Statistical Analysis of Risk Factors for Cardiovascular disease in Malakand Division

Salahuddin
Department of Statistics, University of Peshawar
Peshawar, Pakistan

Fazle Rabbi
Subject Specialist in Statistics, GHSS, Baidara
Swat, Pakistan

Abstract

Several studies have been conducted to investigate the incidence of cardiovascular disease and to determine the possible risk factors for the disease. In this study, a Statistical method of odds ratio analysis was performed to look at the association of one of the type of cardiovascular disease known as myocardial infarction with various risk factors such as diabetes, cholesterol, hypertension, sex, smoking, obesity, family history and age in Malakand division. A total of 700 patients were examined and their personal and medical data were collected. For each patient, the phenomenon of myocardial infarction was studied in relation to different risk factors. The analyses suggest that hypertension, smoking, diabetes, cholesterol level and family history were important risk factors for the occurrence of myocardial infarction.

Introduction

Cardiovascular disease has emerged as one of the major health problem in Pakistan. Every year a large number of people die due to this terrible disease. There are many types of cardiovascular diseases. One of them is the myocardial infarction (abbreviated as MI). A large number of people die every year due to myocardial infarction (also called heart attack). A myocardial infarction occurs when a coronary artery abruptly fails to deliver blood to a part of our heart.

Various studies have been conducted to investigate the incidence of MI and to determine the possible risk factors for the disease (Ruiswyk. *et al* (1993); Rabajoli. *et al* (1996); Capewell. *et al* (1996); Flack. *et al* (1995); Aminian. *et al* (1998)). They all studied various demographic and medical risk factors of MI. Pais. *et al* (2001) studied the increased risk of acute myocardial infarction with beedi and cigarette smoking in India. They conducted their study in a teaching hospital in Banglore. They found that important risk factor prevalent in the study was smoking.

To see the effects of various risk factors on MI in Malakand Division, we have carried out this study based on the data obtained from the cardiology unit Saidu group of teaching hospitals, Saidu Sharif, Swat. The associations of various risk factors with the occurrence of MI were determine through the statistical method of odds ratio analyses.

Methodology

A total of 700 patients were examined from the cardiology unit of Saidu group of teaching hospitals, Saidu Sharif, Swat for MI to investigate the risk factors of

myocardial infarction (which is a binary variable for this study) like diabetes, cholesterol, hypertension, sex, smoking, obesity, family history and age. The statistical analyses were performed through SPSS software package and the method used was odds ratio analyses.

Odds and the Odds Ratio

Odds refer to the ratio of the probability of an event occurring to that of its not occurring. For example, on average 51 boys are born in every 100 births, so the odds of any randomly selected delivery being that of a boy is: number of boys (51)/number of girls (49), equal to approximately 1.04. The odds of success are defined to be the ratio of the probability of success to the probability of failure. Hence if *p* is the true success probability, the odds of a success are ($\frac{p}{1-p}$). If the observed binary data consists of y success in *n* observations, the odds of a success can be estimated by $\frac{\hat{p}}{1-\hat{p}} = \frac{y}{n-y}$.

When two sets of binary data are to be compared, a relative measure of the odds of success in one set relative to that in the other is the odds ratio. Suppose that p_1 and p_2 are the success probabilities in these two sets, so that the odds of a success in the *ith* set is $\frac{p_i}{1-p_i}$, where *i* = 1, 2. The ratio of the odds of a success in one set of binary data relative to the other is usually denoted by ψ, so that $\psi = \frac{p_1/(1-p_1)}{p_2/(1-p_2)}$ is the odds ratio. When the odds of a success in each of the two sets of binary data are identical, ψ is equal to1. This will happen when the two success probabilities are equal. Values of ψ less than 1 suggest that the odds of a success are less in the first set of data than in the second, while an odds ratio greater than 1 indicates that the odds of a success are greater in the first set of data. The odds ratio is the measure of the difference between two success probabilities, (p_1 - p_2) which can assume values in the interval (-1, 1).

In order to estimate the odds ratio, suppose the binary data are arranged as in the following 2×2 contingency table.

	Number of success	Number of failures
Data set 1	*a*	*b*
Data set 2	*c*	*d*

The estimated success probabilities in the two sets are $\hat{p}_1 = \frac{a}{a+b}$ and $\hat{p}_2 = \frac{c}{c+d}$, and so the estimated odds ratio, $\hat{\psi} = \frac{\hat{p}_1/(1-\hat{p}_1)}{\hat{p}_2/(1-\hat{p}_2)} = \frac{ad}{bc}$. This estimate is the ratio of the products of the two pairs of diagonal elements in the above 2×2 table, and for this reason, $\hat{\psi}$ is sometimes referred to as the cross- product ratio.

