# The Role of Sleep in Mediating the Relationship Between BMI and CVD Risk Factors in Samoan Adults

by

# **Molly McKune**

BS in Mathematics, Kennesaw State University, 2017

Submitted to the Graduate Faculty of the

Graduate School of Public Health in partial fulfillment

of the requirements for the degree of

Master of Science in Biostatistics

University of Pittsburgh

2021

#### UNIVERSITY OF PITTSBURGH

#### GRADUATE SCHOOL OF PUBLIC HEALTH

This thesis was presented

by

#### **Molly McKune**

It was defended on

April 26, 2021

and approved by

Jeanine M. Buchanich, PhD, Research Associate Professor, Biostatistics Graduate School of Public Health, University of Pittsburgh

Ada Youk, PhD, Research Associate Professor, Biostatistics Graduate School of Public Health, University of Pittsburgh

Tina Costacou, PhD, Associate Professor, Epidemiology Graduate School of Public Health, University of Pittsburgh

**Thesis Advisor:** Jenna C. Carlson, PhD, Assistant Professor, Biostatistics Graduate School of Public Health, University of Pittsburgh

Copyright © by Molly McKune

2021

#### The Role of Sleep in Mediating the Relationship Between BMI and CVD Risk Factors in Samoan Adults

Molly McKune, MS

University of Pittsburgh, 2021

#### Abstract

<u>Background:</u> Previous research shows an association between higher BMI and higher cardiovascular disease (CVD) risk, as well as an association between poor sleep quality and higher CVD risk, but minimal research has been conducted in this area with Samoan adults. Samoan adults have a higher BMI than average, partially due to a variant found in their *CREBRF* (*CREB 3* regulatory factor) gene. Due to genetic tendency to have higher BMI, it is imperative to identify factors that can mitigate the impact of high BMI on CVD risk, such as sleep quality.

<u>Objective:</u> From a cohort selected from the GWAS which identified the *CREBRF* genetic variant, sleep quality and CVD risk factor data was collected. This cohort is known as The Soifua Manuia (Good Health) Study. The objective of this paper is to identify sleep quality measures which mediate the relationship between BMI and CVD risk using these participants.

<u>Methods:</u> Path Analysis, a form of structural equation modeling, was performed on 503 participants to quantify the relationship between BMI, Sleep Quality, and CVD risk, and identify statistically significant evidence of a causal pathway where sleep quality mediates the relationship between BMI and CVD risk.

<u>Results:</u> Three path analyses were found to be significant. Sleep disordered breathing (AHI 3%) was found to mediate the relationship between BMI and both HDL cholesterol levels and diastolic blood pressure (p < 0.05). Snoring volume was found to mediate the relationship between

BMI and HDL cholesterol (p < 0.2). In all significant models, an improvement in the sleep quality measure was associated with an improvement in CVD risk factor on average.

<u>Conclusions:</u> The data preliminarily shows that improvements in sleep disordered breathing and snoring volume could help mitigate the impact of BMI on HDL cholesterol and diastolic blood pressure. These data present evidence that many Samoan adults have some form of sleep disordered breathing and undiagnosed sleep apnea. The public health significance of these results is to have identified compelling evidence of the need for a public health intervention in Samoan adults to diagnose and treat sleep apnea to improve sleep quality.

# **Table of Contents**

Prefaceix
1.0 Introduction1
1.1 BMI, Sleep Quality, and CVD Risk Factors1
1.2 Samoan Population2
1.3 Path Analysis3
2.0 Methods
2.1 Data
2.1.1 Data Source4
2.1.2 Covariates5
2.1.3 Data Cleaning6
2.2 Analysis7
3.0 Results 11
3.1 Descriptive Statistics 11
3.2 Linear Models to Assess Confounder Relationships 20
3.3 Structural Equation Modeling: Path Analysis22
4.0 Discussion
Appendix A Code
Bibliography

# List of Tables

Table 1: Summary Statistics of Sex & CREBRF Genotype	11
Table 2: Summary Statistics of Age	
Table 3: Summary Statistics of BMI	
Table 4: Outcome Summary Statistics	
Table 5: Mediator Summary Statistics	
Table 6: Sleep Disordered Breathing Distribution	
Table 7: Linear Models Summary	
Table 8: Significant Path Analysis Model Summaries	
Table 9: Path Analysis Model Fit Statistics	

# List of Figures

Figure 1: Missing Data	7
Figure 2: Path Diagram	
Figure 3: Distribution of Age	
Figure 4: Distribution of BMI	
Figure 5: HDL Distribution	
Figure 6: LDL Distribution	
Figure 7: Fasting Glucose Distribution	
Figure 8: Systolic Blood Pressure Distribution	
Figure 9: Diastolic Blood Pressure Distribution	
Figure 10: AHI 3% Distribution	
Figure 11: ESS Distribution	
Figure 12: Average Snoring Volume Distribution	
Figure 13: Weighted Sleep Duration Distribution	
Figure 14: Diagram of Associations Between Confounders and Sleep Q	uality Measures 22
Figure 15: Path Diagram AHI 3% & DBP (Model 1)	
Figure 16: Path Diagram AHI 3% & HDL (Model 2)	
Figure 17: Path Diagram Snore Volume & HDL (Model 3)	

#### Preface

We would like to thank the Samoan participants and local village authorities and research assistants over the years. We acknowledge the support of our research collaboration with the Samoa Ministry of Health; the Samoa Bureau of Statistics; the Samoa Ministry of Women, Community and Social Development; and the American Samoa Department of Health. The Samoan Obesity, Lifestyle, and Genetics Adaptations (OLaGA) Study Group investigators are Ranjan Deka, Jenna C. Carlson, Kima Fa'asalele-Savusa, Nicola L. Hawley, Vaimoana Lupematisila, Stephen T. McGarvey, Ryan L. Minster, Leausa Toleafoa Take Naseri, Muagututi'a Sefuiva Reupena, Melania Selu, John Tuitele, Asiata Satupa'itea Viali and Daniel E. Weeks.

This thesis is dedicated to my incredible and wide-reaching support system: Mom, Dad, Daniel, Nino, and my amazing Biostat professors & cohort.

I'd like to also make a special dedication to my grandmother, Janet, whose spirit and love provide me strength and faith when I need it most.

#### **1.0 Introduction**

#### 1.1 BMI, Sleep Quality, and CVD Risk Factors

It is well-known in research that BMI has an impact on cardiovascular disease (CVD) risk. This impact can be seen by a higher BMI leading to poorer outcomes for CVD risk factors, such as diabetes management, cholesterol levels, and high blood pressure (McAnulty, Scragg). For Samoans, higher BMI is associated with higher blood pressure and lower HDL cholesterol measurements. In recent years, it has also been shown that sleep quality measures, such as Epworth Sleepiness Score and sleep disordered breathing, can impact CVD risk, with poorer sleep quality leading to increased risk of disease (Fernandez-Mendoza). Given the continuing prevalence of CVD, it is important to identify as many ways to lower risk for disease as possible.

Given that both sleep quality and BMI impact CVD risk, it is hypothesized that sleep quality could mediate the relationship between BMI and common CVD risk factors. If BMI's impact on CVD risk is mediated by sleep quality, sleep improvement could become an area to focus interventions to help lower risk factors of CVD.

In current state, lowering BMI is a common suggestion to reduce risk of CVD. However, reducing BMI via weight loss is often difficult, and some recently identified genetic factors can make it even more difficult to lower BMI. If sleep quality improvements could minimize the impact of BMI on CVD risk and be easier to implement than weight loss, it could be a meaningful option to explore for public health impacts of lowering CVD risk across the globe.

1

#### **1.2 Samoan Population**

The population of Samoa could see a significant benefit from identifying alternatives to lowering CVD risk other than reducing BMI. As a population, Samoans tend to have a higher BMI than other regions of the world. Although their body mass is comprised of higher muscle mass compared to other regions, Samoans are still at heightened risk for obesity (Hawley 2014). Previous research has identified a variant (rs373863828) in the *CREBRF* gene prevalent in Samoans which is associated with an increased BMI (Hawley 2020). Although this same variant is also associated with a lower risk of diabetes, it has not been shown to be associated with a lower risk for other CVD risk factors (Wang 2019). Given the history of association between sleep and genetics, it could also be possible for this *CREBRF* gene to influence sleep quality.

The genome-wide association study (GWAS) that identified the variant in *CREBRF* also selected a cohort of individuals to follow up and collect further information, including a comprehensive sleep quality data collection via questionnaires and WatchPat technology, a wearable device that records accurate measurements of sleep characteristics. Via this follow-up, it was identified that a majority of the 500+ participants selected were found to have at least mild sleep disordered breathing and/or sleep apnea syndrome, yet only one participant had been diagnosed with sleep apnea by a physician. Given that a physician diagnosis of sleep apnea is typically needed to receive treatment, many Samoans could be suffering from low quality sleep due to undiagnosed and untreated sleep apnea.

Beyond the simple sleepiness ramifications of poor sleep, it has also been shown that sleep apnea is associated with higher risk for CVD (Floras 2018). This furthers the public health importance in Samoa to determine assess if sleep quality improvement can mediate the relationship between BMI and CVD risk factors. If such a relationship were to exist, a governmental public health effort to diagnose and treat sleep apnea correctly across the Samoan population could result in a large impact on improved CVD risk outcomes for the country.

