RELATIONSHIP OF AGE, GENDER, AND CHOLESTEROL ON TYPE II DIABETES IN AFRICAN AMERICAN ADULTS FROM CENTRAL VIRGINIA

Author: Himshikha Das

This is a term paper for Biostatistics course.

ABSTRACT

Background: About 13% of all US adults have diabetes, and 34.5% of adults meet the criteria for prediabetes. Low cholesterol levels are consistently associated with an increased risk of being diagnosed with diabetes mellitus II. Individuals that are at least 45 years of age, and of the male sex are also at an increased risk of developing diabetes mellitus II.

Methods: A random sample of 1046 African Americans from Central Virginia were interviewed, of which 403 were chosen to be participants in the study. Eighteen measurements were taken in each of the participants, including cholesterol and glycosylated hemoglobin. Other variables, such as age and gender, were also recorded.

Results: Correlation between glycosylated hemoglobin and cholesterol is low, while glycosylated hemoglobin levels and age are moderately correlated. Glycosylated hemoglobin and gender had no association. Out of all the models, we can see that age and cholesterol can best explain the variance in participants' glycosylated hemoglobin levels.

Conclusion: Age and cholesterol level are correlated with glycosylated hemoglobin levels in the study sample. Gender was not associated with glycosylated hemoglobin levels.

INTRODUCTION

According to the CDC, an estimated 13% of all adults in the United States have diabetes and 34.5% of adults meet the criteria for prediabetes (Centers for Disease Control and Prevention, 2020). Diabetes Mellitus Type II (DM II) is the most common form of diabetes, making up for about 90% of all diabetes diagnoses. In individuals with DM II, cells in the body experience a diminished response to insulin, more frequently referred to as *insulin resistance* (Goyal & Jailal, 2021). Insulin is a hormone that allows glucose to enter body cells and be used for energy. In those with insulin resistance, cells struggle to absorb glucose. This leads to an overproduction of insulin by the pancreas, bringing about DM II (*Insulin Resistance and Prediabetes*, 2018). DM II is characterized by high blood sugar levels as they are not regulated in those with diabetes, as a result of insulin resistance. Glycosylated hemoglobin levels, representing the amount of sugar in the blood, can be measured using blood samples. Glycosylated hemoglobin levels greater than 7.0 mg/dL are indicative of DM II.

There are several risk factors for developing DM II. Being African American, Hispanic, American Indian, Asian American, or Pacific Islander increases an individual's risk of developing DM II. Individuals that are 45 years of age or older makes individuals more susceptible to developing DM II (*Type 2 diabetes,* 2021). Recent studies have provided evidence that HDL cholesterol may be a key contributor to glucose homeostasis in the human body through insulin secretion, direct glucose uptake, and enhanced insulin sensitivity (Drew et al., 2012). Studies have also shown that variables such as gender and age have a significant effect on glycosylated hemoglobin levels.

In this analysis, we aim to explore the relationship between age, gender, cholesterol levels, and glycosylated hemoglobin levels among the 403 study subjects selected.

METHODS

Study Design:
At the University of Virginia School of Medicine, 1046 participants were randomly selected and interviewed. Eighteen variables were measured to understand the prevalence of obesity, diabetes, and cardiovascular disease in African Americans from central Virginia. A subset of 403 participants was selected from this population via stratified sampling.

In this study we attempt to assess potential correlation in exposure-outcome relationship between age, gender, glycosylated hemoglobin, and cholesterol levels in the study sample population. The variables in this study were not manipulated or assigned. Therefore by definition, the observational study design is used (Gerstman, 2015).

Measures:
Data on eighteen variables was collected from the 403 participants in the study sample. The variables measured were height, weight, blood pressure, cholesterol levels, glycosylated hemoglobin levels, and more. Variables such as age and gender were noted for each of the participants as well. All subjects with blood pressure that read greater than or equal to 140/90, as well as some participants with blood pressure readings less than or equal to 140/90, a second BP measurement was taken approximately 10 minutes later. Blood was drawn from study subjects and transported on ice to analyze subjects' total cholesterol, high density lipoprotein cholesterol, and glycosylated hemoglobin. Glycosylated hemoglobin levels, reflective of the amount of sugar in the blood, were used as a predictor of diabetes. Subjects with glycosylated hemoglobin levels greater than 7.0 were diagnosed with diabetes.