Statistical Inference based on Odds Ratio

A commonly used technique in the analysis of count data is the examination of odd ratios. Thus, to measure association between a disease and risk factors in observational studies, odd ratios analysis can be used. To test such an association, we generally set the hypothesis: $H_0 : \psi = 1$ or equivalently $H_0 : \ln(\psi) = 0$.

The hypothesis indicates that we are basically testing for the independence of the two dichotomous variables (disease and risk factor). Note that the hypothesis is the same as $H_0 : \beta_1 = 0$ in the logistic regression equation,

$$\ln(\frac{p}{1-p}) = \beta_0 + \beta_1 x .$$

The null hypothesis $H_0 : \psi = 1$ or $\ln(\psi) = 0$ may be tested using the test statistic

$$Z = \frac{\ln(\hat{\psi})}{S.E\{\ln(\hat{\psi})\}},$$

which has an approximate standard normal distribution.

An approximate 100(1- α) % confidence interval for $\ln(\psi)$ is constructed as:

$$\ln(\hat{\psi}) \pm Z_{\alpha/2} S.E\{\ln(\hat{\psi})\} \qquad \text{(a)}$$

For example, a 95% confidence interval for $\ln(\psi)$ is given by

$$\ln(\hat{\psi}) \pm 1.96 S.E\{\ln(\hat{\psi})\}$$

The confidence interval (a) on inversion will give us the confidence interval for ψ as

$$\hat{\psi} e^{-z\alpha/2 S.E} < \psi < \hat{\psi} e^{z\alpha/2 S.E}$$

If the interval contains unity, it indicates independence; otherwise an association is indicated.

Odds Ratio analyses and Interpretation

1) Myocardial Infarction versus Sex

Consider the following table
SEX * MI

Sex	MI		Total
	0	1	
0	56	132	188
1	144	368	512
Total	200	500	700

In this case the estimated odds of having myocardial infarction are 1.084 times more for males as compared to females. Hence the males are more likely to have myocardial infarction than females. The estimated log odds ratio is 0.081

and its asymptotic standard error is 0.187, which is highly significant at 1% level of significance. A 95 % confidence interval for the true log odds ratio is (- 0.286, 0.448) and so the 95 % confidence interval for the true odds ratio is (0.751, 1.564). The confidence interval shows a significant relationship between sex and myocardial infarction. Hence we can conclude that myocardial infarction is associated with sex.

2) Myocardial Infarction versus Smoking (SMOK)

Consider the following table
SMOK*MI

Smoking	MI		Total
	0	1	
0	168	357	525
1	32	143	175
Total	200	500	700

In this case the estimated odds of having myocardial infarction are 2.103 times more for smokers as compared to non-smokers. Hence the smokers are more likely to have myocardial infarction than non-smokers. The estimated log odds ratio is 0.743 and its asymptotic standard error is 0.217, which is highly significant at 1% level of significance. A 95% confidence interval for the true log odds ratio is (0.318, 1.168) and hence the 95 % confidence interval for the true odds ratio is (1.374, 3.218). The confidence interval shows a significant relationship between smoking and myocardial infarction. Hence we can conclude that myocardial infarction is associated with smoking.

3) Myocardial Infarction versus Diabetes (DIAB):

Consider the following table.
DIAB * MI

Diabetes	MI		Total
	0	1	
0	193	357	550
1	7	143	150
Total	200	500	700

In this case the estimated odds of having myocardial infarction are 11.044 times more for diabetic as compared to non-diabetic. Hence the persons with diabetes are more likely to have myocardial infarction than those who do not have diabetes. The estimated log odds ratio is 2.402 and its asymptotic standard error is 0.397, which is highly significant at 1% level of significance. A 95 % confidence interval for the true log odds ratio is (1.624, 3.180) and hence the 95 % confidence interval for the true odds ratio is (5.072, 24.047). The confidence interval shows a significant relationship between diabetes and myocardial infarction. Hence we can conclude that myocardial infarction is associated with diabetes.

4) Myocardial Infarction versus Family History (FH) of Myocardial Infarction:

Consider the following table.
FH *MI

Family History	MI		Total
	0	1	
0	174	336	510
1	26	164	190
Total	200	500	700

In this case the estimated odds of having myocardial infarction are 3.266 times more for those who have a family history of myocardial infarction as compared to those who do not have family history of myocardial infarction. Hence the persons with a family history of myocardial infarction are more likely to have myocardial infarction than those who do not have family history of myocardial infarction. The estimated log odds ratio is 1.184 and its asymptotic standard error is 0.231, which is highly significant at 1% level of significance. A 95 % confidence interval for the true log odds ratio is (0.731, 1.637) and hence the 95 % confidence interval for the true odds ratio is (2.077, 5.136). The confidence interval shows a significant relationship between family history and myocardial infarction. Hence we can conclude that myocardial infarction is associated with family history of myocardial infarction.