#### **1.3 Path Analysis**

Structural equation modelling, specifically path analysis was used to assess whether sleep quality mediates the relationship between BMI and CVD risk. Path Analysis is a method of defining specific equations that incorporate mediating variables and modeling them to find evidence of causation in relationships (Pearl, 2012) (Nayebi, 2020).

The ability to model causation in the form of a directional relationship is the main reason for choosing path analysis over linear regression. In addition, mediation cannot be specifically modeled by linear regression but ratheronly moderation (i.e., interactions) can be modeled.

Using path analysis, the intent is to identify statistically significant models that define a causal pathway between BMI, sleep quality, and CVD risk factors. Through this, there would be evidence to suggest effects of BMI on CVD risk factors can be reduced by improving sleep quality, including treatment of sleep apnea.

3

#### 2.0 Methods

#### 2.1 Data

#### 2.1.1 Data Source

Data used in this analysis comes from The Soifua Manuia (Good Health) study in Samoa. This study followed up with a cohort of 519 individuals between August 2017 and March 2019 from its parent genome-wide association study (GWAS), which took place in 2010. The GWAS identified a gene of interest, denoted as CREBRF, and participants were selected for The Soifua Manuia study based on their genotype at rs373863828, a variant in *CREBRF*, aiming for a ratio of 1 AA: 2 AG: 2 GG. The Soifua Manuia study collected a wide range of information during a 7 to 10-day period. The information collected consisted of "anthropometric measurements (weight, height, circumferences, and skinfolds), body composition assessment (bioelectrical impedance and dual-energy x-ray absorptiometry), point-of-care glycated hemoglobin measurement, a fasting blood draw and oral glucose tolerance test, urine collection, blood pressure measurement, hand grip strength measurement, objective physical activity and sleep apnea monitoring, and questionnaire measures (eg, health interview, cigarette and alcohol use, food frequency questionnaire, socioeconomic position, stress, social support, food and water insecurity, sleep, body image, and dietary preferences)" (Hawley et al., 2020). The sleep apnea monitoring was done via a WatchPat device, which accurately collects various health measurements while sleeping, such as oxygen level and snore volume.

#### 2.1.2 Covariates

Covariates selected from The Soifua Manuia study to be used in this analysis were: age, sex, CREBRF genotype, BMI, Systolic Blood Pressure, Diastolic Blood Pressure, High-density Lipoprotein, Low-density Lipoprotein, Total Cholesterol, Triglycerides, Fasting Glucose, Epworth Sleepiness Score, Apnea-Hypopnea Index with 3% Oxygen Desaturation Cutoff, Average Snoring Volume, and Weighted Sleep Duration.

The cardiovascular disease risk factor outcomes used in this paper from The Soifua Manuia study are Systolic Blood Pressure, Diastolic Blood Pressure, High-density lipoprotein (HDL) cholesterol, Low-density lipoprotein (LDL) cholesterol, and Fasting Glucose. HDL, LDL, and Fasting Glucose were all lab-collected from a fasting blood serum sample. The units of measure for all three are milligrams per deciliter (mg/dL). Systolic and diastolic blood pressure were also collected in a lab setting and are an average of multiple measurements. The units of measure for both blood pressure variables are millimeters of mercury (mmHg).

Sleep quality mediators used are Epworth Sleepiness Score (ESS), Apnea-hypopnea Index with 3% Oxygen Desaturation Cutoff (AHI 3%), Average Snoring Volume, and Weighted Sleep Duration. The WatchPat technology worn during sleep measured AHI 3%, Average Snoring Volume, and Weighted Sleep Duration. Epworth Sleepiness Score is a subjective measurement of sleepiness and self-reported via the questionnaire. Average Snoring Volume is an average of snoring volume over the entire time sleeping with a WatchPat and is measured in decibels. Weighted Sleep Duration is a weighted average of weekday sleep and weekend sleep and is measured in minutes. The Epworth Sleepiness Score measures daytime sleepiness level and ranges from 0 to 24, with a higher number indicating higher average daytime sleepiness (Johns, 1991). AHI 3% is a continuous variable that can be used to categorize sleep disordered breathing. Higher AHI 3% measurements correspond with worse sleep disordered breathing categorization.

BMI was the main predictor of interest. It was lab measured and reported in kg/m<sup>2</sup>. In addition, there were three variables included in the analysis as confounders: sex, age, and CREBRF genotype. Sex was measured by the count of X chromosomes in the genotype: XX = Female, XY = Male; it was coded as 0=Male and 1=Female. Age was measured as a continuous number and includes decimals. CREBRF genotype was measured during the initial GWAS study and shows the number of A alleles: 0=GG, 1=AG, or 2=AA.

#### 2.1.3 Data Cleaning

The data were found in three different Soifua Manuia datasets and merged by participant ID. Once merged and subset to only the variables of interest, missingness was recoded. The original data used a value of -7777 to signify verified missing data. This was changed to be coded as "NA" in R, which indicated a missing data value. In addition, CREBRF was recoded due to a data management mistake. The corrected CREBRF variable had 3 levels: 0=GG, 1=AG, and 2=AA. Fasting glucose was also cleaned to account for participants taking diabetes medication. Anyone taking medication was removed from the Fasting Glucose observations and marked as missing ("NA").

After cleaning the data in steps outlined above, patterns of missingness across the entire dataset were assessed. See Figure 1 below—blue indicates valid values and red indicates missing data. The two variables with the most amount of missing data were Average Snoring Volume (n-miss=177) and AHI 3% (n-miss=175). No imputation methods were used, and any missing data

was left out of analysis to ensure no false relationships were found during structural equation modeling. The total number of participants used for analysis was 503.



Figure 1: Missing Data

#### 2.2 Analysis

It is well-known in healthcare that BMI impacts cardiovascular disease risk factors, such as HDL, LDL, Fasting Glucose, and blood pressure. In recent years, it has also been shown that sleep quality can impact cardiovascular disease risk. However, it has not yet been shown that sleep quality could mediate the relationship between BMI and CVD risk factors. The use of path analysis as a form of structural equation modeling was performed to determine the presence of a mediating effect between BMI and CVD risk factors. First, simple linear models were fit between the confounders and the mediators to assess if any pathways existed between them. A key assumption of path analysis is that each causal relationship only goes in one direction and there are no reciprocating effects, where the relationship is multi-directional (Nayebi, 2020). See Figure 2 for the proposed framework of path analysis. Black lines represent relationships confirmed by literature and research. Green lines represent relationships that were drawn only if linear models between the two were statistically significant and of meaningful magnitude.



**Figure 2: Path Diagram** 

#### \*HDL, LDL, Triglycerides, Total Chol, Fasting Gluc, SBP, DBP; +ESS, AHI 3%, Snore Vol, Sleep Duration

These possible causal pathways were input into the model using the following structure of multiple linear regression models:

CVD Risk Factor ~ b*Sleep Measure + c*BMI + d*CREBRF + e*Sex + f*Age	(2.2.1)
Sleep Measure ~ $a*BMI + g*Sex + h*CREBRF+l*Age$	(2.2.2)
$BMI \sim i^*Age + j^*Sex + k^*CREBRF$	(2.2.3)
BMI Indirect = $a^*b$	(2.2.4)
BMI Direct $= c$	(2.2.5)

**Equation 1: Path Analysis Structural Equation Models** 

BMI Total 
$$= c + (a^*b)$$
 (2.2.6)

Where a, b, ..., l in equations 2.2.1 - 2.2.3 are the weights of interest for each specified path in Figure 2. The equation with an outcome of sleep measure (2.2.2) only included g, h, and l if linear models showed a relationship between the confounders and the sleep measure being assessed in the model.

In total, 20 path analysis models were performed to test all combinations of CVD risk factors and sleep quality measures. For each model of path analysis, the indirect, direct, and total effects of BMI on the CVD risk factor were assessed for statistical significance using p-values and magnitude of effect (equations 2.2.4 - 2.2.6). For a model to be considered statistically significant, the indirect effect must have had a p-value less than 0.2 and a total effect less than 0.05. The threshold for the indirect effect only required borderline significance in an effort to identify all possible meaningful indirect effects. This is because the indirect effect of BMI is of most interest in this analysis and was used to quantify how sleep quality mediates the relationship between BMI and CVD risk factors. In contrast, the total effect of BMI required the standard level of significance in order to ensure BMI does play a role in impacting the CVD risk factor in this sample of Samoan adults.

After selecting the models of interest, they were assessed for fit using a global test statistic to test overall model significance and comparative fit index (CFI), which compares the model being tested to a null model and ranges from 0 to 1 (Glen 2017). In general, the CFI should be higher than 0.9 and the test statistic should be less than 0.05. If multiple models predicting an outcome are observed to be significant, those models were also compared using AIC and BIC.

With path analysis a causal model can be specified through a defined set of multiple linear regression models. The goal of performing this analysis is to interpret an indirect causal path between BMI and CVD risk factors through sleep quality. In addition, the total effect of BMI on CVD risk should also be considered in interpretation. Individual standardized path values between all identified pathways were also be shown through diagrams and were used to suggest further analysis steps.

