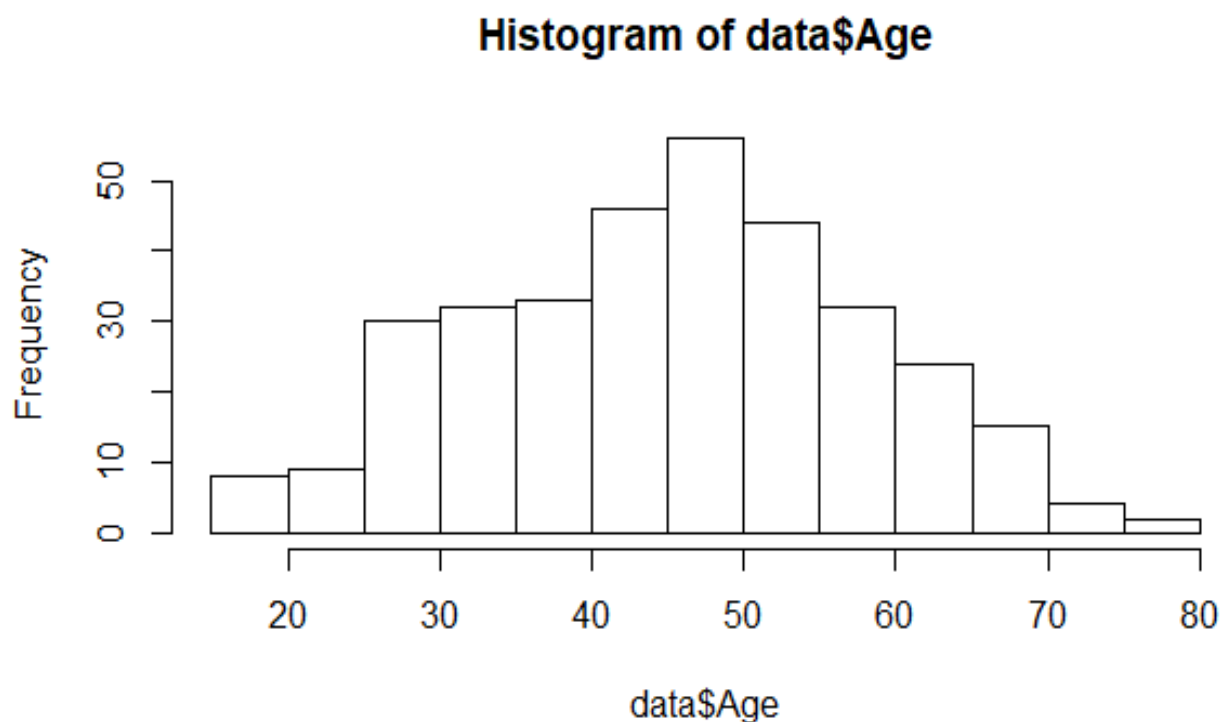


Introduction

A pharmaceutical company in Ireland have proposed a new drug that could reduce a patient's risk of cardiovascular disease. Researchers at St Vincent's Hospital Dublin have been selected to test this claim. They got 335 applicants with known high blood pressure and record their various measurements. Doctors assess each applicant and label them as at Low, Medium, or High risk of cardiovascular disease. The applicants are randomly selected to a group: control or placebo, so that the effects of the new drug can be significantly tested. After 6 months on this new drug, the patient's risk is reassessed and recorded. The body of this report is focused on the use of statistical methods to analyse the findings of these assessments. Most of the results stated in the report were obtained using the software RStudio.