5) Myocardial Infarction versus Cholesterol (CHL)

Consider the following table.
CHL * MI

Cholesterol	MI		Total
	0	1	
0	187	390	577
1	13	110	123
Total	200	500	700

Here the estimated odds of having myocardial infarction are 4.057 times more for those who have a high cholesterol level as compared to those having low cholesterol level. Hence the persons with a high cholesterol level are more likely to have myocardial infarction than those who have a low cholesterol level. The estimated log odds ratio is 1.400 and its asymptotic standard error is 0.306, which is highly significant at 1% level of significance. A 95 % confidence interval for the true log odds ratio is (0.800, 1.999) and hence a 95 % confidence interval for the true odds ratio is (2.227, 7.391). The confidence interval shows a significant relationship between high cholesterol level and myocardial infarction. Hence we can conclude that myocardial infarction is associated with high cholesterol level.

6) Myocardial Infarction versus Hypertension (HYP)

Consider the following table.

HYP * MI

Hypertension	MI		Total
	0	1	
0	174	286	460
1	26	214	240
Total	200	500	700

In this case the estimated odds of having myocardial infarction are 5.008 times more for persons having hypertension as compared to those who do not have hypertension. Hence the persons with hypertension are more likely to have myocardial infarction than those who do not have hypertension. The estimated log odds ratio is 1.611 and its asymptotic standard error 0.229, which is highly significant at 1% level of significance. A 95 % confidence interval for the true log odds ratio is (1.162, 2.060) and hence the 95 % confidence interval for the true odds ratio is (3.197, 7.845). The confidence interval shows a significant relationship between the hypertension and myocardial infarction. Hence we can conclude that myocardial infarction is associated with the presence of hypertension.

7) Myocardial Infarction versus Obesity (OBS)

Consider the following table.

OBS * MI

Obesity	MI		Total
	0	1	
0	184	427	611
1	16	73	89
Total	200	500	700

In this case the estimated odds of having myocardial infarction are 1.966 times more for persons having obesity than those who have no obesity. Hence the persons having obesity are more likely to have myocardial infarction than those who have no obesity. The estimated log odds ratio is 0.676 and its asymptotic standard error is 0.290, which is highly significant at 1% level of significance. A 95 % confidence interval for the true log odds ratio is (0.108, 1.244) and hence the 95 % confidence interval for the true odds ratio is (1.114, 3.471). The confidence interval shows a significant relationship between obesity and myocardial infarction. Hence we can conclude that myocardial infarction is associated with obesity.

8) Myocardial Infarction versus Age (> or = 60 years)

Consider the following table.

AGE * MI

Age	MI		Total
	0	1	
0	52	160	212
1	148	340	488
Total	200	500	700

In this case the estimated odds of having myocardial infarction are 0.747 times more for persons having age less than 60 years than those having age equal to or greater than 60 years. Hence the persons having age equal to or greater than 60 years are more likely to have myocardial infarction than those who have age less than 60 years. The estimated log odds ratio is -0.292 and its asymptotic standard error is 0.188, which is highly significant at 5% level of significance. A 95 % confidence interval for the true log odds ratio is

(-0.660, 0.076) and hence the 95 % confidence interval for the true odds ratio is (0.517, 1.079). The confidence interval shows a significant relationship between age and myocardial infarction. Hence we can conclude that myocardial infarction is associated with age (equal to or greater than 60 years).

Conclusion

The main objective of the present study was to determine the most likely risk factors of myocardial infarction and to model the incidence of myocardial infarction in patients arriving at cardiology unit Saidu group of teaching hospitals Saidu Sharif Swat. A total of 700 patients were examined and their personal and medical data were taken. For each patient, the phenomena of myocardial infarction was studied in relation to different risk factors, like hypertension, cholesterol, diabetes, smoking, sex, obesity, age and family history of myocardial infarction.