#### **3.0 Results**

# **3.1 Descriptive Statistics**

The count and distribution of sex and genotype of *CREBRF* variant rs373863828 can be seen in Table 1. Of the 503 participants used in analysis, 46% were male and 54% were female. There were 42.5% who had no copies of the A allele (GG), 39.6% who had one copy of the A allele (AG), and 17.9% who had two copies of the A allele (AA).

	Male	Female	Total
GG	99 (19.7%)	115 (22.9%)	214 (42.5%)
AG	91 (18.1%)	108 (21.5%)	199 <i>(39.6%)</i>
AA	41 (8.2%)	49 (9.7%)	90 <i>(17.9%)</i>
Total	231 (45.9%)	272 (54.1%)	503

Table 1: Summary Statistics of Sex & CREBRF Genotype

The distribution of age can be seen in Figure 3 and Table 2. Ages of the sample ranged from 30.68 to 72.7 years. The average age of the sample was 52.14 years. The distribution of age is acceptably symmetric in the sample.

#### Table 2: Summary Statistics of Age

Min	Q1	Median	Mean	Q3	Max
30.68	44.12	51.85	52.14	59.88	72.70

#### **Distribution of Age**



Figure 3: Distribution of Age

The distribution of BMI can be seen in Table 3 and Figure 4 below. The average BMI of the sample is 35.7, with the BMI of the sample ranging from 20.2 to 76.69. As seen in the figure below, the distribution of BMI is skewed right. The large values present were confirmed to be correctly measured.

#### Table 3: Summary Statistics of BMI

Min	Q1	Median	Mean	Q3	Max	N Missing
20.20	30.76	35.31	35.74	40.01	76.69	2

# **Distribution of BMI**



**Figure 4: Distribution of BMI** 

The summary statistics of the outcomes HDL, LDL, Fasting Glucose, Systolic BP, and Diastolic BP are listed in Table 4 below. The distributions of these variables are shown in Figures 5-9, respectively.

Table 4:	Outcome	Summary	Statistics
----------	---------	---------	------------

	Min	Q1	Median	Mean	Q3	Max	N Missing
HDL (mg/dL)	21.00	43.00	48.00	49.04	54.00	109.00	5
LDL (mg/dL)	25.0	105.5	128.0	127.9	149.5	285.0	12
Triglycerides (mg/dL)	43.0	86.25	115.0	131.94	155.75	643.0	5
Total Cholesterol (mg/dL)	99.0	176.0	202.0	202.6	227.8	395.0	5
Fasting Glucose (mg/dL)	54.50	94.3	103.60	118.8	121.1	476.6	51
Systolic BP (mmHg)	87.0	115.5	126.5	129.9	140.0	221.5	12



Figure 5: HDL Distribution



LDL Cholesterol Distn

Figure 6: LDL Distribution

Both HDL and LDL cholesterol measures are skewed slightly right.



# **Fasting Glucose**

Figure 7: Fasting Glucose Distribution

Fasting glucose is heavily skewed right. The glucose measures were confirmed to be valid for extreme measures. The average fasting glucose is 129.12 mg/dL.



Figure 8: Systolic Blood Pressure Distribution





Figure 9: Diastolic Blood Pressure Distribution

Both SBP and DBP are slightly skewed right, but close to symmetric. The average SBP measurement is 129.9, and the average DBP measurement is 79.97.

The summary statistics for the mediators Apnea-Hypopnea Index with 3% Oxygen Desaturation (AHI 3%), Epworth Sleepiness Score (ESS), Average Snore Volume, and Weighted Sleep Duration are listed in Table 5 below. The distributions for the mediators are shown in Figures 10-13, respectively.

	Min	Q1	Median	Mean	Q3	Max	N Missing
AHI 3%	0.300	8.375	16.550	19.808	26.300	76.900	175
ESS	0	4	7	6.853	10	16	34
Avg Snoring Vol	40	41	42	44.31	46	65	177
Sleep Duration	120	360	420	422.3	480	745.7	14

#### **Table 5: Mediator Summary Statistics**



Sleep Disordered Breathing Distribution

Figure 10: AHI 3% Distribution

AHI 3% measurements are skewed right toward increasing sleep disordered breathing severity. The average AHI 3% measurement is 19.808, which coincides with moderate sleep apnea. This is confirmed in the table below:

<b>Breathing Category</b>	Count (Column %)
Normal Breathing	40 (12.2%)
Mild SDB	108 (32.9%)
Moderate SDB	113 (34.5%)
Severe SDB	67 (20.4%)

# Table 6: Sleep Disordered Breathing Distribution

# **Epworth Sleepiness Score Distribution**



Figure 11: ESS Distribution

Values for ESS are distributed somewhat uniformly. The average ESS value for the sample is 6.853, which coincides with Higher Normal Daytime Sleepiness (Johns, 1991).

# Avg Snoring Volume



Figure 12: Average Snoring Volume Distribution

The average snoring volume distribution is skewed extremely positively toward louder volumes of snoring. The average volume of the sample is 44.31 decibels, and the maximum is 65.

# **Avg Sleep Duration**



**Figure 13: Weighted Sleep Duration Distribution** 

Weighted sleep duration is symmetrically distributed in the sample. The average is 422.3 minutes, and the range of sleep length is 120 to 745.7 minutes.

# 3.2 Linear Models to Assess Confounder Relationships

To determine if relationships exist between the potential confounders and the mediators, simple linear models were performed for the following comparisons:

Equation 2: Simple Linear Regression Models for Counfounders and Sleep Quality

AHI 3% = 
$$\beta_0 + \beta_1 * X$$
 (3.2.1)

$$ESS = \beta_0 + \beta_1 * X \tag{3.2.2}$$

Snoring Volume = 
$$\beta_0 + \beta_1 * X$$
 (3.2.3)

Sleep Duration = 
$$\beta_0 + \beta_1 * X$$
, (3.2.4)

Where X was sex, age, or CREBRF genotype. The results are summarized in Table 7 below.

	$oldsymbol{eta}_1$ Estimate	P-value
АНІ 3% ~ Sex	-4.453	0.0074*
AHI 3% ~ Age	0.21575	0.0102*
AHI 3% ~ CREBRF	2.8602	0.0014*
ESS ~ Sex	-0.9013	0.0175*
ESS ~ Age	0.001742	0.9260
ESS ~ CREBRF	0.4950	0.0179*
Snore Vol ~ Sex	0.3063	0.5620
Snore Vol ~ Age	0.009063	0.7350
Snore Vol ~ CREBRF	0.6499	0.0215*
Sleep Dur ~ Sex	10.865	0.2060
Sleep Dur ~ Age	0.1439	0.7360
Sleep Dur ~ CREBRF	1.798	0.7040

#### **Table 7: Linear Models Summary**

\* denotes a statistically significant result (p <0.05)

Statistically significant results (p < 0.05) were observed between AHI and Sex, AHI and age, AHI and CREBRF, ESS and Sex, ESS and CREBRF, and Snore Volume and CREBRF. Even though the results were statistically significant, the relationships between AHI and Age as well as Snore Volume and CREBRF were excluded from path analysis due to the small magnitude of the effects.

The final relationships in the structural equation modeling path analysis between confounders and mediators are shown by orange arrows as follows:



Figure 14: Diagram of Associations Between Confounders and Sleep Quality Measures \*HDL, LDL, Triglycerides, Total Chol, SBP, DBP, Fasting Glucose; <sup>+</sup>AHI 3% & ESS only

Note only mediators ESS and AHI 3% are associated with confounders sex and *CREBRF* genotype. These relationships will be accounted for in models for all outcomes.

#### 3.3 Structural Equation Modeling: Path Analysis

Each path analysis assessed how a sleep quality measure (AHI 3%, ESS, Avg Snoring Volume, or Sleep Duration) mediated the relationship between BMI and a CVD risk outcome (HDL, LDL, Fasting Glucose, SBP, DBP). In total, 28 models were run. The indirect, direct, and

total effects of BMI on the outcomes were assessed for statistical significance (p < 0.05). The results are shown in Table 8 below for the models whose indirect BMI effects met the specified significance level or were borderline line statistically significant (p < 0.2), as that was the effect of interest.

	BMI Indirect Estimate	BMI Direct Estimate	BMI Total Estimate		
	(p-value)	(p-value)	(p-value)		
1. BMI, AHI 3%, & DBP	0.179 (0.007*)	0.255 (0.050*)	0.434 (0.000*)		
2. BMI, AHI 3%, & HDL	-0.089 (0.049*)	-0.116 (0.196)	-0.205 (0.009*)		
3. BMI, Snore Vol, & HDL	-0.049 (0.190+)	-0.154 (0.075)	-0.203 (0.010*)		
*Denotes $p < 0.05$ ; + Denotes $p < 0.2$					

#### **Table 8: Significant Path Analysis Model Summaries**

Each individual coefficient was also measured in the model and included in the structural equation modeling path diagram in standardized form.



Figure 15: Path Diagram AHI 3% & DBP (Model 1)



Figure 16: Path Diagram AHI 3% & HDL (Model 2)



Figure 17: Path Diagram Snore Volume & HDL (Model 3)

Three final models were chosen based on significant indirect effects of BMI on CVD risk factors. From the results, there is evidence to suggest AHI 3% and Average Snoring Volume mediate the relationship between BMI and HDL cholesterol levels. For both of these indirect

effects, greater BMI is associated with lower average HDL. For AHI 3% (model 2), each 1-unit increase in BMI indirectly results in a 0.089 mg/dL lower average HDL. The total effect of a 1-unit higher BMI in this model is a decrease of 0.205 mg/dL of HDL on average. For Average Snoring Volume (model 3), each 1-unit increase in BMI results in a 0.049 mg/dL lower average HDL. In this model, the total effect of 1-unit higher BMI is a decrease of 0.203 mg/dL of HDL on average.

