A Statistical Study to Identify the Risk Factors of Heart Attack

Zoha Fatima, Itrat Batool Naqvi* and Sharoon Hanook
Department of Statistics, Forman Christian College, Lahore, Pakistan

Abstract

A statistical study has been conducted to identify the risk factors of heart attack. The study design used in this research is an observational cross-sectional. A semi-structured questionnaire was designed and surveyed consisting of 25 questions which were filled from 246 patients from two hospitals ‘Gulab Devi’ and ‘Jinnah Hospital’ Lahore, Pakistan. Respondents were asked questions regarding some of the possible reasons that may cause heart attack.

Out of 246 patients, 123 were cases (people who had a heart attack) and remaining 123 were control (people who only had chest pain). We took 123 patients in each group because we needed comparison. Spss and R SOFTWARE were used to determine results of this research. By using univariate, bivariate and multivariate analysis it was observed that the significant factors from model are diabetes blood pressure, sweating, heart attack before, age, severity of pain, medication and pressure of the work.

Keywords: Statistical study; Heart attack; Patients; Risk factors

Introduction

Loss of a loved one can never be forgotten. Forgetting the pain of losing someone is almost as difficult as forgetting the person. With increased death rate in our society caused by myocardial infarction also known as heart attack, researchers are finding ways to get to the root cause of it so that more and more lives can be saved. A heart attack is a confusing, frightening and stressful experience that shakes us to the core of our existence. Some of the causes of heart attack are coronary heart disease, spasm of coronary artery and spontaneous coronary dissection.

Jarrah et al. [1] explored the association between the serum level of vitamin D and ischemic heart disease. Body mass index and dyslipidemia showed a significant association and other risk factors for ischemic heart disease which included smoking, diabetes, hypertension, angina showed no significant association. Pusceddu et al. [2] examined the relation between telomere length and mortality in the Ludwigsfalen Risk and Cardiovascular health study. Author concluded from this study that short telomere increased the risk of cardiovascular disease and all-cause mortality where age plays an important covariate. Somuncu et al. [3] assessed the relationship between aspirin resistance and carotid in young patients with ST-segment elevated myocardial infarction (STEMI). Physicians should be careful about AR (Aspirin Resistance) in patients with STEMI who have carotid plaque and thicker carotid intima-media thickness (CIMT).

Sedky et al. [4] assessed that genetic variants of CYP2R1 are key regulators of serum vitamin D levels and incidence of myocardial infarction (MI) in middle aged Egyptians. Author concluded that genetic variants of CYP2R1 are key determinants of serum 25OHD levels and are highly associated with the risk of MI.

Materials and Methods

Primary data will be collected through a semi-structured questionnaire. The univariate, bivariate and multivariate analysis has been applied on the data. In univariate analysis the single variable has been uniquely examined by means of graph and percentages. In bivariate analysis, the relationships among variables were explored and in multivariate analysis, try to identify risk factors of the heart attack along with their corresponding probabilities according to the survey conducted.

Determination of sample size

\[ n = \frac{p(1-p)z^2}{e^2} \]

where,

\[ p=\text{prevalence rate of the patients}=0.8, \]

[1]

\[ z=\text{tabulated value from normal distribution at 0.05 level of significance}=1.96, \]

\[ e=\text{difference between estimated and actual value that is absolute precision that is required on either side of the proportion } p=0.05. \]

Sampling plan

Target Population: All patients with the complain of pain in their chests.

Sampled Population: Patients that came to Gulab Devi and Jinnah with complain of chest pain.

Sample Size: 250 patients with a prevalence rate of 0.8.

Sampling tool: A semi-structured questionnaire.

Sampling Method: Convenient sampling (Five weeks of cohort study conducted).

Logistic regression model

When dependent variable is dichotomous logistic regression can be used for analysis purpose.

The Logistic Regression model is given as:

\[ F(z) = \text{Prob(event)} = \left( \frac{e^z}{1+e^z} \right) \]

*Corresponding author: Itrat Batool Naqvi, Department of Statistics, FC College, Lahore, Pakistan, Tel: 9242992315818; E-mail: itratnaqvi@fccollege.edu.pk

Received March 08, 2019; Accepted June 10, 2019; Published June 17, 2019


Copyright: © 2019 Fatima Z, et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.
The significant risk factors of heart attack at 5 and 10% level of significance are:

- Diabetes,
- Blood Pressure,
- Sweating,
- Heart attack before,
- Age,
- Severity of pain,
- Medication,
- Pressure of the work.

The fitted model is:

\[ P(\text{Heart Attack}) = \log\left(\frac{\pi}{1-\pi}\right) \]

\[ Z = \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 \]

\[ Z = -4.776 + 0.025 X_1 + 1.184 X_2 + 1.927 X_3 + 0.846 X_4 + 1.021 X_5 + 1.395 X_6 + 0.560 X_7 + 1.100 X_8 \]

### Interpretation

From Table 5 the risk of Heart attack is 2.331 times more if the patient is diabetic, 1.751 times more if the case has a blood pressure problem, 3.268 times more if during the chest pain he is feeling sweating, 6.870 times more in the cases who already had at least one time heart attack, 1.025 times more depending on the age, 2.777 times more depending on the severity of pain, 4.035 times more if the patient is on medication and 3.005 times more if he is having pressure of the work.

\[ P(\text{Heart Attack}) = \log\left(\frac{\pi}{1-\pi}\right) \]

\[ Z = \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \beta_4 X_4 \]

\[ Z = -4.776 + 0.025 X_1 + 1.184 X_2 + 1.927 X_3 + 0.846 X_4 + 1.021 X_5 + 1.395 X_6 + 0.560 X_7 + 1.100 X_8 \]

\[ X_1 = \text{Age}, \]

### Results

Univariate analysis, in this paper the histogram for the ages of case(s) and control(s) are presented, although the rest univariate results (charts, percentages) can be provided upon request of the readers [6-10].