Statistical Methods:
The data was collected via stratified sampling. Tabulated data obtained and figures generated using Microsoft Excel. Data cleaning followed by an exploration of data was done. Statistical data analysis was performed using the R programming language version 4.1.1 (2021-08-10) via RStudio. The statistical analysis portion of this study was conducted with RStudio.

We will begin to evaluate the data via tables created by RStudio software. Then, we will identify patterns in distribution between the parameters of interest including age, gender, cholesterol, and glycosylated hemoglobin. Next, we will examine the correlation between age, gender, cholesterol, with glycosylated hemoglobin in the sample population. To do this, we will perform statistical analysis on the dataset, and determine whether the results reject or fail to reject our hypothesis at alpha=0.05 level.

For regression analysis, we achieve this by using the lm() function in R and the output is called using the summary() function on the model. The dependent variable is glycosylated hemoglobin

and the independent variables are age/gender/cholesterol, together with the data being used being the diabetes dataset.

RESULTS

Descriptive Characteristics of Sample:
A total of 403 participants, with an average age of 46.9 years (SD = 16.3), were recruited for this study and evaluated on four variables: age, gender, cholesterol level (chol), and glycosylated hemoglobin level. Descriptive statistics for the variables are reported in Table 1.

Table 1: Summarized descriptive data for study participants

Cholesterol	
Mean (SD)	208 (44.4)
Median [Min, Max]	204 [78.0, 443]
Glycosylated Hemoglobin	
Mean (SD)	5.59 (2.21)
Median [Min, Max]	4.87 [2.68, 16.1]
Age	
Mean (SD)	46.9 (16.3)
Median [Min, Max]	45.0 [19.0, 92.0]
Gender	
Mean (SD)	234 (58.1%)
Median [Min, Max]	169 (41.9%)

Figure 1: Glycosylated hemoglobin levels relative to age

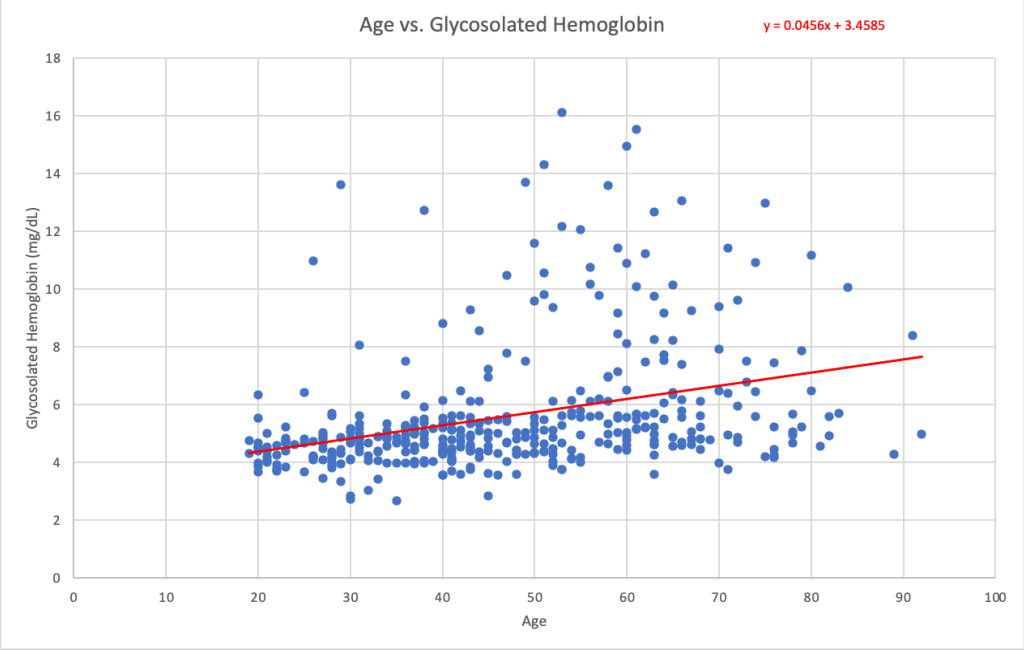


Figure 2: Glycosylated hemoglobin levels relative to gender.