A significant model was also seen for AHI 3% in mediating the relationship between BMI and Diastolic BP (model 1). This indirect effect increases the diastolic blood pressure as BMI increases. Each one unit increase in BMI indirectly leads to a 0.179 mmHg higher diastolic blood pressure on average. The total effect of a one unit increase in BMI is 0.434 mmHg higher diastolic blood pressure on average.

In all three significant models, it is seen that the majority of the impact of BMI goes through the sleep mediator. This suggests that improvement in AHI 3% and snoring volume could lead to improvement in HDL and DBP in a way similar to the direct impact of lowering BMI. The direct, indirect, and total influence of BMI can be compared using the standardized predicted weights shown in Figures 15 - 17.

Although some anticipated direct effects existed, there was no statistically significant evidence to suggest any indirect effect of BMI through sleep mediators for LDL, Fasting Glucose, or Systolic Blood Pressure. Some significant effects did exist between BMI and sleep measures, but those are not specifically reported on or of interest in this analysis.

The total model fits were assessed using a global fit test, AIC, BIC, and CFI. The fit assessments are shown in the table below.

#### **Table 9: Path Analysis Model Fit Statistics**

Fit Test P-Value AIC BIC CFI

1. BMI, AHI 3%, & DBP	< 0.001	7201.822	7254.579	0.894
2. BMI, AHI 3%, & HDL	< 0.001	7205.515	7258.531	0.901
3. BMI, Snore Vol, & HDL	0.033	6446.024	6491.393	0.957

Based on AIC, BIC, and CFI, the best model to assess the total impact of BMI on HDL uses snore volume as the mediator. Although the p-value was significant for the indirect effect of BMI on DBP through AHI 3%, the CFI falls slightly short of the acceptable 0.9 cutoff.

#### 4.0 Discussion

There were three models determined to be meaningful and worth investigating further. In these three models, the significant indirect effects of BMI suggest that improving the sleep quality measures identified could serve to improve the corresponding CVD risk factors. The first model, where AHI 3% mediates the relationship between BMI and diastolic blood pressure, suggests that improvements in sleep disordered breathing could lower diastolic blood pressure on average. The second model, where AHI 3% mediates the relationship between BMI and HDL cholesterol, suggests that improvements in sleep disordered breathing could lower diastolic blood pressure on average. The second model, where AHI 3% mediates the relationship between BMI and HDL cholesterol, suggests that improvements in sleep disordered breathing could lead to higher levels of HDL cholesterol on average—this is a positive impact, since HDL is considered the "good cholesterol," higher levels of it are better. For the third and final significant model, snoring volume mediates the relationship between BMI and HDL cholesterol, suggesting that lowering snoring volume can also result in higher HDL on average. These results align with literature that BMI has a negative association with HDL and positive association with DBP (McAnulty). In addition, sleep disordered breathing (measured by AHI 3%) is also associated with increased risk of CVD in the literature, which aligns with these results (Floras).

Each of these models reinforces the existing research that diagnosing and treating sleep apnea in Samoa is important in lowering cardiovascular disease risk. Given that both higher AHI 3% and higher snoring volumes are indicators of sleep apnea, these models suggest that beyond the direct impact of sleep quality on CVD risk, improving AHI 3% and snoring volume through sleep apnea treatment can also mitigate the effects of BMI on CVD risk.

Some weaknesses of the research included high levels of missingness for snoring volume and AHI 3%, which are the two mediators that returned statistically significant results. This combined with a sample size of 503 for this novel method leads to a recommendation that further research be done to confirm the relationships found to be significant.

However, there is enough evidence to recommend beginning to plan and implement a public health intervention in Samoa to accurately diagnose sleep apnea and provide proper treatment. This diagnosis and treatment initiative is clearly imperative to Samoans' health, as seen in the data provided by the WatchPat device, where almost all (87.8%) participants have some form of sleep disordered breathing (SDB).

In addition to identifying the indirect effect of BMI on CVD risk factors, the path analysis performed was also able to quantify the effect of BMI on sleep quality. From the global fit results in Table 8 and standardized weights in Figures 14-16 above, it can be seen that BMI has a significant effect on AHI 3% and average snoring volume. This relationship between BMI and sleep quality in Samoan adults is worth investigating further in this cohort with additional analysis in other research initiatives.

Overall, this cohort was able to provide meaningful insight into the ways sleep quality can mediate the relationship between BMI and CVD risk factors in Samoan adults. With Samoans' unique prevalence of genetic variants that lead to higher BMI, it's important to act on findings of this analysis. There appears to be substantial room to improve AHI 3% and snoring volume in Samoan adults through accurate sleep apnea diagnosis. Improvement in these areas of sleep quality is shown to be likely to mitigate the impact of BMI and lead to improved HDL and DBP measures. More research is always recommended, but compelling evidence exists to support implementing public health initiatives and continuing to improve the lives and health of Samoans.

#### **Appendix A Code**

R Studio

R version 3.6.0 (2019-04-26)

Platform: x86 64-redhat-linux-gnu (64-bit)

```
Running under: CentOS Linux 7 (Core)
# Load Libraries
```{r load libraries,message=FALSE}
library(tidyverse)
library(readxl)
library(expss)
library(VIM)
```{r}
load("samoa final.Rdata")
samoa_final = apply_labels(samoa_final, Sex = c("Male" = 0, "Female" = 1), Genotype_code =
c("GG" = 0, "AG" = 1, "AA" = 2))
# Session Information
{}^{r}
sessionInfo()
# Investigate Missingness
```{r}
#Remove fasting glucose values where particiapnt takes diabetes medication
#diabinfo <- subset(samoa alldata, select = c(IDNumber, help diabetes))
#diabinfo$help_diabetes[diabinfo$help_diabetes == -7777] <- NA
#summary(as.factor(diabinfo$help_diabetes))
#diabinfo$gluc_flag2 = ifelse(diabinfo$help_diabetes == 2, 1, ifelse(diabinfo$help_diabetes == 3,
```

1, ifelse(diabinfo\$help\_diabetes == 5, 1, ifelse(diabinfo\$help\_diabetes == 6,1, ifelse(diabinfo\$help\_diabetes == 7,1, ifelse(diabinfo\$help\_diabetes == 9,1,0)))))) #summary(as.factor(diabinfo\$gluc\_flag2)) #diabinfo2 <- subset(diabinfo, select=c(IDNumber,gluc\_flag2))</pre>

```
#samoa_final <- merge(samoa_final,diabinfo2, by="IDNumber")
```

#samoa\_final\$Fasting\_Gluc\_mgdL2 = samoa\_final\$Fasting\_Gluc\_mgdL

#ahi\_3p,epworth\_sleepiness\_score,sleep\_length\_minutes\_weighted,avg\_snoring\_volume, #Predictors

#Fasting\_Gluc\_mgdL,L\_AverageSystolic, L\_AverageDiastolic, LdlC\_mgdL, HdlC\_mgdL, TotChol\_mgdL, Tag\_mgdL)) #outcomes #saved dataset with all edits

#### $(r) \{r\}$

#samoa\_final = subset(samoa\_sleep\_all, select = c(IDNumber,Sex,DecAge,Genotype\_code, #Demographic/Confounders

#L\_BMI, #Mediator

#ahi\_3p,epworth\_sleepiness\_score,sleep\_length\_minutes\_weighted,avg\_snoring\_volume, #Predictors

#Fasting\_Gluc\_mgdL,L\_AverageSystolic, L\_AverageDiastolic, LdlC\_mgdL, HdlC\_mgdL)) #outcomes

#### (r)

#Replacing -7777 with N/A
#samoa\_final[samoa\_final == -7777] <- NA
#samoa\_final\$sleep\_length\_minutes\_weighted[samoa\_final\$sleep\_length\_minutes\_weighted ==
"NA"] <- NA
#samoa\_final\$avg\_snoring\_volume[samoa\_final\$avg\_snoring\_volume == "NA"] <- NA
#library(dplyr)
#samoa\_final <- samoa\_final %>% mutate(crebrf = ifelse(Genotype\_code == 0, 0,
ifelse(Genotype\_code == 1, 2, ifelse(Genotype\_code == 2, 1, NA))))
#samoa\_final\$Genotype\_code <- samoa\_final\$crebrf</pre>

#adding total cholesterol and triglycerides to the model #addchol = subset(samoa\_sleep\_all, select=c(IDNumber, TotChol\_mgdL, Tag\_mgdL)) #addchol[addchol == -7777] <- NA</pre>

