**Statistical Analysis Plan (SAP)**

|  |  |  |
| --- | --- | --- |
| **Title** | | The relationship between glucose and MI using Framingham dataset |
| **CRU/Department/Division/Center** | | Duke University |
| **IRB Number** | |  |
| **Investigators:** | |  |
| **Lead Investigator** | |  |
| **Mentors** | |  |
| **Biostatistician(s)** | | Hayley Nemeth, Zach Frere, Zhenhui Xu |
| **Supervising Biostatistician** | | Dr. Elizabeth Hauser |
| **Original Creation Date** | | 2020-04-06 |
| **Version Date** | | 2020-04-06 |
| **Project Folder Location** | |  |
| **Project Goal(s)** | |  |
| **Submission Deadline(s)** | | 2020-04-06 |
| **Investigator Agreement** | All statistical analyses included in an abstract or manuscript should reflect the work of the biostatistician(s) listed on this SAP. No changes or additional analyses should be made to the results or findings without discussing with the project biostatistician(s).  All biostatisticians on this SAP should be given sufficient time to review the full presentation, abstract, manuscript, or grant and be included as co-authors on any abstract or manuscript resulting from the analyses.  If substantial additional analysis is necessary or the aims of the project change, a new SAP will need to be developed.  Publications resulting from this SAP are supported in part by the Duke CTSA and must cite grant number UL1TR002553 and be submitted to PubMed Central.  I have reviewed the SAP and understand that any changes must be documented.  *Acknowledged by:* Click or tap here to enter text.  *Date:* Click or tap to enter a date. | |
| **Activity Log** |  | |

# Study Overview

Previous studies have shown that high level blood glucose has prognostic value for patients with myocardial infarction (MI) [1]. These studies primarily focus on the admission blood glucose and development of MI. An elevated blood glucose is a risk factor for increased mortality and in-hospital complications whether these patients have diabetes or not [2]. Knowledge of MI risk factors can aid in clinicians’ decision making and death rate prediction [3].

## Study Aims

1. To describe the distribution of glucose at Period 1 and incidence of MI at Period 2.
2. To test the association between glucose and the incidence of MI.

## Study Hypotheses

## Primary Hypotheses

1. We hypothesize that higher glucose is associated with higher incidence of MI.

where is the population incidence of MI among people with higher glucose and is the population incidence of MI among people with lower glucose.

## Gaps in Knowledge

1. Although previous studies have shown that glucose is a indicative of mortality rate in patients with MI, few data are available to evaluate glucose as a risk factor of incident MI for those without MI.

# Study Population

Our study population consists of 4,434 participants in the community of Framingham, Massachusetts, enrolled in 1948. Each participant was followed up for 24 years to collect cardiovascular outcomes. Each participant has 1 to 3 observations and thus there are 11,627 observations on 4,434 participants. The inclusion and exclusion criteria could be seen in Framingham study.

## Data Acquisition

*Fill in all relevant information:*

|  |  |
| --- | --- |
| Study design | A retrospective longitudinal study |
| Data source/how the data were collected | A subset of data collected in Framingham study |
| Contact information for team member responsible for data collection/acquisition |  |
| Date or version (if downloaded, provide date) | 2020-02-26 |
| Data transfer method and date | Sakai |
| Where dataset is stored | Sakai |

# Outcomes, Exposures, and Additional Variables of Interest

## Primary Outcome(s)

|  |  |  |  |
| --- | --- | --- | --- |
| **Outcome** | **Description** | **Variables and Source** | **Specifications** |
| MI at Period 2 | Prevalent Myocardial Infarction | Variable Name: PREVMI, PERIOD | 0-No prevalent MI at Period 2  1-Prevalent MI at Period 2  PERIOD=2 |

## Additional Variables of Interest

|  |  |  |  |
| --- | --- | --- | --- |
| **Variable** | **Description** | **Variables and Source** | **Specifications** |
| glucose | Casual serum glucose (mg/dL) | Variable Name: GLUCOSE  Period =1 | Continuous variable  Range: 39-478 |
| age | Age at exam (years) | Variable Name: AGE  Period = 1 | Continuous variable  Range: 32-81 |
| hypertension | Subject was defined as hypertensive if treated or if second exam at which mean systolic was >=140 mmHg or mean Diastolic >=90 mmHg | Variable Name: PREVHYP  Period = 1 | Binary Variable |

# Statistical Analysis Plan

## Demographic and Clinical Characteristics (“Table 1”)

1. Baseline patient characteristics at the time of period 1
2. Variables will be summarized by mean and standard deviation for continuous variables and count and percentile for categorical variables.

