Comparing the Shock-Index based on Cholesterol Level or Diabetic Status

Carolin Poschen

Introduction

This report analyses the influence of the cholesterol level for women as well as the diabetic status for all genders on the heart health. The analysis is based on the Framingham Heart Study[1] which gathers information like age and gender as well as health data like BMI, cholesterol level or if the participant takes Blood Pressure medications.

For this study, we create a new variable *Shock-Index* for each participant which is the person's heart rate divided by the person's systolic blood pressure. We compare this value for women based on their cholesterol level as well as for all participants based on whether or not they have diabetes.

For both groups, women split on their cholesterol level as well as all participants based on their diabetic status, we conduct a t-test to test if there is a statistically significant difference between the subgroups.

Different Shock-Indices for Women based on their Cholesterol Level

We first split all women in this study, for which the cholesterol level was measured, in the subgroups *high* or *low*. *High* (or *low*) is here defined as a cholesterol level above (or below) the average cholesterol level for all women in this study.

With an average cholesterol level of 239.62 mg/dL, we get 1059 women with a low and 976 women with a high cholesterol level. Figure 1 shows the Box-Plot[2] of the calculated Shock-Index for both groups.

We calculate the average shock-index for each group (low Chol.: $\bar{x_l}$, high Chol.: $\bar{x_h}$) and the estimate difference between means from the group (x_{diff} = $\bar{x_h}$ - $\bar{x_l}$). Using sampling with replacement, we gather 10,000 bootstrap samples for each group and calculate a 90% confidence interval (CI) for the true difference in means from these samples. The results are displayed in table 1. We assume that the data is from a random sample and therefor is independent. With more than 30 subjects in each subgroup, we assume the data comes from a normal distribution.

Table 1: Summary Statistics Shock-Index for Women based on Chol. Level

Sample Mean $\bar{x_l}$	0.61004
Sample Mean $\bar{x_h}$	0.57232
x_{diff}	-0.03772
90% CI	(-0.0424 -0.0331)

Table 2: T-Test results

Test Statistic	**Value**
t-statistic	-7.0175
df	2026.7
p-value	3.067e-12

We conduct a *t-test* for differences in means with a significance level of $\alpha = 0.05$. The null hypothesis states that there is no difference between means of each subgroup,

[1] https://en.wikipedia.org/wiki/Framingham_Heart_Study

[2] https://en.wikipedia.org/wiki/Box$_p$*lot*

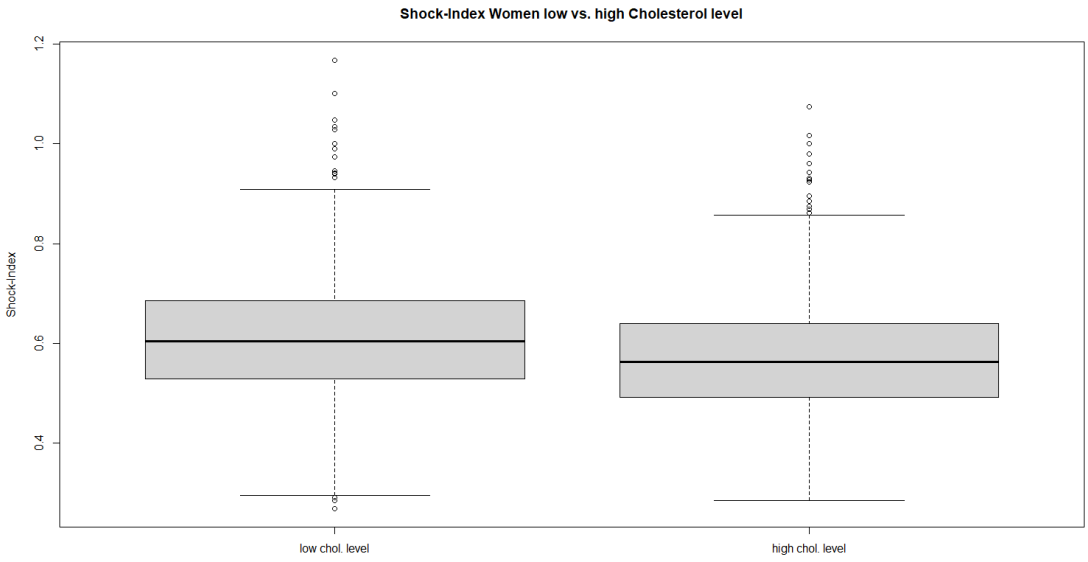

Figure 1: Comparison Shock-Index for Women with high or low Cholesterol level

while the alternative hypothesis states that there is some difference. The test results are listed in table 2. With a *t-statistic* of -7.0175 and a *p-value* of 3.067e-12 we reject the null hypothesis that there is no difference between the means of both groups and conclude that we have strong statistically evidence that there is a difference between both subgroups. With a confidence interval of (-0.0424 -0.0331), we conclude that a low cholesterol level leads to a higher shock-index for women in this population.