```{r}

#samoa\_final = merge(samoa\_final, addchol, by = "IDNumber")
#save(samoa\_final, file="/home/mem344/samoa\_final.Rdata")

```{r}

aggr\_plot <- aggr(samoa\_final, col=c('navyblue','red'), numbers=TRUE, sortVars=TRUE, labels=names(samoa\_final), cex.axis=.4, gap=3, ylab=c("Histogram of missing data","Pattern"))

# Descriptive Statistics
## Sex
```{r}
sextable <- table("Sex"=samoa\_final\$Sex)
sextable2 <- prop.table(sextable)
kable(cbind("Count"= sextable,"Proportion" = sextable2))
```</pre>

## Age
```{r}
summary(samoa\_final\$DecAge)
hist(samoa\_final\$DecAge, main="Distribution of Age", xlab="Age")
```

#### ##BMI

Average BMI for the entire sample is 35.75 ```{r} #Replacing BMI -7777 (verified missing) with NA samoa\_final\$L\_BMI[samoa\_final\$L\_BMI == -7777] <- NA

#Summary of BMI
summary(samoa\_final\$L\_BMI)
hist(samoa\_final\$L\_BMI, main="Distribution of BMI", xlab="BMI")
```

## CREBRF
```{r}
summary(as.factor(samoa\_final\$Genotype\_code))
genetable <- table("CREBRF"=samoa\_final\$Genotype\_code)
genetable2 <- prop.table(genetable)
kable(cbind("Count"= genetable,"Proportion" = genetable2))</pre>

#Sex and CREBRF
table3 <- table(samoa\_final\$Genotype\_code,samoa\_final\$Sex)
table3
prop.table(table3)</pre>

```
## Epworth Sleepiness Score
```{r}
summary(samoa_final$epworth_sleepiness_score)
```

```
hist(samoa final$epworth sleepiness score,main="Epworth
                                                                 Sleepiness
                                                                                    Score
Distribution",xlab="ESS")
## AHI 3%
\sum{r}
summary(samoa_final$ahi_3p)
hist(samoa_final$ahi_3p,main="Sleep Disordered Breathing Distribution",xlab="AHI 3%")
## Sleep Duration
```{r}
samoa final$sleep length minutes weighted
  =
as.numeric(samoa_final$sleep_length_minutes_weighted)
hist(samoa final$sleep length minutes weighted, main="Avg Sleep Duration", xlab="Minutes
per Night")
summary(samoa_final$sleep_length_minutes_weighted)
## Snoring Volume
(r) \{r\}
samoa_final$avg_snoring_volume = as.numeric(samoa_final$avg_snoring_volume)
hist(samoa_final$avg_snoring_volume, main="Avg Snoring Volume", xlab="Decibel Volume")
summary(samoa final$avg snoring volume)
• • •
## Glucose
\sum{r}
hist(samoa_final$Fasting_Gluc_mgdL, main="Fasting Glucose", xlab="mg/dL")
summary(samoa final$Fasting Gluc mgdL)
plot(samoa final$Fasting Gluc mgdL,
  samoa final$L BMI,xlab="Fasting
Glucose", ylab="BMI")
## Systolic BP
```{r}
hist(samoa_final$L_AverageSystolic, main="Systolic BP", xlab="mmHg")
summary(samoa final$L AverageSystolic)
plot(samoa_final$L_AverageSystolic, samoa_final$L_BMI,xlab="Systolic BP",ylab="BMI")
## Diastolic BP
(r) \{r\}
hist(samoa_final$L_AverageDiastolic, main="Diastolic BP", xlab="mmHg")
summary(samoa_final$L_AverageDiastolic)
```

```
plot(samoa_final$L_AverageDiastolic, samoa_final$L_BMI,xlab="Diastolic BP",ylab="BMI")
```

• • • •

#### ## HDL

```{r}
hist(samoa\_final\$HdlC\_mgdL, main="HDL Cholesterol Distn", xlab="mg/dL")
summary(samoa\_final\$HdlC\_mgdL)
plot(samoa\_final\$HdlC\_mgdL, samoa\_final\$L\_BMI,xlab="HDL",ylab="BMI")
```

# ## LDL

```{r}

hist(samoa\_final\$LdlC\_mgdL, main="LDL Cholesterol Distn", xlab = "mg/dL") summary(samoa\_final\$LdlC\_mgdL) plot(samoa\_final\$LdlC\_mgdL, samoa\_final\$L\_BMI,xlab="LDL",ylab="BMI")

## Total Cholesterol
```{r}
hist(samoa\_final\$TotChol\_mgdL, main="Total Cholesterol Distn", xlab = "mg/dL")
summary(samoa\_final\$TotChol\_mgdL)
plot(samoa\_final\$TotChol\_mgdL, samoa\_final\$L\_BMI,xlab="TotChol",ylab="BMI")
```

## Triglycerides
```{r}
hist(samoa\_final\$Tag\_mgdL, main="Triglyceride Distn", xlab = "mg/dL")
summary(samoa\_final\$Tag\_mgdL)
plot(samoa\_final\$Tag\_mgdL, samoa\_final\$L\_BMI,xlab="Triglycerides",ylab="BMI")
```

# Linear Models to Assess Relationships Between Confounders/Predictors

## Sex
```{r}
summary(lm(epworth\_sleepiness\_score ~ as.factor(Sex), data=samoa\_final))
summary(lm(ahi\_3p ~ as.factor(Sex), data=samoa\_final))
summary(lm(avg\_snoring\_volume ~ as.factor(Sex), data=samoa\_final))
summary(lm(sleep\_length\_minutes\_weighted ~ as.factor(Sex), data=samoa\_final))
```

## Age

```{r}
summary(lm(epworth\_sleepiness\_score ~ DecAge, data=samoa\_final))
summary(lm(ahi\_3p ~ DecAge, data=samoa\_final))
summary(lm(avg\_snoring\_volume ~ DecAge, data=samoa\_final))
summary(lm(sleep\_length\_minutes\_weighted ~ DecAge, data=samoa\_final))

## CREBRF

```{r}
summary(lm(epworth\_sleepiness\_score ~ Genotype\_code, data=samoa\_final))
summary(lm(ahi\_3p ~ Genotype\_code, data=samoa\_final))
summary(lm(avg\_snoring\_volume ~ Genotype\_code, data=samoa\_final))
summary(lm(sleep\_length\_minutes\_weighted ~ Genotype\_code, data=samoa\_final))
```

Based on the magnitude and significance level of each relationship, I've determined a noted relationship (path) should exist for sex to AHI 3%, sex to Epworth Sleepiness Score, CREBRF to AHI 3%, and CREBRF to Epworth Sleepiness Score.

# Path Analysis

```{r}
#install.packages("lavaan")
library(lavaan)
library(semPlot)

#Predictors:
#epworth\_sleepiness\_score, ahi\_3p, avg\_snoring\_volume, sleep\_length\_minutes\_weighted

#Outcomes: #Fasting\_Gluc\_mgdL, L\_AverageSystolic, L\_AverageDiastolic, LdlC\_mgdL, HdlC\_mgdL

#Mediator: #L\_BMI

#Confounders:
#IDNumber, Sex, DecAge, Genotype\_code

## HDL

### Simple Test Model AHI 3%, BMI, and HDL \*\*Done to help understand LaVaAn Package\*\*

# ```{r} model1 <- "HdlC\_mgdL ~ a\*ahi\_3p + b\*L\_BMI ahi\_3p ~ c\*L\_BMI</pre>

BMI\_indirect :=  $c^*a$ BMI\_direct := bBMI\_total :=  $b + (c^*a)$ 

```
results <- sem(model1, data=samoa_final)
summary(results, fit.measures=TRUE, standardized=TRUE, rsquare=TRUE)
semPaths(results, "std",
     layout = "tree2", nCharNodes = 0, sizeMan = 15,
     edge.label.cex = 1, edge.label.position = 0.22,
     weighted = FALSE, edge.color = "navy",
     residuals = FALSE, exoCov = FALSE)
...
### Model For AHI 3%
```{r}
model2 <- "HdlC_mgdL ~ b*ahi_3p + c*L_BMI + d*Genotype_code + e*Sex + f*DecAge
      ahi 3p ~ a*L BMI + g*Sex + h*Genotype code
      L BMI \sim i*DecAge + j*Sex + k*Genotype code
     BMI_indirect := a*b
     BMI_direct := c
     BMI total := c + (a*b)
results2 <- sem(model2, data=samoa final)
summary(results2, fit.measures=TRUE, standardized=TRUE, rsquare=TRUE)
semPaths(results2, "std",
     layout = "tree2", nCharNodes = 0, sizeMan = 15,
     edge.label.cex = 1, edge.label.position = 0.22,
     weighted = FALSE, edge.color = "navy",
     residuals = FALSE, exoCov = FALSE)
...
### Model For Epworth Sleepiness Score
```{r}
model3 <- "HdlC mgdL
                              ~ b*epworth sleepiness score + c*L BMI + d*Genotype code
+ e^*Sex + f^*DecAge
      epworth sleepiness score ~ a*L BMI + g*Sex + h*Genotype code
                        \sim i*DecAge + j*Sex + k*Genotype_code
      L_BMI
     BMI_indirect := a*b
     BMI direct := c
     BMI_total := c + (a*b)
results3 <- sem(model3, data=samoa final)
summary(results3, fit.measures=TRUE, standardized=TRUE, rsquare=TRUE)
semPaths(results3, "std",
```

"