## Analyses Plan for Aim 1

1. Association between glucose and the incidence of MI
2. Two sample t test will be used to test the difference of glucose between people with MI and without MI if the distribution of glucose is normal. Mann Whitney test will be used to test the difference of glucose between people with MI and without MI if the distribution of glucose is not normal.

*Interpretation:* If we find a statistically significant difference of glucose between people with and without MI (p value < 0.05), this would imply that patients with higher glucose have higher risk of incident MI.

1. Logistic regression models will be used to test the association between glucose and the risk of diabetes. The multivariate model will be as follows:

where p(X) represents the incidence of MI.

*Interpretation:* If we find a statistically significant association between glucose and incidence of MI (p value <0.05), this would imply that patients with higher glucose have higher risk of incident MI.

# References

1. Admission glucose concentrations independently predict early and late mortality in patients with acute myocardial infarction treated by primary or rescue percutaneous coronary intervention.
2. Admission Glucose and Mortality in Elderly Patients Hospitalized With Acute Myocardial Infarction. Mikhail Kosiborod, MD , Saif S. Rathore, MPH , Silvio E. Inzucchi, MD , Frederick A. Masoudi, MD , Yongfei Wang, MS , Edward P. Havranek, MD , and Harlan M. Krumholz, MD, SM
3. Ohman EM, Granger CB, Harrington RA, Lee KL: Risk stratification and therapeutic decision making in acute coronary syndromes.

**Data Analysis Report**

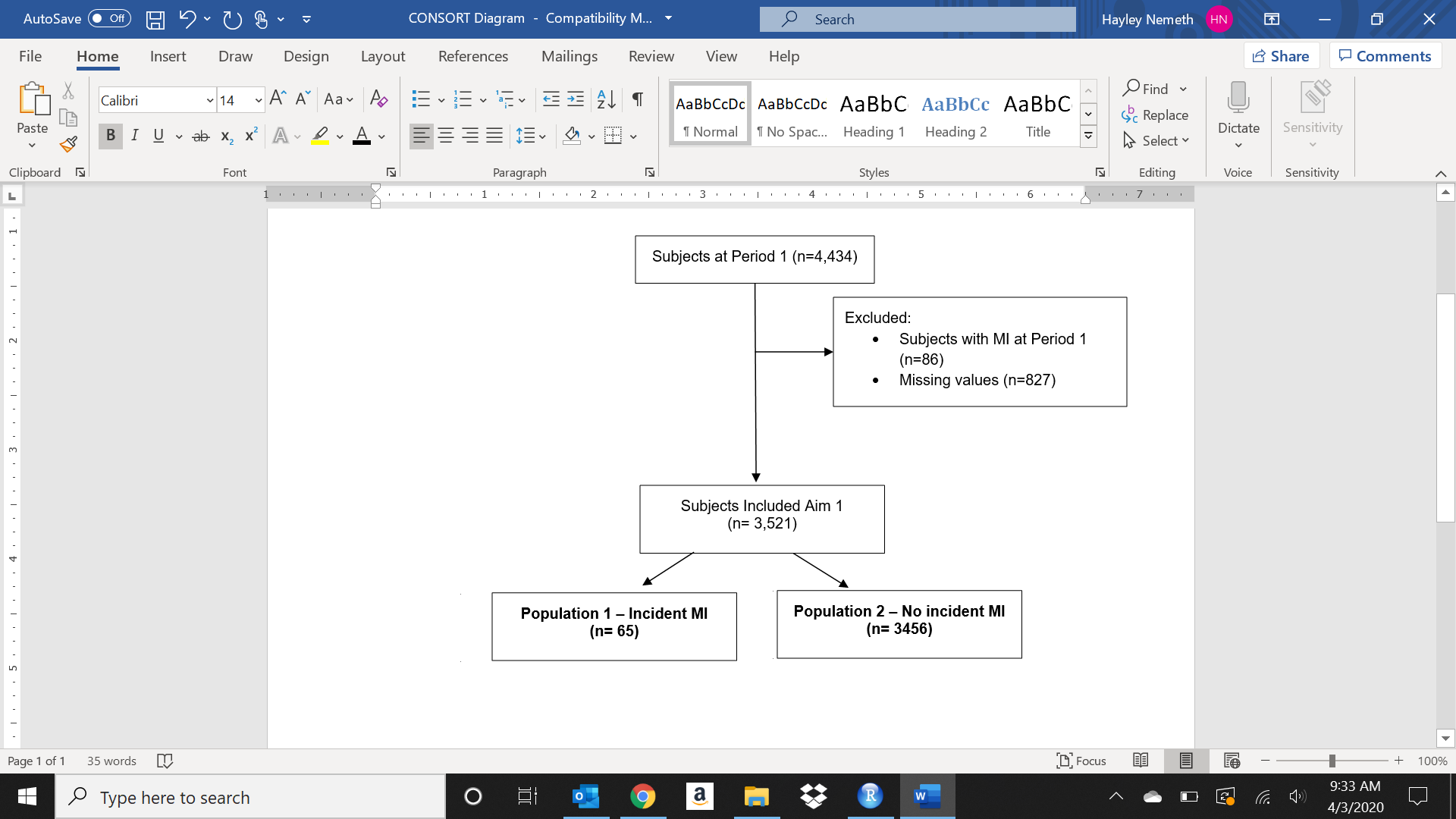
1. **Demographics and Clinical Characteristics**
2. There are 4434 observations in period 1 examination. Subjects with MI at Period 1 are excluded (n=86). There are 827 missing values in glucose measurements and prevalent myocardial infarction. In study aim 1, 3521 subjects are included in the statistical analysis.

