Journal of Biometrics & Biostatistics
ISSN: 2155-6180
Fatima et al., J Biom Biostat 2019, 10:2
Research Article Open Access

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 Hosptial' 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 Ludwigshafen Risk and Cardiovascular health study. Author concluded from this study that short telomere length 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 *CYP2R*1 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

$$\mathbf{n} = \frac{p(1-p)z^2}{e^2} where,$$

p=prevalence rate of the patients=0.8,

z=tabulated value from normal distribution at 0.05 level of significance=1.96,

e=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:

$$F(z) = \text{Prob}(\text{event}) = \left(\frac{e^z}{1-e^z}\right)$$

***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

Citation: Fatima Z, Naqvi IB, Hanook S (2019) A Statistical Study to Identify the Risk Factors of Heart Attack. J Biom Biostat 10: 430.

Where $z=\beta_0+\beta_1X_1+\beta_2X_2+\ldots$ for multiple model and $\beta_0, \beta_1, \beta_2, \ldots$ are coefficients and xi are the independent variables.

$\text{Prob}(\text{Event}) = log_e(\frac{p}{1-p})$, p is the proportion of the event of "yes" or "no" and e is the base of natural logarithms Hanif et al. [5].

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).

Variables	Chi-square association	p-value
Diabetes and Medication	12.135	0.000
Diabetes and Blood pressure	8.972	0.011
Blood pressure and Severity of pain	11.379	0.023

Table 1: Bivariate table for cases.

Variables	Chi-square association	p-value
Diabetes and medication	26.702	0.000
Sweating and restlessness	15.035	0.000
Blood pressure and pressure of the work	7.213	0.027

Table 2: Bivariate table for controls.

	Chi-square	df	Sig.
Model	146.519	23	0.000

Table 3: Omnibus Tests of Model Coefficients.

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:

$$\text{P}\left(\text{Heart Attack}\right) = Log\ (\frac{\pi}{1-\pi})$$

$$Z=\beta_0+\beta_1X_1+\beta_2X_2+\beta_3X_3+\beta_kX_k$$

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

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.

$$\text{P}\left(\text{Heart Attack}\right) = Log\ (\frac{\pi}{1-\pi})$$

$$Z=\beta_0+\beta_1X_1+\beta_2X_2+\beta_3X_3+\beta_kX_k$$

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

X_1=Age,

Observed (heart attack)	Predicted		
	Heart Attack		Percentage Correct
	No	Yes	
No	96	27	78
Yes	19	104	84.6
Overall Percentage			81.3

Table 4: Classification table.

Variables	Co-efficient	Odds Ratio	Significance
Diabetes	0.846	2.331	0.093
Blood Pressure	0.560	1.751	0.011
Sweating	1.184	3.268	0.011
Heart attack before	1.927	6.870	0.013
Age	0.025	1.025	0.054
Severity of pain	1.021	2.777	0.000
Medication	1.395	4.035	0.033
Pressure of the work	1.100	3.005	0.012
Constant	-4.776	0.008	0.005

Table 5: Table of variable in the equation.

X_2=Sweating,

X_3=Heart attack before,

X_4=Diabetes,

X_5=Severity of pain,

X_6=Medication,

X_7=Blood pressure,

X_8=Pressure of the work.

Now the model probability is determined if:

X_1=age=50 y,

X_2=sweating=1 (yes), the code 1 represents that a person was sweating at the time of pain and likewise,

X_3=heart attack before=1 (yes),

X_4=diabetes=1 (yes),

X_5=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=medication=1 (yes),

X_7=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=pressure of the work=1 (yes).

$$P\left(\text{Heart Attack}\right) = Log\left(\frac{\pi}{1-\pi}\right),$$

$$Z=\beta_0+\beta_1X_1+\beta_2X_2+\beta_3X_3+\beta_kX_k,$$

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

P (Heart Attack)=0.9992,

It seems that with an age of 50 years and having the above stated variables with their respected codes have 99.92 risk of heart attack.

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

For Z=3.697

P (Heart Attack)=97.6%.

The probability of his or her risk decreased by approximately 2% that 97.6%. Hence we cannot deny the risk of the other factors sweating, heart attack before, diabetes, medication, and pressure of the work.

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

Figure 1: From the above histogram, we can observe that most of the respondents (cases) are falling within the age limit of 30 to 70 years.

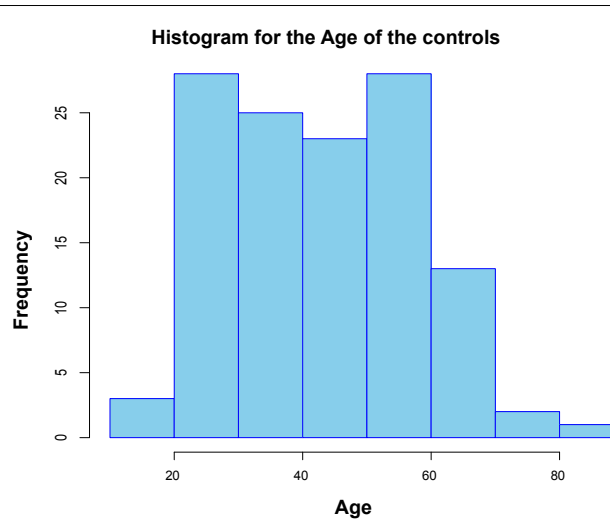

Figure 2: From the above histogram, we can observe that most of the respondents (controls) are falling within the age limit of 20 to 70 years.

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 started after the meal and blood pressure. Variables associated with patient taking any kind of medicine were other disease, diabetes, kidney problem, 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

1. Jarrah MI, Mhaidat NM, Alzoubi KH, Alrabadi N, Alsatari E, et al. (2018) The association between the serum level of vitamin D and ischemic heart disease: a study from Jordan. Vasc Health Risk Manag 14: 119-127.
2. Pusceddu I, Kleber M, Delgado G, Herrmann W, März W, et al. (2018) Telomere length and mortality in the Ludwigshafen Risk and Cardiovascular Health study. PLoS ONE 13: e0198373.
3. Somuncu MU, Demir AR, Karakurt H, Serbest NG, Kalayci B, et al. (2018) The relationship between The Relationship Between Aspirin Resistance and Carotid Imaging in Young Patients With ST-Segment Elevated Myocardial Infarction: A Cross-Sectional Study. Clin Appl Thromb Hemost 24: 1358-1364.
4. Sedky NK, Abdel Rahman MF, Hassanein SI, Gad MZ (2018) Genetic Variants of CYP2R1 Are Key Regulators of Serum Vitamin D Levels and Incidence of Myocardial Infarction in Middle-Aged Egyptians. Curr Pharm Biotechnol 19: 265-273.
5. Hanif M, Ahmad M, Abdelfattah EH (2014) Biostatistics for Health Students with Manual on Software Application. ISOSS Publication.
6. Edgingtion ES (1987) Randomization Tests. Statistics, textbooks and monographs (2ndedn), New York: Dekker, USA.
7. Armitage P, Colton T (2005) Encyclopedia of biostatistics (2ndedn), John Wiley & Sons, Chichester, England.
8. Daldaard P (2008) Statistics and Computing: Introductory statistics with R, Statistics and Computing, Springer Verlag, New York, USA, pp: 1-364.
9. Bland JM, Bland DG (1994) General practice: Statistical Notes one sided and two sided test of significance. British Medical Journal 309: 873-874.
10. Chaudhary SM, Kamal S (1996) Solved problems: Introduction to Statistical Theory Part 2, Ilmi Kitab Khana, Urdu Bazar.