Different Shock-Indices based on Diabetic Status

We split all participants in two subgroups, the *diabetic group* consists of participants that have diabetes, while the *non-diabetic group* consists of participants that do not have diabetes. We get 99 subjects in the diabetic group and 3359 subjects in the non-diabetic group. Figure 2 shows the Box-Plot for the shock-index for both subgroups.

We calculate the average shock-index for each subgroup (diabetic group: $\bar{x_d}$, non-diabetic group $\bar{x_{nd}}$) and the estimate difference between means from the group ($x_{diff} = \bar{x_d}$ - $\bar{x_{nd}}$). Using sampling with replacement, we gather 10,000 bootstrap samples for each subgroup and calculate a 90% confidence interval (CI) for the true difference in means from these samples. The results are displayed in table 3. We assume that the data is from a random sample and therefor is independent. With more than 30 subjects in each subgroup, we assume the data comes from a normal distribution.

We conduct a *t-test* for differences in means with a significance level of $\alpha = 0.05$. The null hypothesis states that there is no difference between means of each subgroup, while the alternative hypothesis states that there is some difference. The test results

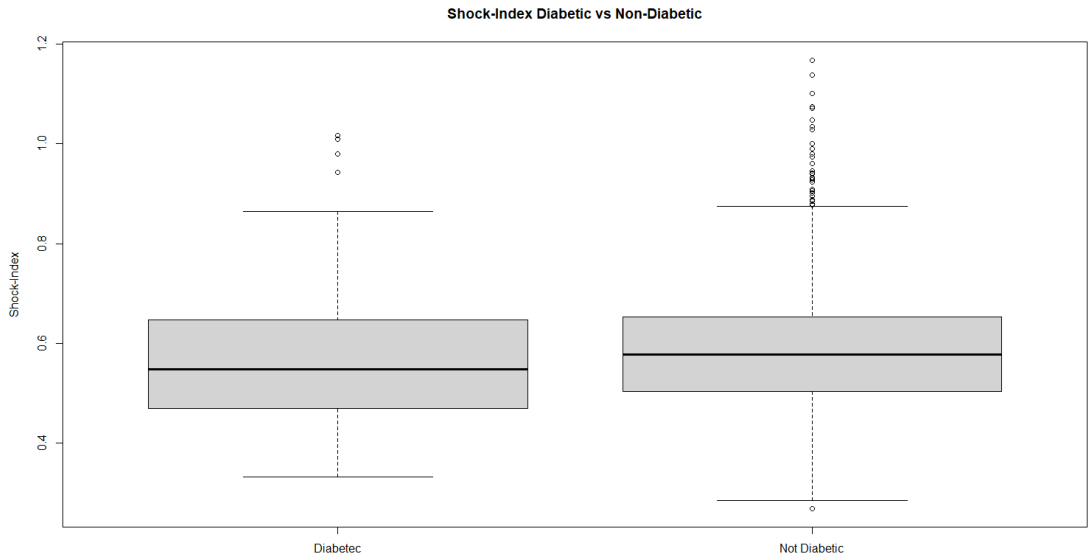

Figure 2: Overview Shock-Index for Participants with and without Diabetes

Table 3: Summary Statistics Shock-Index based on Diabetic Status

Sample Mean $\bar{x_d}$	0.5705
Sample Mean $\bar{x_{nd}}$	0.5843
x_{diff}	-0.01384
CI	(-0.01887 -0.00861)

Table 4: T-Test results

Test Statistic	**Value**
t-statistic	-0.9242
df	101.33
p-value	0.3576

are listed in table 4. With a *t-statistic* of -0.9242 and a *p-value* of 0.3576 we fail reject the null hypothesis that there is no difference between the means of both groups. While the confidence interval of (-0.01887 -0.00861) does not include 0, our test results leave us with inconclusive results.

Discussion

While our results show that there is a difference in means for the shock-index based on their cholesterol level, our results were inconclusive for the diabetic status. While the subset of data for women based on the cholesterol level is nearly balanced, the subset for the diabetic status is very unbalanced. We therefor suggest further studies for the influence of the diabetic status which should also investigate more influences instead of just the shock-index as the person's heart rate divided by the person's blood pressure.