Figure 1: Consort Diagram of inclusion and exclusion of participants

Table1: Baseline patient characteristic for aim 1(Incidence of Period 2 MI Given Period 1 Glucose)

|  |  |  |  |
| --- | --- | --- | --- |
|  | MI  (n=65) | No MI  (n=3456) | Overall  (n=3521) |
| **PREVHYP** |  |  |  |
| hypertensive | 35 (53.8%) | 1034 (29.9%) | 1069 (30.4%) |
| Non-hypertensive | 30 (46.2%) | 2422 (70.1%) | 2452 (69.6%) |
| **AGE** |  |  |  |
| Mean (SD) | 52.9 (7.66) | 49.4 (8.57) | 49.4 (8.57) |
| Median (Min, Max) | 53.0 (38.0, 67.0) | 48.0 (32.0, 70.0) | 49.0 (32.0, 70.0) |
| **GLUCOSE** |  |  |  |
| Mean (SD) | 82.2 (16.8) | 81.3 (21.3) | 81.3 (21.2) |
| Median (Min, Max) | 80.0 (50.0, 163) | 78.0 (40.0, 394) | 78.0 (40.0, 394) |

1. **The association between glucose and the incidence of MI**
2. Mann Whitney Test

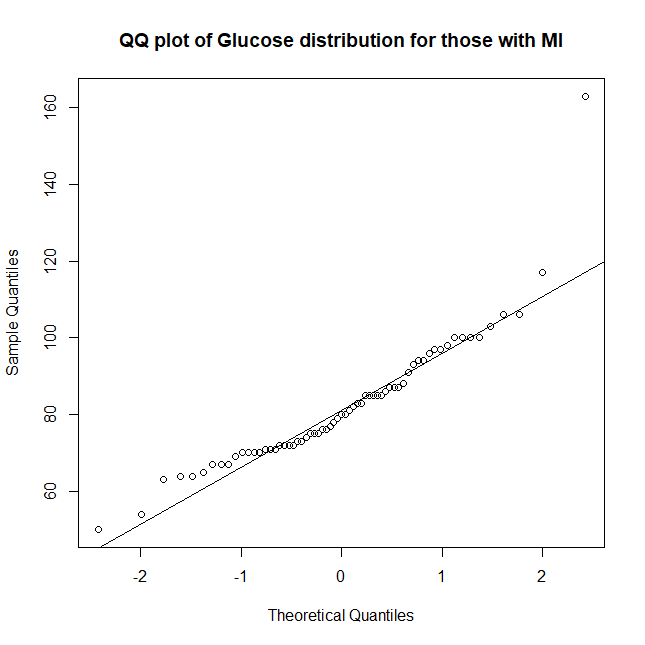
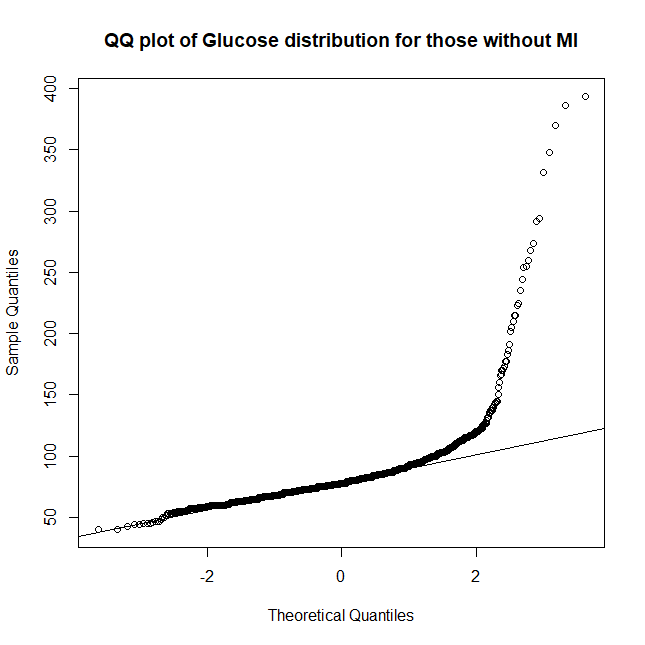
The distribution of glucose for those with and without MI is not normal. Mann Whitney test is applied since it transforms the variable to rank. The p value of Mann Whitney test is 0.377. There is no enough evidence to conclude that there is significant difference in glucose between those with MI and those without MI. However, Mann Whitney test turns continuous variables into ranks which reduces the power of the test.

Figure 2: QQ plot of glucose for those with and without MI

Table 2: The test statistic and the p value of Mann Whitney Test

|  |  |  |
| --- | --- | --- |
|  | Test Statistic | P Value |
| Glucose | 1.19e+05 | 0.377 |

1. Logistic Regression Model