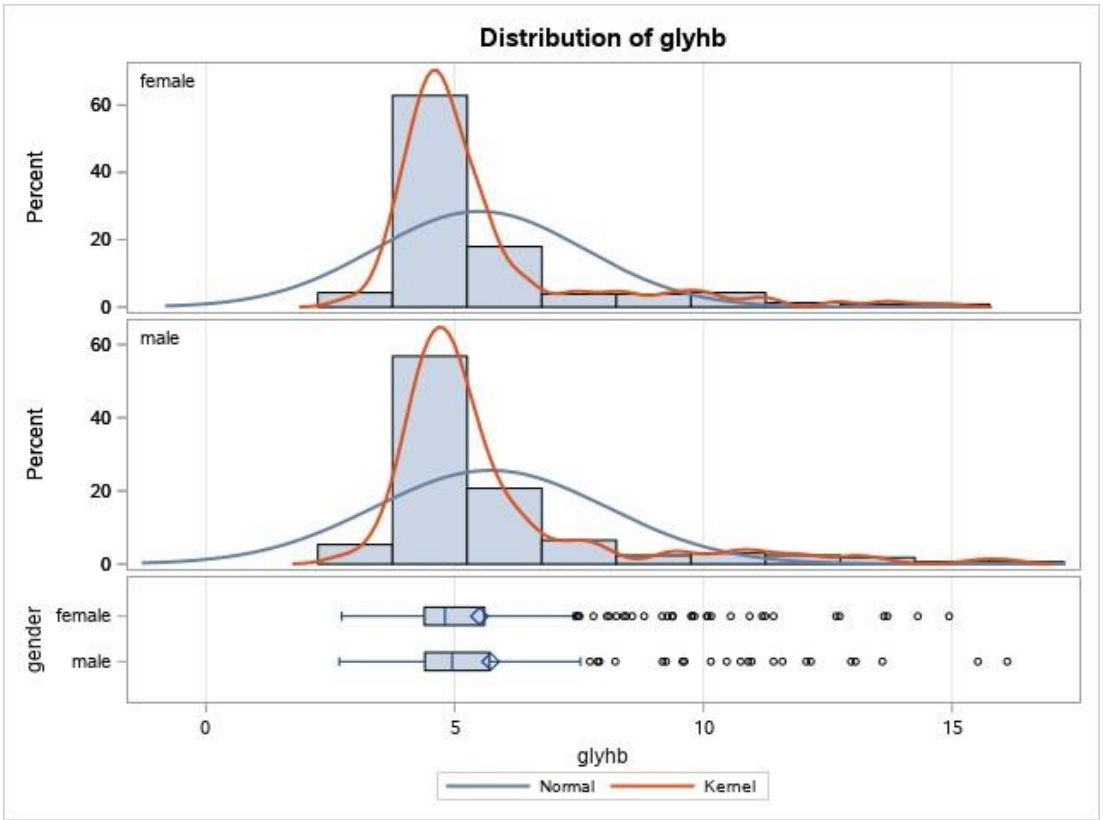
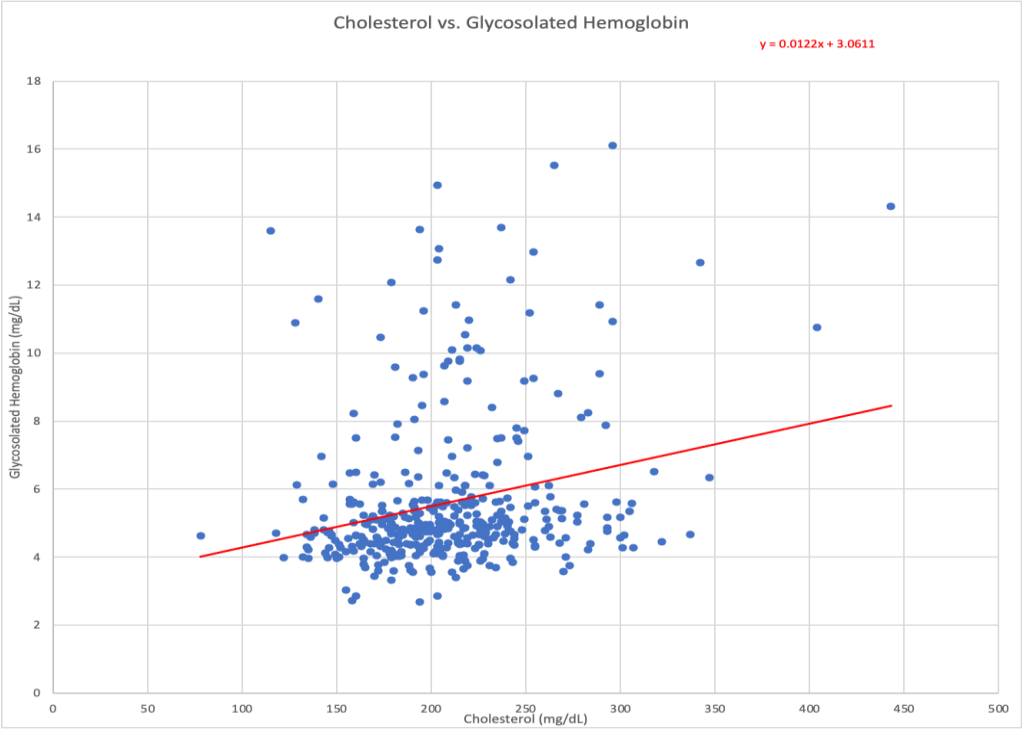