```
layout = "tree2", nCharNodes = 0, sizeMan = 15,
     edge.label.cex = 1, edge.label.position = 0.22,
     weighted = FALSE, edge.color = "navy",
     residuals = FALSE, exoCov = FALSE)
...
### Model for Snore Volume
```{r}
model4 <- "HdlC_mgdL
                           \sim b*avg_snoring_volume + c*L_BMI + d*Genotype_code + e*Sex
+ f*DecAge
      avg_snoring_volume ~ a*L_BMI
      L_BMI
                     \sim i*DecAge + j*Sex + k*Genotype_code
     BMI_indirect := a*b
     BMI direct := c
     BMI total := c + (a*b)
results4 <- sem(model4, data=samoa final)
summary(results4, fit.measures=TRUE, standardized=TRUE, rsquare=TRUE)
semPaths(results4, "std",
     layout = "tree2", nCharNodes = 0, sizeMan = 15,
     edge.label.cex = 1, edge.label.position = 0.22,
     weighted = FALSE, edge.color = "navy",
     residuals = FALSE, exoCov = FALSE)
...
### Model for Sleep Duration
\sum{r}
model5 <- "HdlC_mgdL
                                       ~ b*sleep length minutes weighted + c*L BMI +
d*Genotype\_code + e*Sex + f*DecAge
      sleep_length_minutes_weighted ~ a*L_BMI
      L BMI
                     \sim i*DecAge + j*Sex + k*Genotype_code
     BMI indirect := a*b
     BMI direct := c
     BMI_total := c + (a*b)
results5 <- sem(model5, data=samoa_final)
summary(results5, fit.measures=TRUE, standardized=TRUE, rsquare=TRUE)
semPaths(results5, "std",
     layout = "tree2", nCharNodes = 0, sizeMan = 15,
     edge.label.cex = 1, edge.label.position = 0.22,
     weighted = FALSE, edge.color = "navy",
     residuals = FALSE, exoCov = FALSE)
```

• • • •

## LDL

```
### Model for AHI 3%
\sum{r}
model6 <- "LdlC_mgdL ~ b*ahi_3p + c*L_BMI + d*Genotype_code + e*Sex + f*DecAge
      ahi_3p \sim a*L_BMI + g*Sex + h*Genotype_code
      L_BMI ~ i*DecAge + j*Sex + k*Genotype_code
     BMI indirect := a*b
     BMI_direct := c
     BMI total := c + (a*b)
results6 <- sem(model6, data=samoa final)
summary(results6, fit.measures=TRUE, standardized=TRUE, rsquare=TRUE)
semPaths(results6, "std",
     layout = "tree2", nCharNodes = 0, sizeMan = 15,
     edge.label.cex = 1, edge.label.position = 0.22,
     weighted = FALSE, edge.color = "navy",
     residuals = FALSE, exoCov = FALSE)
...
### Model for ESS
(r) \{r\}
model7 <- "LdlC_mgdL
                            ~ b*epworth_sleepiness_score + c*L_BMI + d*Genotype_code +
e^*Sex + f^*DecAge
      epworth\_sleepiness\_score \sim a*L\_BMI + g*Sex + h*Genotype\_code
                    \sim i*DecAge + j*Sex + k*Genotype code
      L BMI
     BMI_indirect := a*b
     BMI direct := c
     BMI total := c + (a*b)
results7 <- sem(model7, data=samoa final)
summary(results7, fit.measures=TRUE, standardized=TRUE, rsquare=TRUE)
semPaths(results7, "std",
     layout = "tree2", nCharNodes = 0, sizeMan = 15,
     edge.label.cex = 1, edge.label.position = 0.22,
     weighted = FALSE, edge.color = "navy",
     residuals = FALSE, exoCov = FALSE)
...
### Model for Snore Volume
```

```
```{r}
model8 <- "LdlC_mgdL
                           \sim b*avg_snoring_volume + c*L_BMI + d*Genotype_code + e*Sex
+ f*DecAge
      avg_snoring_volume ~ a*L_BMI
                     \sim i*DecAge + j*Sex + k*Genotype_code
      L BMI
     BMI_indirect := a*b
     BMI_direct := c
     BMI_total := c + (a*b)
results8 <- sem(model8, data=samoa_final)
summary(results8, fit.measures=TRUE, standardized=TRUE, rsquare=TRUE)
semPaths(results8, "std",
     layout = "tree2", nCharNodes = 0, sizeMan = 15,
     edge.label.cex = 1, edge.label.position = 0.22,
     weighted = FALSE, edge.color = "navy",
     residuals = FALSE, exoCov = FALSE)
...
### Model for Sleep Duration
(r) \{r\}
model9 <- "LdlC_mgdL
                                       ~ b*sleep length minutes weighted + c*L BMI +
d*Genotype code + e*Sex + f*DecAge
      sleep length minutes weighted ~ a*L BMI
      L BMI
                     \sim i*DecAge + j*Sex + k*Genotype code
     BMI indirect := a*b
     BMI direct := c
     BMI total := c + (a*b)
results9 <- sem(model9, data=samoa_final)
summary(results9, fit.measures=TRUE, standardized=TRUE, rsquare=TRUE)
semPaths(results9, "std",
     layout = "tree2", nCharNodes = 0, sizeMan = 15,
     edge.label.cex = 1, edge.label.position = 0.22,
     weighted = FALSE, edge.color = "navy",
     residuals = FALSE, exoCov = FALSE)
...
## Fasting Glucose
### Model for AHI 3%
(r) \{r\}
```

```
model10 <- "Fasting Gluc mgdL ~ b*ahi 3p + c*L BMI + d*Genotype code + e*Sex +
f*DecAge
                   \sim a*L_BMI + g*Sex + h*Genotype_code
      ahi_3p
                     \sim i*DecAge + j*Sex + k*Genotype_code
      L_BMI
     BMI_indirect := a*b
     BMI_direct := c
     BMI_total := c + (a*b)
results10 <- sem(model10, data=samoa final)
summary(results10, fit.measures=TRUE, standardized=TRUE, rsquare=TRUE)
semPaths(results10, "std",
     layout = "tree2", nCharNodes = 0, sizeMan = 15,
     edge.label.cex = 1, edge.label.position = 0.22,
     weighted = FALSE, edge.color = "navy",
     residuals = FALSE, exoCov = FALSE)
...
### Model for ESS
\sum{r}
model11 <- "Fasting Gluc mgdL ~ b*epworth sleepiness score + c*L BMI + d*Genotype code
+ e^*Sex + f^*DecAge
      epworth sleepiness score ~ a*L BMI + g*Sex + h*Genotype code
                    \sim i*DecAge + j*Sex + k*Genotype code
      L BMI
     BMI_indirect := a*b
     BMI direct := c
     BMI_total := c + (a*b)
results11 <- sem(model11, data=samoa final)
summary(results11, fit.measures=TRUE, standardized=TRUE, rsquare=TRUE)
semPaths(results11, "std",
     layout = "tree2", nCharNodes = 0, sizeMan = 15,
     edge.label.cex = 1, edge.label.position = 0.22,
     weighted = FALSE, edge.color = "navy",
     residuals = FALSE, exoCov = FALSE)
•••
### Model for Snore Volume
(r) \{r\}
model12 <- "Fasting_Gluc_mgdL ~ b*avg_snoring_volume + c*L_BMI + d*Genotype_code +
e^*Sex + f^*DecAge
      avg_snoring_volume ~ a*L_BMI
                     \sim i*DecAge + j*Sex + k*Genotype_code
      L BMI
```