As it is shown in table 3, there is no evidence to conclude glucose and incidence of MI are associated (p value = 0.722). Similarly, there is no evidence to conclude that there is significant association between glucose and the incidence of MI (p value of = 0.912) after adjusted for age and hypertension.

Table 3: The estimated coefficients and the p value of logistic regression

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | Univariate Regression Model | | Multivariate Regression Model | |
|  | Estimated Coefficients (±SE) | P value | Estimated Coefficients (±SE) | P value |
| Hypertension | 1.01 (±0.252) | 6.44e-05 | 0.822 (±0.266) | 0.00199 |
| Age | 4.74e-02 (±0.0146) | 0.00116 | 0.0328 (±0.0155) | 0.0345 |
| Glucose | 1.85e-03 (±0.00521) | 0.722 | -6.31e-04 (±0.00572) | 0.912 |

1. **Conclusion**

Our analysis indicates insufficient evidence to support the null hypothesis that higher glucose is associated with higher incidence of myocardial infarction.

1. The Mann Whitney test shows that there is no significant difference in glucose for those with MI and without MI.
2. The logistic regression model shows that after adjusting for age and hypertension, there is no significant association between glucose and incidence of MI.
3. **Discussion**
4. In this study, glucose was tested by a random blood sugar test. The blood samples were taken at a random time which might influence the measured value.
5. The proportion of incident MI was low in the sampling population. The difference of proportion of incident MI by glucose might become harder to detect.