Figure 3: Relationship between levels of glycosylated hemoglobin levels and cholesterol



BIVARIABLE ANALYSIS

Figure 1: Glycosylated hemoglobin and age

Elevated levels of glycosylated hemoglobin become more frequent as age increases until about 65 years old when glycosylated hemoglobin levels and the overall number of participants begin to decrease. Age can explain 11.34% variation in glycosylated hemoglobin values. For every 1 unit increase in the age of a subject, glycosylated hemoglobin increases by 0.045. 11.12% of the variance found in the dependent variable can be explained by the independent variable. The F-statistic shows that the null hypothesis can be rejected and there is a relationship between glycosylated hemoglobin and age ($p<0.001$). The correlation value for pearson is .33 and .43 for spearman and signifies medium correlation between age and glycosylated hemoglobin. [glycosylated hemoglobin = 3.45 +.045*age]

Welch's t-test is a two-sample location test that is used to test the hypothesis that two populations have equal means. Here we are looking at the glycosylated hemoglobin and gender. Our assumption is that when using p-values to determine the results of a test the general rule is no significant differences were detected if $p > 0.05$, and based on the results the means of the two sample groups are NOT significantly different. Here, t statistic = -0,98363, degree of freedom is 338.78, p-value = 0.326. The mean in the female group is 5.499274 and the mean in the group male is 5.722130.

Figure 2: Glycosylated hemoglobin and gender

For the male subjects, glycosylated hemoglobin increases by 0.22. Gender can explain 0.2% variation in glyhb values which is insignificant. 1.428×10^{-4}% of the variance found in the dependent variable can be explained by the independent variable. The F-statistic shows that the null hypothesis can be rejected and there is no relationship between glyhb and gender. ($p=0.3178$). [glycosylated hemoglobin = 5.49+.22*gender]

Figure 3: Glycosylated hemoglobin and cholesterol

As shown in figure 3, the data is mostly concentrated around 200 mg/dL for cholesterol levels, and between 4 and 6 mg/dL for glycosylated hemoglobin. Since the majority of the data lies between 100 to 300, a few outliers could also be seen. Cholesterol level can explain 6% variation in glycosylated hemoglobin values which is insignificant. For every 50 mg/dL in the chol of a subject, glycosylated hemoglobin increases by .01. The F-statistic shows that the null hypothesis can be rejected and there is a relationship between glycosylated hemoglobin and chol ($p<0.001$). The correlation value for pearson is .24 and .23 for spearman and signifies low correlation between cholesterol and glycosylated hemoglobin. [glycosylated hemoglobin = 3.06+.01*chol]

MULTIVARIABLE ANALYSES:

A multivariable linear regression was performed to determine whether glycosylated hemoglobin is related to age and cholesterol. Glycosylated hemoglobin is the dependent variable, while age and cholesterol are independent variables. The relationship between glycosylated hemoglobin, age, and cholesterol was tested against the claim of no association. The analysis resulted with Adjusted R-squared of 0.1384 which shows that 13.84% of variation within glycosylated

hemoglobin is explained by age and cholesterol. [glycosylated hemoglobin = 1.899 +.008*chol + .04 age] where the slope explains that for every 10 year increase in the age of a subject, glycosylated hemoglobin increases by a scale of .04, and for every 50 mg/dL increase in chol, glycosylated hemoglobin increases by a scale of .008. F-statistic of 33.29 indicates that the null should be rejected. P-value of 4.251×10^{-14}, which is less than 0.05 indicating that at least one coefficient is not equal to 0, so the null is rejected. At alpha=0.05 level, the data provide evidence that there is a relationship between **glycosylated hemoglobin with age and cholesterol** (P-value 4.251×10^{-14}).

A multivariate linear regression is used to determine whether there is an association between the dependent explanatory variable glycosylated hemoglobin and the independent variables gender and cholesterol level. The relationship between cholesterol and gender with regard to glycosylated hemoglobin levels was tested against the claim of no association. Our adjusted R-square shows that only 5.82% of the variation found in the glycosylated hemoglobin is explained by gender and cholesterol level [glycosylated hemoglobin = 2.951+0.01*chol + 0.24*gender].The slope term in our model is saying that for the male subjects, glycosylated hemoglobin increases by 0.24mg/dl, and for every 50 mg/dLincrease in cholesterol.The F-statistic shows that the null hypothesis can be rejected and there is a relationship between glycosylated hemoglobin and gender and cholesterol. The P-value is 2.274×10^{-6} which is less than 0.05, therefore we reject the null hypothesis. We conclude that at alpha=0.05 level, our data provides significant evidence that there is a positive correlation between **glycosylated hemoglobin and gender and cholesterol** (P= 2.274).

A multivariate linear regression (shown in model 6) was used to assess the correlation between age and gender with regard to glycosylated hemoglobin levels. The relationship between age and gender with regard to glycosylated hemoglobin levels was tested against the claim of no association. The slope term in our model determined that for every subsequent 10 year increase in the age of a subject, glycosylated hemoglobin increases by 0.09 mg/dL. Model 6 additionally found glycosylated hemoglobin increases by 0.224 mg/dL for every male subject. An adjusted R-squared value of 0.1096 was calculated, inferring that 10.96% of the variance found in the dependent variable can be explained by the independent variable. The calculated F-statistic (25.73) shows that the null hypothesis can be rejected and there is a relationship between **glycosylated hemoglobin and agel and gender**; which is further with a significant p-value of 3.065×10^{-11}.

A multivariate linear regression was used to assess the relationship between age, gender, and cholesterol with regard to glycosylated hemoglobin levels. Glycosylated hemoglobin is the dependent variable, while age and cholesterol are independent variables. The relationship between glycosylated hemoglobin, age, gender, and cholesterol was tested against the claim of no association. The analysis resulted with equation [glycosylated hemoglobin = 1.8459 + 0.039age + 0.138 gender + 0.0088chol] where the slope explains that every 10 year increase in the age of a subject, glycosylated hemoglobin increases by a scale of 0.039, and for every male subject glycosylated hemoglobin increases by a scale of 0.138 and for every 50 mg/dL increase in cholesterol, glycosylated hemoglobin increases by a scale of 0.0088. Adjusted R-squared of

0.1372 indicates that 13.72 % of the variance found in the dependent variable (glycosylated hemoglobin) can be explained by the independent variables (age, gender, cholesterol). At 2.253×10^{-13}, the P-value is less than alpha 0.05, so the null is rejected. The resulting F-statistic is 22.31 indicating that the null hypothesis can be rejected and there is a relationship between glycosylated hemoglobin, age, cholesterol and gender. At alpha=0.05 level, the data provide evidence that there is a relationship between **glycosylated hemoglobin with age, gender, and cholesterol** (P-value 2.253×10^{-13}).

DISCUSSION

Glycosylated hemoglobin levels appear to be relatively steady across different age groups, but there is notable variation as the sample participants' age increases. From about 40 to 60 years old, data points potentially indicate a slight positive correlation between glycosylated hemoglobin levels and age. Data after age 60 shows a possibly negative trend with progressively fewer participants in successive age groups. The youngest study participants appear to be protected against elevated glycosylated hemoglobin levels. Correlation between glycosylated hemoglobin and cholesterol is low. Gender and glycosylated hemoglobin levels were not found to be correlated. Our findings helped us to understand how age, gender, and cholesterol levels are related to glycosylated hemoglobin levels.

