elevated $\geq 130/85$ mm Hg or drug treatment

Triglycerides: ≥ 150 mg/dL or drug treatment

HDL-C: <40 mg/dL men and <50 mg/dL women or drug treatment

Fasting Glucose: ≥ 100 mg/dL or drug treatment

Secondary: BMI

Possible Covariates:

Gender (male, female), Ethnicity (Non-Hispanic Black, Hispanic, Non-Hispanic Other, Non-Hispanic White), age, education, income, alcohol use, smoking status, physical activity.

Inclusion Criteria:

18-65 years olds. Serum triglyceride levels and serum fasting glucose levels were measured on NHANES examinees who were examined in the morning session only. Eligible NHANES examinees were randomly assigned to the morning fasting sample. Only individuals in the morning session will therefore be included.

Exclusion Criteria:

Pregnant Women.

Data and Variables:

All datasets contain the SEQN variable that will be used to link individuals between datasets.

Food Insecurity: Adult Food Security Category from Food Security Questionnaire. Variable: 1999-2002: ADfdsec, 2003-2014: FSDAD.

Waist Circumference: BMIWAIST from Body Measures Data Set

Blood Pressure: 1999-2002: BPXDAR from Blood Pressure Dataset 2003-2014 Systolic Average: BPXSY1, BPXSY2, BPXSY3, BPXSY4 Diastolic Average: BPXDI1, BPXDI2, BPXDI3, BPXDI4

Hypertension medication: BPQ050A from Blood Pressure & Cholesterol

Triglycerides: LBXTR from Cholesterol- LDL & Triglycerides

HDL-C: Variable: 1999- 2004: LBDHDL from Cholesterol Total & HDL dataset 2005-2014: LBDHDD from Cholesterol-HDL dataset.

Cholesterol Medication: Answers yes to BPQ090D and BPQ100D from Blood Pressure and Cholesterol Questionnaire

Fasting Glucose: Variable: LBXGLU from Plasma Fasting Glucose, Serum C-peptide & Insulin Dataset (1999-2004), Plasma Fasting Glucose & Insulin Dataset (2005-2012) and Plasma Fasting Glucose Dataset (2013-2014).

Medication to lower blood sugar: DIQ070 Now taking pills to lower blood sugar from Diabetes questionnaire.

BMI: BMXBMI from Body Measures Dataset

Gender: RIAGENDR from the Demographics and Sample Weights Data Set

Ethnicity: 1999-2010: RIDRETH1 from the Demographics and Sample Weights Data Set 2011-2014: RIDRETH3.

Age: RIDAGEYR from the Demographics and Sample Weight Data Set

Alcohol Use: Never – will be defined as answering no to ALQ 110 - Had at least 12 alcohol drinks/lifetime Variable: ALQ110.

Moderate – Reports drinking 2 or less drinks per day and 0 for # days have 5 or more drinks/past 12 mos. Variables: ALQ120, ALQ130 and ALQ140

Heavy – Reports drinking more than 2 drinks per day or greater than 0 for # days have 5 or more drinks/past 12 mos. Variables: ALQ120, ALQ130 and ALQ140

Smoking status: Never – answered no to SMQ020 smoked at least 100 cigarettes in life.

Former- answered yes to SMQ020 SMQ020 smoked at least 100 cigarettes in life and Not at all to SMQ040 Do you now smoke cigarettes.

Current – answered yes to SMQ020 smoked at least 100 cigarettes in life and Every Day or Some Days to SMQ040 Do you now smoke cigarettes

Physical Activity: Answered yes to any of the moderate activity questions (PAD320, PAQ620, PAQ665) using Physical Activity Questionnaire

Income: Variable: INDFMINC (Annual Family Income) (before 2007) INDFMIN2(after 2007) from Demographics and Sample Weight Data Set

Pregnancy: Variable: RIDEXPRG from Demographics and Sample Weight Data Set

Education: DMDEDUC3 from 18-19 years olds, DMDEDUC2 for ≥ 20 year olds.

Analysis:

Primary analysis will be performed using logistic regression. A likelihood ratio test will be used to test the significance of the food insecurity exposure variable with full food security used as the reference category. Odds ratios for metabolic syndrome in the marginal, low, and very low food security categories compared to the reference full food security category will be calculated. Covariates that are not confounders or effect modifiers will be considered as precision variables or for face validity.

Sampling weights will be calculated according to the following formula: if sddsrvy in (1,2) then SAF16YR = $1/4 * WTS\text{AF}4\text{YR}$, if sddsrvy in (3,4,5,6,7,8) then SAF16YR = $1/8 * WTS\text{AF}2\text{YR}$, where sddsrvy is the survey year variable (1 = 1999-2000, 2 = 2001-2002, 3 = 2003-2004, 4 = 2005-2006, 5 = 2007-2008, 6 = 2009-2010, 7 = 2011-2012, 8 = 2012-2014). The 'survey' package in R will be used to account for the complex survey design. The variable SDMVPSU (primary sampling unit) will be used as the 'id', the variable SDMVSTRA (sampling strata) will be used as the 'strata' and the variable SAF16YR (sampling weights) will be used as the 'weights' in the svydesign function in the survey package. The design object created by this function will be passed to svyglm for the logistic regression. All analysis will be performed using R and version controlled using Git. All files except data will be stored in a repository on github.

Tables and Graphs:

Table 1. Characteristics of Study Participants by Food Security Category

	Full Food Securitivity	Marginal Food Security	Low Food Security	Very Low Food Security	p-value
	N(%)	N(%)	N(%)	N(%)	
Gender					
Male					
Ethnicity					
Asian					
Black					
Hispanic					
Other					
Non-Hispanic White					
Education					
Less than 9 th Grade					
Some High School					
High School Grad					
Some College or AA degree					
College Graduate or above					
Smoking Status					
Never					
Former					
Current					
Alcohol					
Never					
Occasional					
Moderate					
Heavy					
Physical Activity					
	Mean(SD)	Mean(SD)	Mean(SD)	Mean(SD)	
Age					
Income					

Table 2. Crude and adjusted odds ratios for metabolic syndrome vs Full Food Security (reference)

	Crude OR (95% CI)	Adjusted OR (95% CI)
Marginal Food Security		
Low Food Security		
Very Low Food Security		