```
BMI_indirect := a*b
     BMI_direct := c
     BMI_total := c + (a*b)
results12 <- sem(model12, data=samoa_final)
summary(results12, fit.measures=TRUE, standardized=TRUE, rsquare=TRUE)
semPaths(results12, "std",
     layout = "tree2", nCharNodes = 0, sizeMan = 15,
     edge.label.cex = 1, edge.label.position = 0.22,
     weighted = FALSE, edge.color = "navy",
     residuals = FALSE, exoCov = FALSE)
• • • •
### Model for Sleep Duration
{}^{(n)}{r}
model13 <- "Fasting Gluc mgdL
                                     ~ b*sleep length minutes weighted + c*L BMI +
d*Genotype\_code + e*Sex + f*DecAge
      sleep length minutes weighted ~ a*L BMI
                     \sim i*DecAge + j*Sex + k*Genotype code
      L BMI
     BMI indirect := a*b
     BMI direct := c
     BMI_total := c + (a*b)
results13 <- sem(model13, data=samoa_final)
summary(results13, fit.measures=TRUE, standardized=TRUE, rsquare=TRUE)
semPaths(results13, "std",
     layout = "tree2", nCharNodes = 0, sizeMan = 15,
     edge.label.cex = 1, edge.label.position = 0.22,
     weighted = FALSE, edge.color = "navy",
     residuals = FALSE, exoCov = FALSE)
...
## SBP
### Model for AHI 3%
```{r}
model14 <- "L_AverageSystolic ~ b*ahi_3p + c*L_BMI + d*Genotype_code + e*Sex +
f*DecAge
      ahi 3p
                   \sim a*L_BMI + g*Sex + h*Genotype_code
      L_BMI
                    \sim i*DecAge + j*Sex + k*Genotype_code
```

```
BMI_indirect := a*b
```

```
BMI direct := c
     BMI_total := c + (a*b)
results14 <- sem(model14, data=samoa_final)
summary(results14, fit.measures=TRUE, standardized=TRUE, rsquare=TRUE)
semPaths(results14, "std",
     layout = "tree2", nCharNodes = 0, sizeMan = 15,
     edge.label.cex = 1, edge.label.position = 0.22,
     weighted = FALSE, edge.color = "navy",
     residuals = FALSE, exoCov = FALSE)
...
### Model for ESS
\sum{r}
model15 <- "L AverageSystolic ~ b*epworth sleepiness score + c*L BMI + d*Genotype code
+ e^*Sex + f^*DecAge
      epworth sleepiness score ~ a*L BMI + g*Sex + h*Genotype code
                    \sim i*DecAge + j*Sex + k*Genotype_code
      L BMI
     BMI indirect := a*b
     BMI direct := c
     BMI total := c + (a*b)
results15 <- sem(model15, data=samoa final)
summary(results15, fit.measures=TRUE, standardized=TRUE, rsquare=TRUE)
semPaths(results15, "std",
     layout = "tree2", nCharNodes = 0, sizeMan = 15,
     edge.label.cex = 1. edge.label.position = 0.22.
     weighted = FALSE, edge.color = "navy",
     residuals = FALSE, exoCov = FALSE)
• • •
### Model for Snore Volume
(r) \{r\}
model16 <- "L_AverageSystolic ~ b*avg_snoring_volume + c*L_BMI + d*Genotype_code +
e*Sex + f*DecAge
      avg_snoring_volume ~ a*L_BMI
                     \sim i*DecAge + j*Sex + k*Genotype_code
      L BMI
     BMI indirect := a*b
     BMI direct := c
     BMI_total := c + (a*b)
results16 <- sem(model16, data=samoa_final)
```

```
summary(results16, fit.measures=TRUE, standardized=TRUE, rsquare=TRUE)
semPaths(results16, "std",
     layout = "tree2", nCharNodes = 0, sizeMan = 15,
     edge.label.cex = 1, edge.label.position = 0.22,
     weighted = FALSE, edge.color = "navy",
     residuals = FALSE, exoCov = FALSE)
...
### Model for Sleep Duration
```{r}
model17 <- "L_AverageSystolic
                                     ~ b*sleep_length_minutes_weighted + c*L_BMI +
d*Genotype_code + e*Sex + f*DecAge
      sleep_length_minutes_weighted ~ a*L_BMI
      L BMI
                     \sim i*DecAge + j*Sex + k*Genotype code
     BMI_indirect := a*b
     BMI direct := c
     BMI_total := c + (a*b)
results17 <- sem(model17, data=samoa final)
summary(results17, fit.measures=TRUE, standardized=TRUE, rsquare=TRUE)
semPaths(results17, "std",
     layout = "tree2", nCharNodes = 0, sizeMan = 15,
     edge.label.cex = 1, edge.label.position = 0.22,
     weighted = FALSE, edge.color = "navy",
     residuals = FALSE, exoCov = FALSE)
• • • •
## DBP
### Model for AHI 3%
```{r}
model18 <- "L_AverageDiastolic ~ b*ahi_3p + c*L_BMI + d*Genotype_code + e*Sex +
f*DecAge
      ahi_3p
                   \sim a*L_BMI + g*Sex + h*Genotype_code
      L_BMI
                    \sim i*DecAge + j*Sex + k*Genotype code
     BMI indirect := a*b
     BMI_direct := c
     BMI_total := c + (a*b)
results18 <- sem(model18, data=samoa_final)
summary(results18, fit.measures=TRUE, standardized=TRUE, rsquare=TRUE)
```

```
semPaths(results18, "std",
     layout = "tree2", nCharNodes = 0, sizeMan = 15,
     edge.label.cex = 1, edge.label.position = 0.22,
     weighted = FALSE, edge.color = "navy",
     residuals = FALSE, exoCov = FALSE)
...
### Model for ESS
(r) \{r\}
model19 <- "L AverageDiastolic ~ b*epworth sleepiness score + c*L BMI + d*Genotype code
+ e^*Sex + f^*DecAge
      epworth_sleepiness_score ~ a*L_BMI + g*Sex + h*Genotype_code
                    \sim i*DecAge + j*Sex + k*Genotype code
      L BMI
     BMI indirect := a*b
     BMI direct := c
     BMI_total := c + (a*b)
results19 <- sem(model19, data=samoa_final)
summary(results19, fit.measures=TRUE, standardized=TRUE, rsquare=TRUE)
semPaths(results19, "std",
     layout = "tree2", nCharNodes = 0, sizeMan = 15,
     edge.label.cex = 1, edge.label.position = 0.22,
     weighted = FALSE, edge.color = "navy",
     residuals = FALSE, exoCov = FALSE)
...
### Model for Snore Volume
```{r}
model20 <- "L_AverageDiastolic ~ b*avg_snoring_volume + c*L_BMI + d*Genotype_code +
e*Sex + f*DecAge
      avg_snoring_volume ~ a*L_BMI
                     \sim i*DecAge + j*Sex + k*Genotype code
      L BMI
     BMI indirect := a*b
     BMI_direct := c
     BMI total := c + (a*b)
results20 <- sem(model20, data=samoa final)
summary(results20, fit.measures=TRUE, standardized=TRUE, rsquare=TRUE)
semPaths(results20, "std",
     layout = "tree2", nCharNodes = 0, sizeMan = 15,
     edge.label.cex = 1, edge.label.position = 0.22,
     weighted = FALSE, edge.color = "navy",
```

residuals = FALSE, exoCov = FALSE)

• • • •

```
### Model for Sleep Duration
(r) \{r\}
model21 <- "L_AverageDiastolic
                                      ~ b*sleep_length_minutes_weighted + c*L_BMI +
d*Genotype\_code + e*Sex + f*DecAge
      sleep_length_minutes_weighted ~ a*L_BMI
                     \sim i*DecAge + j*Sex + k*Genotype_code
      L BMI
     BMI_indirect := a*b
     BMI_direct := c
     BMI total := c + (a*b)
results21 <- sem(model21, data=samoa_final)
summary(results21, fit.measures=TRUE, standardized=TRUE, rsquare=TRUE)
semPaths(results21, "std",
     layout = "tree2", nCharNodes = 0, sizeMan = 15,
     edge.label.cex = 1, edge.label.position = 0.22,
     weighted = FALSE, edge.color = "navy",
     residuals = FALSE, exoCov = FALSE)
...
## Triglycerides
### Model for AHI 3%
(r) \{r\}
model22 <- "Tag_mgdL ~ b*ahi_3p + c*L_BMI + d*Genotype_code + e*Sex + f*DecAge
      ahi 3p ~ a*L BMI + g*Sex + h*Genotype code
      L BMI
              \sim i*DecAge + j*Sex + k*Genotype_code
     BMI indirect := a*b
     BMI direct := c
     BMI_total := c + (a*b)
results22 <- sem(model22, data=samoa_final)
summary(results22, fit.measures=TRUE, standardized=TRUE, rsquare=TRUE)
semPaths(results22, "std",
     layout = "tree2", nCharNodes = 0, sizeMan = 15,
     edge.label.cex = 1, edge.label.position = 0.22,
     weighted = FALSE, edge.color = "navy",
     residuals = FALSE, exoCov = FALSE)
...
```

```
### Model for ESS
(r) \{r\}
model23 <- "Tag_mgdL
                            ~ b*epworth_sleepiness_score + c*L_BMI + d*Genotype_code +
e^*Sex + f^*DecAge
      epworth_sleepiness_score ~ a*L_BMI + g*Sex + h*Genotype_code
                    ~ i*DecAge + j*Sex + k*Genotype_code
      L BMI
     BMI_indirect := a*b
     BMI direct := c
     BMI total := c + (a*b)
results23 <- sem(model23, data=samoa_final)
summary(results23, fit.measures=TRUE, standardized=TRUE, rsquare=TRUE)
semPaths(results23, "std",
     layout = "tree2", nCharNodes = 0, sizeMan = 15,
     edge.label.cex = 1, edge.label.position = 0.22,
     weighted = FALSE, edge.color = "navy",
     residuals = FALSE, exoCov = FALSE)
• • •
### Model for Snore Volume
(r)
model24 <- "Tag mgdL
                           ~ b*avg snoring volume + c*L BMI + d*Genotype code + e*Sex
+ f*DecAge
      avg snoring volume ~ a*L BMI
      L BMI
                     \sim i*DecAge + j*Sex + k*Genotype_code
     BMI indirect := a*b
     BMI direct := c
     BMI total := c + (a*b)
results24 <- sem(model24, data=samoa final)
summary(results24, fit.measures=TRUE, standardized=TRUE, rsquare=TRUE)
semPaths(results24, "std",
     layout = "tree2", nCharNodes = 0, sizeMan = 15,
     edge.label.cex = 1, edge.label.position = 0.22,
     weighted = FALSE, edge.color = "navy",
     residuals = FALSE, exoCov = FALSE)
• • • •
### Model for Sleep Duration
(r) \{r\}
model25 <- "Tag_mgdL
                                       ~ b*sleep_length_minutes_weighted + c*L_BMI +
d*Genotype\_code + e*Sex + f*DecAge
```