Our findings can be applied towards predicting the possibility of developing diabetes in different age groups as well as early testing of individuals in the higher age group for cholesterol and glycosylated hemoglobin levels. The data provide evidence that age and cholesterol can explain the variance with glycosylated hemoglobin values. We can use these results to develop intervention strategies and assist in the early diagnosis of Type II diabetes that could potentially improve the quality of life for many people in America.

There are limitations that relate to the study conducted, impacting the findings of the study. The study was conducted focusing on African American adults in central Virginia. This may impact the generalizability of the results, as the sample was from a particular geographical area in only one racial group. The location of these subjects may have an impact on the findings, potentially impacting access to health care. There were several variables considered in this report, making it impossible to determine causation. Only a correlation can be determined between the variables involved (age, gender, glycosylated hemoglobin levels, and cholesterol levels). Cholesterol levels and glycosylated hemoglobin levels were determined using blood samples drawn from each individual participant. The blood was carried to a laboratory on ice, but it cannot be guaranteed that this was sufficient in keeping the samples at a specific temperature. If temperatures were not controlled properly, measurements may not be as accurate as possible. Also, this study only looked at three independent variables which is not enough to explain the onset of diabetes, hence we should look at a larger sample. Also, the sample size of 403 subjects is very small for clinical study purposes.

For future studies examining correlation between age, gender, glycosylated hemoglobin levels, and cholesterol levels, we suggest using a larger sample. To understand the impact on the African American population as a whole, it would be more effective to include participants from different geographical areas. This would account for differences in care, genetic factors, and so forth. It may be important to look at other factors that may correlate with glycosylated hemoglobin levels, indicating a diabetes diagnosis.

REFERENCES

Centers for Disease Control and Prevention. (2020). *National Diabetes Statistics Report, 2020*. CDC. Retrieved November 29, 2021, from https://www.cdc.gov/diabetes/library/features/diabetes-stat-report.html.

Drew, B., Rye, KA., Duffy, S. *et al.* The emerging role of HDL in glucose metabolism. *Nat Rev Endocrinol* **8,** 237–245 (2012). https://doi.org/10.1038/nrendo.2011.235

Gerstman, B. Burt. "Chapter 2." *Basic Biostatistics: Statistics for Public Health Practice*, Jones & Bartlett Learning, Burlington, MA, 2015.

Goyal, R., & Jailal, I. (2021, May 20). *Diabetes mellitus type 2*. StatPearls [Internet]. Retrieved September 23, 2021, from https://www.ncbi.nlm.nih.gov/books/NBK513253/.

Liu, H., Ma, Q., Shan, W., Xiang, G., & Xing, W. (2016) Association between glycated hemoglobin A1c levels with age and gender in Chinese adults with no prior diagnosis of diabetes mellitus.*Biomedical reports*, *4*(6), 737–740. doi:10.3892/br.2016.643

Rezigalla, Assad A. "Observational Study Designs: Synopsis for Selecting an Appropriate Study Design." *Cureus*, Cureus, 17 Jan. 2020, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6970097/.

Type 2 diabetes: What is it, causes, symptoms, risk factors & treatments. Cleveland Clinic. (2021, March 25). Retrieved September 23, 2021, from https://my.clevelandclinic.org/health/diseases/21501-type-2-diabetes.

U.S. Department of Health and Human Services. (2018). *Insulin Resistance & Prediabetes*. National Institute of Diabetes and Digestive and Kidney Diseases. Retrieved November 29, 2021, from https://www.niddk.nih.gov/health-information/diabetes/overview/what-is-diabetes/prediabetes-insulin-resistance.