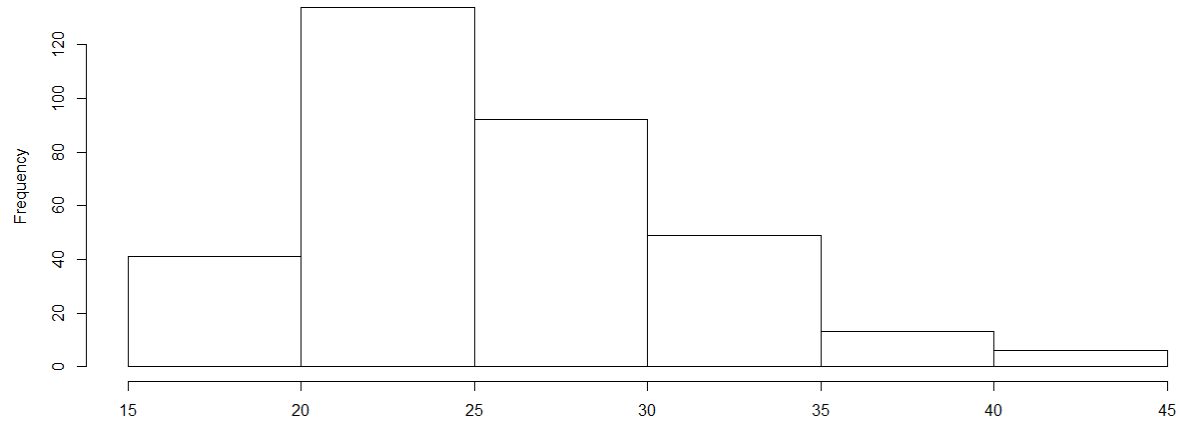
1.Exploratory Analysis

The number of patients in the Control group was found to be 182, with 153 in the Placebo group. A histogram was obtained in RStudio to describe the distribution of age groups in the sample and is



BMI is w^2/h where w is weight in kg and h in height measured in metres. This quantity was also calculated for each of the applicants of the sample, a histogram provided to describe its distribution in the sample, and the mean BMI was calculated for male and female applicants. This histogram is pictured below in Figure 2 and it was obtained that mean BMI of the males in the sample was 24.37975 and the mean BMI of the females was 26.72904.

Histogram of data\$BMI



County	Mean Weight (kg)
<i>Kilkenny</i>	70.99333
<i>Kerry</i>	75.80833
<i>Clare</i>	71.58333
<i>Sligo</i>	68.51111
<i>Cavan</i>	67.84
<i>Louth</i>	70.87333
<i>Galway</i>	71.55
<i>Cork</i>	71.74118
<i>Roscommon</i>	64.49375
<i>Wexford</i>	63.66429
<i>Waterford</i>	78.27273
<i>Longford</i>	67.22
<i>Carlow</i>	63.12143
<i>Dublin</i>	65.78889
<i>Westmeath</i>	66.47778
<i>Tipperary</i>	69.23125
<i>Wicklow</i>	77.33077
<i>Limerick</i>	71.56923
<i>Mayo</i>	65.23529
<i>Monaghan</i>	71.88824
<i>Donegal</i>	67.35455
<i>Offaly</i>	69.96667
<i>Laois</i>	68.89167
<i>Kildare</i>	65.91429
<i>Meath</i>	64.25
<i>Leitrim</i>	74.89

Table 2: Counties and their respective mean weights

It was found that 45 females were members of Blood Group O and that 26.06061% of the males in the study belonged to Blood Group B. The total cholesterol difference from before the study was 72, i.e. the total cholesterol level of the patients decreased by 72. We found Carlo w has the lowest mean with 63.12143 and Waterford highest mean 78.27273.

2.Hypothesis Testing

I used a Chi Squared Association test to see if there is a connection between gender and cardiovascular disease risk.

Cardiovascular Risk	High	Low	Medium
Female	61	44	65
Male	56	65	44

Table 3:

The test statistic obtained was $\chi^2 = 8.2326$, with 2 degrees of freedom. critical value was 5.991.

so, at the 5% significant level, the test statistic does not fall within the rejection region, and so we fail to reject H_0 . Therefore, we conclude that there is not a difference in the risk of contracting cardiovascular disease between males and females.

We use a 2 sample T test comparing the weights in the Control and Placebo groups.

The test statistic obtained was $z = -0.74152$ The corresponding p-value was 0.2295, which is greater than alpha. Therefore, as the p - value is greater than the significance level, we fail to reject the null hypothesis. Therefore, we conclude that there is not enough evidence to support the claim that there is a difference in weight between the control and placebo groups.

Again, we use a 2 sample T test, this time, comparing the cholesterol difference in the Treatment and Placebo groups.