```
sleep_length_minutes_weighted ~ a*L_BMI
      L BMI
                      \sim i*DecAge + j*Sex + k*Genotype_code
     BMI_indirect := a*b
     BMI_direct := c
     BMI_total := c + (a*b)
results25 <- sem(model25, data=samoa_final)
summary(results25, fit.measures=TRUE, standardized=TRUE, rsquare=TRUE)
semPaths(results25, "std",
     layout = "tree2", nCharNodes = 0, sizeMan = 15,
     edge.label.cex = 1, edge.label.position = 0.22,
     weighted = FALSE, edge.color = "navy",
     residuals = FALSE, exoCov = FALSE)
• • •
## Total Cholesterol
### Model for AHI 3%
{}^{(r)}
model26 <- "TotChol_mgdL ~ b*ahi_3p + c*L_BMI + d*Genotype_code + e*Sex + f*DecAge
      ahi 3p ~ a*L BMI + g*Sex + h*Genotype code
      L BMI \sim i*DecAge + i*Sex + k*Genotype code
     BMI_indirect := a*b
     BMI direct := c
     BMI_total := c + (a*b)
results26 <- sem(model26, data=samoa final)
summary(results26, fit.measures=TRUE, standardized=TRUE, rsquare=TRUE)
semPaths(results26, "std",
     layout = "tree2", nCharNodes = 0, sizeMan = 15,
     edge.label.cex = 1, edge.label.position = 0.22,
     weighted = FALSE, edge.color = "navy",
     residuals = FALSE, exoCov = FALSE)
•••
### Model for ESS
(r) \{r\}
                               ~ b*epworth_sleepiness_score + c*L_BMI + d*Genotype_code
model27 <- "TotChol_mgdL
+ e^*Sex + f^*DecAge
      epworth\_sleepiness\_score \sim a*L\_BMI + g*Sex + h*Genotype\_code
                    \sim i*DecAge + i*Sex + k*Genotype_code
      L BMI
```

```
46
```

```
BMI_indirect := a*b
     BMI_direct := c
     BMI_total := c + (a*b)
results27 <- sem(model27, data=samoa_final)
summary(results27, fit.measures=TRUE, standardized=TRUE, rsquare=TRUE)
semPaths(results27, "std",
     layout = "tree2", nCharNodes = 0, sizeMan = 15,
     edge.label.cex = 1, edge.label.position = 0.22,
     weighted = FALSE, edge.color = "navy",
     residuals = FALSE, exoCov = FALSE)
• • • •
### Model for Snore Volume
```{r}
model28 <- "TotChol_mgdL
                                 \sim b*avg snoring volume + c*L BMI + d*Genotype code +
e*Sex + f*DecAge
      avg snoring volume ~ a*L BMI
                     \sim i*DecAge + j*Sex + k*Genotype code
      L BMI
     BMI indirect := a*b
     BMI direct := c
     BMI_total := c + (a*b)
results28 <- sem(model28, data=samoa_final)
summary(results28, fit.measures=TRUE, standardized=TRUE, rsquare=TRUE)
semPaths(results28, "std",
     layout = "tree2", nCharNodes = 0, sizeMan = 15,
     edge.label.cex = 1, edge.label.position = 0.22,
     weighted = FALSE, edge.color = "navy",
     residuals = FALSE, exoCov = FALSE)
...
### Model for Sleep Duration
(r) \{r\}
model29 <- "TotChol_mgdL
                                        ~ b*sleep_length_minutes_weighted + c*L_BMI +
d*Genotype code + e*Sex + f*DecAge
      sleep_length_minutes_weighted ~ a*L_BMI
      L BMI
                     \sim i*DecAge + j*Sex + k*Genotype_code
     BMI indirect := a*b
     BMI direct := c
     BMI_total := c + (a*b)
```

results29 <- sem(model29, data=samoa\_final) summary(results29, fit.measures=TRUE, standardized=TRUE, rsquare=TRUE)

```
semPaths(results29, "std",
    layout = "tree2", nCharNodes = 0, sizeMan = 15,
    edge.label.cex = 1, edge.label.position = 0.22,
    weighted = FALSE, edge.color = "navy",
    residuals = FALSE, exoCov = FALSE)
```

• • • •

"

#### **Bibliography**

- Chellappa, Sarah L., Nina Vujovic, Jonathan S. Williams, and Frank A.J.L. Scheer. "Impact of Circadian Disruption on Cardiovascular Function and Disease." Trends in Endocrinology & Metabolism 30, no. 10 (October 2019): 767–79. https://doi.org/10.1016/j.tem.2019.07.008.
- Chen, Xiaoli, Rui Wang, Phyllis Zee, Pamela L. Lutsey, Sogol Javaheri, Carmela Alcántara, Chandra L. Jackson, Michelle A. Williams, and Susan Redline. "Racial/Ethnic Differences in Sleep Disturbances: The Multi-Ethnic Study of Atherosclerosis (MESA)." SLEEP, June 1, 2015. https://doi.org/10.5665/sleep.4732.
- Ezeamama, Amara E., Satupaitea Viali, John Tuitele, and Stephen T. McGarvey. "The Influence of Socioeconomic Factors on Cardiovascular Disease Risk Factors in the Context of Economic Development in the Samoan Archipelago." Social Science & Medicine 63, no. 10 (November 1, 2006): 2533–45. https://doi.org/10.1016/j.socscimed.2006.06.023.
- Fernandez-Mendoza, Julio, Alexandros N. Vgontzas, Ilia Kritikou, Susan L. Calhoun, Duanping Liao, and Edward O. Bixler. "Natural History of Excessive Daytime Sleepiness: Role of Obesity, Weight Loss, Depression, and Sleep Propensity." Sleep 38, no. 3 (March 1, 2015): 351–60. https://doi.org/10.5665/sleep.4488.
- Floras, John S. "Sleep Apnea and Cardiovascular Disease," n.d., 24. https://doi.org/10.1161/CIRCRESAHA.118.310783.
- Galy, Olivier, Emilie Paufique, Akila Nedjar-Guerre, Fabrice Wacalie, Guillaume Wattelez, Pierre-Yves Le Roux, Solange Ponidja, et al. "Living in Rural and Urban Areas of New Caledonia: Impact on Food Consumption, Sleep Duration and Anthropometric Parameters Among Melanesian Adolescents." Nutrients 12, no. 7 (July 10, 2020): 2047. https://doi.org/10.3390/nu12072047.
- Hawley, Nicola L., Ryan L. Minster, Daniel E. Weeks, Satupaitea Viali, Muagututia Sefuiva Reupena, Guangyun Sun, Hong Cheng, Ranjan Deka, and Stephen T. Mcgarvey.
  "Prevalence of Adiposity and Associated Cardiometabolic Risk Factors in the Samoan Genome-Wide Association Study: Gwas 2010: NCD & Associated Risk Factor Prevalence." American Journal of Human Biology 26, no. 4 (July 2014): 491–501. https://doi.org/10.1002/ajhb.22553.
- Hawley, Nicola L, Alysa Pomer, Anna C. Rivara, Samantha L Rosenthal, Rachel L Duckham, Jenna C Carlson, Take Naseri, et al. "Exploring the Paradoxical Relationship of a Creb 3 Regulatory Factor Missense Variant With Body Mass Index and Diabetes Among Samoans: Protocol for the Soifua Manuia (Good Health) Observational Cohort Study." JMIR Research Protocols 9, no. 7 (July 23, 2020): e17329. https://doi.org/10.2196/17329.

- Huang, Tianyi, Sara Mariani, and Susan Redline. "Sleep Irregularity and Risk of Cardiovascular Events." Journal of the American College of Cardiology 75, no. 9 (March 2020): 991–99. https://doi.org/10.1016/j.jacc.2019.12.054.
- Johns, Murray W. "A New Method for Measuring Daytime Sleepiness: The Epworth Sleepiness Scale." Sleep 14, no. 6 (November 1, 1991): 540–45. https://doi.org/10.1093/sleep/14.6.540.
- MSc, Judy McAnulty, and Robert Scragg MBBS PhD. "Body Mass Index and Cardiovascular Risk Factors in Pacific Island Polynesians and Europeans in New Zealand." Ethnicity & Health 1, no. 3 (September 1, 1996): 187–95. https://doi.org/10.1080/13557858.1996.9961787.
- Nayebi, Hooshang. "Path Analysis." In Advanced Statistics for Testing Assumed Casual Relationships, by Hooshang Nayebi, 47–78. University of Tehran Science and Humanities Series. Cham: Springer International Publishing, 2020. https://doi.org/10.1007/978-3-030-54754-7\_2.
- Nicola L. Hawley1\*, Rachel L. Duckham, Jenna C. Carlson, Take Naseri, Muagututia Sefuiva Reupena, and Viali Lameko. "CREBRF and Body Composition in Samoans," n.d.
- Pearl, Judea. "The Causal Foundations of Structural Equation Modeling:" Fort Belvoir, VA: Defense Technical Information Center, February 16, 2012.
- Stephanie Glen. "Index of Fit / Fit Indices: Simple Definitions for Popular Measures." Statistics How To, December 6, 2017. https://www.statisticshowto.com/index-of-fit/.
- Wang, Heming, Jacqueline M. Lane, Samuel E. Jones, Hassan S. Dashti, Hanna M. Ollila, Andrew R. Wood, Vincent T. van Hees, et al. "Genome-Wide Association Analysis of Self-Reported Daytime Sleepiness Identifies 42 Loci That Suggest Biological Subtypes." Nature Communications 10, no. 1 (December 2019): 3503. https://doi.org/10.1038/s41467-019-11456-7.