APPENDIX:

Correlation:

compute the Pearson correlation between cholestrol and glycosylated hemoglobin
cor(diabetes$chol, diabetes$glycosylated hemoglobin, method = "pearson")
cor(diabetes$chol, diabetes$glycosylated hemoglobin, method = "spearman")
Figure 5

```
> ## compute the Pearson correlaton between cholestrol and glyhb
> cor(diabetes$chol, diabetes$glyhb, method = "pearson")
[1] 0.2449847
> cor(diabetes$chol, diabetes$glyhb, method = "spearman")
[1] 0.2367104
```

The correlation value for Pearson is .24 and .23 for spearman and signifies low correlation between cholesterol and glycosylated hemoglobin. (Ref: https://www.statisticssolutions.com/free-resources/directory-of-statistical-analyses/pearsons-correlation-coefficient/)

compute the Pearson correlation between age and glycosylated hemoglobin
cor(diabetes$age, diabetes$glycosylated hemoglobin, method = "pearson")
cor(diabetes$age, diabetes$glycosylated hemoglobin, method = "spearman")
Figure 6

```
> ## compute the Pearson correlaton between age and glyhb
> cor(diabetes$age, diabetes$glyhb, method = "pearson")
[1] 0.3367019
> cor(diabetes$age, diabetes$glyhb, method = "spearman")
[1] 0.4306941
```

The correlation value for pearson is 0.3367 and spearman is 0.4306941

compute the point biserial correlation between gender and glycosylated hemoglobin.We will recode gender as it is categorical now
diabetes$gender.Dummy<-ifelse(diabetes$gender=="male",1,0)
colnames(diabetes)
summary(diabetes)
Figure 7

```
        Pearson's product-moment correlation

data:  diabetes$gender.Dummy and diabetes$glyhb
t = 1.0003, df = 401, p-value = 0.3178
alternative hypothesis: true correlation is not equal to 0
95 percent confidence interval:
 -0.04803002  0.14685971
sample estimates:
       cor
0.04988975
```

The correlation value is 0.049 and signifies almost no correlation between gender and glycosylated hemoglobin (Ref: https://www.statisticssolutions.com/free-resources/directory-of-statistical-analyses/pearsons-correlation-coefficient/)

```
> t.test( diabetes1$glyhb ~ diabetes1$gender, paired = FALSE, var.equal = FALSE )

        Welch Two Sample t-test

data:  diabetes1$glyhb by diabetes1$gender
t = -0.98363, df = 338.78, p-value = 0.326
alternative hypothesis: true difference in means between group female and group male is not equal to 0
95 percent confidence interval:
 -0.6685105  0.2227972
sample estimates:
mean in group female   mean in group male
            5.499274             5.722130
```

Regression:

mylogit1 = lm(glyhb~age, data=diabetes)
summary(mylogit1)

```
Call:
lm(formula = glyhb ~ age, data = diabetes)

Residuals:
    Min      1Q  Median      3Q     Max
-3.2328 -1.1173 -0.4495  0.1671 10.2372

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 3.458458   0.315567  10.960  < 2e-16 ***
age         0.045554   0.006362   7.161 3.87e-12 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.081 on 401 degrees of freedom
Multiple R-squared:  0.1134,    Adjusted R-squared:  0.1112
F-statistic: 51.27 on 1 and 401 DF,  p-value: 3.867e-12
```

We see that for dependent variable glycosylated hemoglobin and independent variable age,
F statistic = 51.27
Degrees of freedom = 401,
P value : 3.846×10^{-12}

mylogit2 = lm(glyhb~gender.Dummy, data=diabetes)
summary(mylogit2)
F statistics = 1.001 , degrees of freedom = 401,
P-values : 3.867×10^{-12}

```
> summary(mylogit2)

Call:
lm(formula = glyhb ~ gender.Dummy, data = diabetes)

Residuals:
    Min      1Q  Median      3Q     Max
-3.0421 -1.1893 -0.7321  0.0857 10.3879

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)     5.4993     0.1443   38.12   <2e-16 ***
gender.Dummy1   0.2229     0.2228    1.00    0.318
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.207 on 401 degrees of freedom
Multiple R-squared:  0.002489,  Adjusted R-squared:  1.428e-06
F-statistic: 1.001 on 1 and 401 DF,  p-value: 0.3178
```