### Histogram for cases

**Bivariate table for cases:** In Tables 1 and 2 we are discussing the relationship between two variables along with their p-value for the significant variables, in the multivariate analysis, only. The p value here corresponding to chi-square test of association (a bivariate statistic tool) has been automatically generated in SPSS (statistical packages of social sciences).

It can be observed from the above table that diabetes and cases taking any kind of medicine are positively associated at 5% level of significance. It can be concluded that if a person is having diabetes then he’s also taking some kind of medicine. Similarly it is same for blood pressure and severity of pain.

It can be observed from the above table that diabetes and controls taking any kind of medicine are positively associated at 5% level of significance. It can be concluded that if a person is having diabetes then he’s also taking some kind of medicine. Similarly it is same for blood pressure and severity of pain (Table 3).

### Multivariate analysis

In multivariate analysis, the risk factors of the heart attack along with their corresponding probabilities are identified according to the survey conducted.

The value of model chi-square is 146.519 with p value is 0.000 which shows a high significance and best fit. Therefore, we are 95% confident that fitted model is appropriate.

From the above classification table, we can observe that 96 (78%) of the respondent (control) not suffering from heart disease were correctly predicted whereas 104 respondent (case) 84.6% suffering from heart disease were correctly predicted. Overall 81% of 246 respondents were correctly classified (Table 4).

### Table 1: Bivariate table for cases.

<table>
<thead>
<tr>
<th>Variables</th>
<th>Chi-square association</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diabetes and Medication</td>
<td>12.135</td>
<td>0.000</td>
</tr>
<tr>
<td>Diabetes and Blood pressure</td>
<td>8.972</td>
<td>0.011</td>
</tr>
<tr>
<td>Blood pressure and Severity of pain</td>
<td>11.379</td>
<td>0.023</td>
</tr>
</tbody>
</table>
\( X_2 = \text{Sweating}, \)
\( X_3 = \text{Heart attack before}, \)
\( X_4 = \text{Diabetes}, \)
\( X_5 = \text{Severity of pain}, \)
\( X_6 = \text{Medication}, \)
\( X_7 = \text{Blood pressure}, \)
\( X_8 = \text{Pressure of the work}. \)

Now the model probability is determined if:
\( X_1 = \text{age}=50 \text{ y}, \)
\( X_2 = \text{sweating}=1 \text{ (yes)}, \)
\( X_3 = \text{heart attack before}=1 \text{ (yes)}, \)
\( X_4 = \text{diabetes}=1 \text{ (yes)}, \)
\( X_5 = \text{severity of pain}=3, \) here 3 means a high pain intensity which we have coded in this research as follows 1=0-2, 2=3-5, 3=6-8, 4=more than eight,
\( X_6 = \text{medication}=1 \text{ (yes)}, \)
\( X_7 = \text{blood pressure}=2, \) here 2 means a high blood pressure and we have coded in this research as follows 0=normal, 1=low and 2=high,
\( X_8 = \text{pressure of the work}=1 \text{ (yes)}. \)

Now we would like to predict the probability from the above model that the person of age 40 years with no blood pressure complaint and the pain level is mild keeping, the rest variables as stated above, then,

For \( Z=3.697 \)

\( P (\text{Heart Attack})=97.6\%. \)

Discussion

The purpose of this research was to identify the risk factors of heart attack by collecting data of 246 patients, out of which 123 were cases and the remaining 123 were control. Cases were the patients who had the chest pain and later suffered from heart attack and control were the group of people who came to hospital with the problem of chest pain but didn’t have the heart attack. Univariate, bivariate and multivariate analysis was done on the data of these patients (Figures 1 and 2).

In univariate analysis, single variable by means of graph and percentages were observed. Histogram and pie charts were drawn with the help of R and SPSS. In bivariate analysis, the relationship among variables was explored with the help of Pearson correlation coefficients, Spearman rank and Chi-square correlation. From bivariate analysis of cases, we concluded that variables associated with diabetes at 5% level of significance are obesity, patient taking any kind of medicine, patient having some other type of disease and blood pressure. Similarly, we observed that variables associated with blood pressure are severity of pain, smoking and other diseases. Variables that were associated with sweating during chest pain were maternal side, paternal side and smoking. Variables that were associated with heart attack before were paternal side, pressure of the work and gender. Variables that were associated with severity of pain were smoking, pain radiating towards, gender of the patient and variables associated with pressure of the work are genders of the patient, heart attack before, restlessness and other diseases.

From the control, we concluded that variables that are associated with diabetes at 5% level of significance are obesity, restlessness, pressure...
of the work, other diseases and patient taking any kind of medicine. Variables associated with blood pressure were restlessness, pressure of the work, problem of irregular pulse rate and kidney problem. There was a positive association between sweating and restlessness. Another association was seen in kidney problem and heart attack before.

Conclusion

In multivariate analysis, high-risk factors of the heart attack were identified along with their corresponding probabilities according to the survey conducted at 5% and 10% level of significance. We concluded that the risk of Heart attack is 2.331 times more if the patient is diabetic, 1.751 times more if the case has a blood pressure problem, 3.268 times more if during the chest pain he is feeling sweating, 6.870 times more in the cases who already had at least one time heart attack, 1.025 times more depending on the age, 2.777 times more depending on the severity of pain, 4.035 times more if the patient is on medication and 3.005 times more if he is having pressure of the work. The significant factors from the predicted model are diabetes blood pressure, sweating, heart attack before, age, severity of pain, medication and pressure of the work.

References