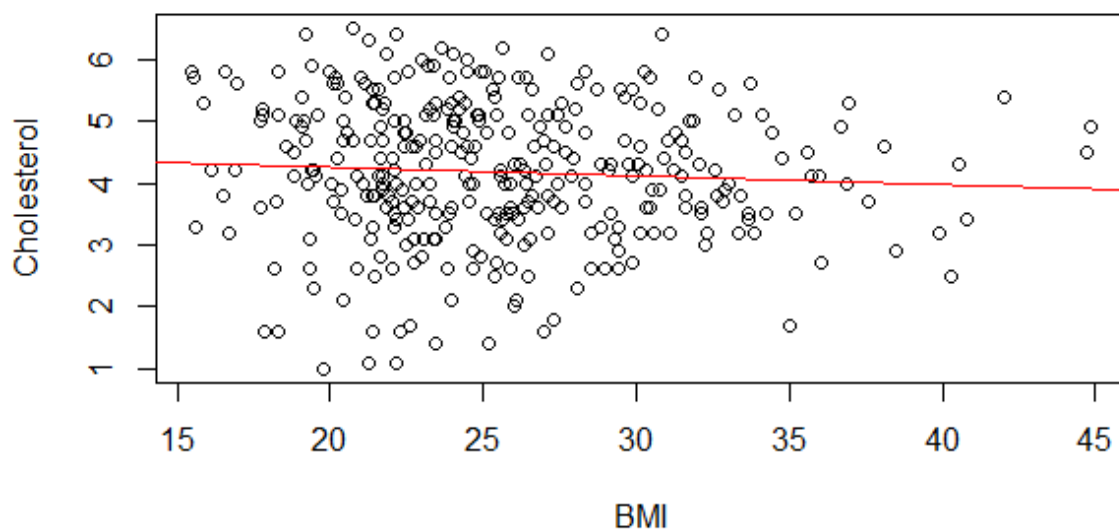
The test statistic obtained was -0.55574 The corresponding p-value was 0.5788, which is again greater than alpha. Thus, we again fail to reject the null hypothesis H_0 . So, we conclude that there is not enough evidence to support the claim that the new drug reduces cholesterol compared to the placebo drug.

3.Model Fitting

The correlation coefficient was computed between BMI and the initial cholesterol levels before the study of cholesterol1 and it was found that $r = -0.06706034$. This suggests a very weak positive relationship correlation between the two variables. regression line fitted to the data on a scatter plot is included below, with the equation of the regression line calculated to be $\hat{y} = 4.55044 - 0.01412x$. With parameter values $a = y\text{-intercept} = 4.55044$, and $b = \text{slope} = -0.01412$ The value of the coefficient of determination was found to be $R^2 = 0.004497$ which is very small and so suggests that the regression model is not a particularly good fit to the data.

We now perform a model utility test to assess the utility of the regression model for this data. We test the hypothesis that

at the 5% significance level, where beta is the slope parameter of the line in the population. The test statistic was computed in RStudio to be $t = -1.226$ with a corresponding p-value of 0.2209. As the p-value alpha, we fail to reject the null hypothesis H_0 , and conclude that there is insufficient evidence that BMI is useful at describing the variation in cholesterol levels through a linear relation.



4.Secondary Analysis

Blood Group	Mean Height (cm)
<i>A</i>	165.618
<i>AB</i>	164.1274
<i>B</i>	165.2027
<i>O</i>	165.0993

Table 4: Mean heights for each level of the Blood Group factor variable

We now use one-way ANOVA (Analysis of Variance) at the 5% significance level to test the hypothesis that the mean height was different across all the different blood groups.

The test statistic was computed in RStudio and it was found that $F = 0.638$, with a corresponding p-value of 0.591, and with 3 degrees of freedom. As the p-value alpha, we fail to reject the null hypothesis H_0 and conclude that there is insufficient evidence to claim that any of the means differ from each other.

Conclusion

Analysis carried out in this report offered key insights into characteristics of the sample. The connection between some physical variables the patients were assessed under and the usefulness of the new drug in the decrease of a patient's cholesterol level. Some key results include that firstly, there wasn't enough evidence to suggest a connotation between gender and cardiovascular disease risk. Also, there did not appear to be a linear relationship between Body Mass Index and cholesterol levels in a patient. Most, particularly of all however, it was found that the new drug proposed by the pharmaceutical company did not consistently lower cholesterol levels when compared to a placebo group. It could be believed, as such, that the new drug is ineffective, and the pharmaceutical company may want to consider any other alternatives.