Out of the total of 700 patients, 512 were male and 188 were female patients. Out of 700 patients, 500 had myocardial infarction and 200 had no myocardial infarction. Of 512 male patients, the number of myocardial infarction patients were 368 and 144 patients had no myocardial infarction. Of 188 female patients, 132 had myocardial infarction and 56 had no myocardial infarction. Out of 700 patients, 240 had hypertension and 460 had no hypertension. But out of 500 myocardial infarction patients 214 had severe hypertension. In these 214 sever hypertensive patients, 152 were male and 62 were female. Out of 200 patients who had no myocardial infarction, only 26 were hypertensive and the rest of 174 were non-hypertensive patients. Another important risk factor considered in this study for myocardial infarction was cholesterol. Out of total patients 123 had high cholesterol level. Out of these 123 patients, 110 had myocardial infarction among which 80 were male and 30 were female patients. There were a total of 175 smokers and 525 were non-smokers. Of 500 myocardial infarction patients, 143 were smokers. Of 175 smokers 143 patients had myocardial infarction and 32 had no myocardial infarction. This means that smokers are more likely to have myocardial infarction than non-smokers. In these 143 patients, who are smokers and had myocardial infarction as well, 137 were male and 6 were female patients. Out of total of 700 patients, 150 were diabetic, of which 143 patients had myocardial infarction and only 7 patients had no myocardial infarction. This shows a high effect of diabetes on myocardial infarction. In these 143 patients 112 were male and 31 were female. A total of 190 patients had family history of myocardial infarction. Of these 190 patients, 164 had myocardial infarction and only 26 patients had no myocardial infarction. So those patients who had family history of myocardial infarction were more likely to have myocardial infarction as

compared to those who had no family history of myocardial infarction. In those164 patients who have family history of myocardial infarction and had myocardial infarction as well, 123 were male and 41 were female. There were 488 patients of age equal to or greater than 60 years; of which 340 patients had MI. 212 patients were of age less than 60 years of which 160 had MI. This indicates a risk of MI for old age people.

Odds ratio analyses were performed to look at the association of MI with different risk factors of MI. The odds of MI for male patients were 1.084 times more than for female patients. This shows an association between MI and sex. The odds of MI for smoker patients were 2.103 times more than for non-smokers. This indicates an association between MI and smoking. The odds of MI for diabetics were 11.044 times more than for non-diabetics. This shows a significant relationship between MI and diabetes. The odds of MI for patients having family history of MI were 3.266 times more than for those having no family history of MI. This also shows a significant association between MI and family history of MI. There was a significant association between MI and high cholesterol level. The odds of MI for high cholesterol were 4.057 times more than those patients having low cholesterol level. The odds of MI for hypertensive patients were 5.008 times more than for non-hypertensive patients. This also indicates a significant association between MI and hypertension. There was also an association between obesity and MI. The odds of MI for those patients who had obesity were 1.966 times more than for those who had no obesity. Age and MI also proved to be associated with each other. The odds of MI for patients of age less than 60 years were 0.747 times more than for those of age equal to or greater than 60 years.

In summary, we conclude from the analysis of random sample of 700 patients that MI has a strong association with hypertension, cholesterol, diabetes, smoking and family history of myocardial infarction.

References

1. Aminian, B., Habibzadeh, F. and Moarref, A.R., (1998), *Factor analysis of cardiogenic shock complicating acute myocardial infarction*, National Iranian Oil Company out patient polyclinics, Shiraz, Iran.
2. Capewell, S., Kendrick, S., Boyd, J., Cohen, G., Juszczak, E. and Clarke, J., (1996), *Measuring outcomes: one month survival after acute myocardial infarction in Scotland*, British medical Journal, 76, pp: 70-75.
3. Draper, N. and Smith, H., (1981), *Applied Regression Analysis*, Second edition. John Willey and sons: New York.
4. Flack, J.M., Neatson, J., Grimm, R., Shih, J., Cutler, J., Ensrud, K. and MacMahon, S., (1995), *Blood pressure and mortality Among Men with prior Myocardial Infarction*, Journal of American Heart Association, 92, pp: 2437-2445.
5. Pais, P., Fay, M.P. and Yusuf, S., (2001), *Increased Risk of Myocardial Infarction Associated With Beedi and Cigarette Smoking in India*, Indian Heart Journal; 53: pp: 731-735.
6. Rabajoli, F., Arneodo, D., Balzola, F., Leo, L. and Vineis, P., (1996), *Moderate alcohol intake and risk of myocardial infarction among non-smokers*, European journal of Public Health, 6(3), pp: 227-230.
7. Ruiswyk, J.V., Hartz, A., Kuhun, E., Krakuer, H., Young, M. and Rimm, A., (1993), *A measure of mortality risk for elderly patients with acute myocardial infarction*, Division of General Internal Medicine, Medical college of Wisconsin, Milwaukee.13 (2), pp: 152-160.