mylogit3 = lm(glyhb~chol, data=diabetes)
summary(mylogit3)
At alpha= 0.05, F-statistic = 25.6
Degree of freedom = 401
P-value= 0.3178

```
> mylogit3 = lm(glyhb~chol, data=diabetes)
> summary(mylogit3)

Call:
lm(formula = glyhb ~ chol, data = diabetes)

Residuals:
    Min      1Q  Median      3Q     Max
-2.7698 -1.1972 -0.6343  0.1704  9.4435

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 3.061139   0.511569   5.984 4.84e-09 ***
chol        0.012180   0.002407   5.060 6.40e-07 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.142 on 401 degrees of freedom
Multiple R-squared:  0.06002,   Adjusted R-squared:  0.05767
F-statistic:  25.6 on 1 and 401 DF,  p-value: 6.397e-07
```

mylogit4 = lm(glyhb~chol+age, data=diabetes)
summary(mylogit4)
At alpha = 0.05,
F statistic = 33.29,
degrees of freedom = 400
P-value = 4.251×10^{-14}

```
> mylogit4 = lm(glyhb~chol+age, data=diabetes)
> summary(mylogit4)

Call:
lm(formula = glyhb ~ chol + age, data = diabetes)

Residuals:
    Min      1Q  Median      3Q     Max
-3.8142 -1.1482 -0.3937  0.2657  9.4997

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept) 1.899262   0.523724   3.626 0.000324 ***
chol        0.008753   0.002367   3.698 0.000247 ***
age         0.040001   0.006441   6.210 1.33e-09 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.049 on 400 degrees of freedom
Multiple R-squared:  0.1427,    Adjusted R-squared:  0.1384
F-statistic: 33.29 on 2 and 400 DF,  p-value: 4.251e-14
```

mylogit5 = lm(glyhb~chol+gender.Dummy, data=diabetes)
summary(mylogit5)
At alpha = 0.05,
F statistic = 13.43

Degree of freedom= 400
P-value= 2.274×10^{-6}
glycosylated hemoglobin=

```
> mylogit5 = lm(glyhb~chol+gender.Dummy, data=diabetes)
> summary(mylogit5)

Call:
lm(formula = glyhb ~ chol + gender.Dummy, data = diabetes)

Residuals:
    Min      1Q  Median      3Q     Max
-2.9118 -1.1773 -0.6095  0.1110  9.5072

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  2.951783   0.520822   5.668 2.78e-08 ***
chol         0.012222   0.002407   5.078 5.85e-07 ***
gender.Dummy1 0.240045   0.216237   1.110    0.268
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.142 on 400 degrees of freedom
Multiple R-squared:  0.0629,    Adjusted R-squared:  0.05822
F-statistic: 13.43 on 2 and 400 DF,  p-value: 2.274e-06
```

mylogit6 = lm(glyhb~age+gender.Dummy, data=diabetes)
summary(mylogit6)

```
> mylogit6 = lm(glyhb~age+gender.Dummy, data=diabetes)
> summary(mylogit6)

Call:
lm(formula = glyhb ~ age + gender.Dummy, data = diabetes)

Residuals:
    Min      1Q  Median      3Q     Max
-3.1749 -1.1178 -0.4573  0.1716 10.1731

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  3.422888   0.322755  10.605  < 2e-16 ***
age          0.045303   0.006385   7.095  5.9e-12 ***
gender.Dummy1 0.112905   0.210803   0.536    0.593
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.083 on 400 degrees of freedom
Multiple R-squared:  0.114,     Adjusted R-squared:  0.1096
F-statistic: 25.73 on 2 and 400 DF,  p-value: 3.065e-11
```

mylogit7 = lm(glyhb~age+gender.Dummy+chol, data=diabetes)
summary(mylogit7)

```
> mylogit7 = lm(glyhb~age+gender.Dummy+chol, data=diabetes)
> summary(mylogit7)

Call:
lm(formula = glyhb ~ age + gender.Dummy + chol, data = diabetes)

Residuals:
    Min      1Q  Median      3Q     Max
-3.7464 -1.1319 -0.4229  0.2123  9.4163

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)   1.845920   0.530108   3.482 0.000552 ***
age           0.039658   0.006466   6.133 2.07e-09 ***
gender.Dummy1 0.138992   0.207625   0.669 0.503605
chol          0.008807   0.002370   3.716 0.000231 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.05 on 399 degrees of freedom
Multiple R-squared:  0.1436,    Adjusted R-squared:  0.1372
F-statistic: 22.31 on 3 and 399 DF,  p-value: 2.253e-13
```