❒   1615

A predictive model for prediction of heart surgery procedure

Aufzalina Mohd Yusof[1], Nor Azura Md. Ghani[2], Khairul Asri Mohd Ghani[3],
Khairul Izan Mohd Ghani[4]

[1,2]Center for Statistical and Decision Sciences Studies, Faculty of Computer & Mathematical Sciences, Universiti Teknologi MARA, Malaysia
[3]Department of Surgery, Faculty of Medicine & Health Sciences, Universiti Putra Malaysia, Malaysia
[4]Columbia Asia Malaysia, Malaysia

Article Info	ABSTRACT
	Coronary heart disease (CHD) is a disease in which plague in the form of waxy substance builds up inside the coronary arteries. Coronary artery bypass grafting (CABG) is used as treatment on CHD patients but the role of CABG has been challenged by percutaneous coronary intervention (PCI) when it was introduced in 1977. Drug eluting stents (DES) was introduced with the development of PCI. The purpose of this study was to find the potential risk factors that associated with the procedures (CABG and DES) and to model procedure (CABG vs DES) on coronary heart disease male patients aged 45 years old and below. The study sample was among male patients aged 45 years old and below who has undergone CABG or DES procedure at either IJN or HUKM from January 2007 until December 2010. Logistic regression was used to model treatment selection on coronary heart disease with 87.3% of the classification rate. Patient who i) smoke, ii) obese, or ii) had dyslipidemia was significantly associated with DES, and the other factors were prone to have CABG as their treatment.

Corresponding Author:

Nor Azura Md. Ghani,
Center for Statistical and Decision Sciences Studies,
Faculty of Computer & Mathematical Sciences,
Universiti Teknologi MARA,
40450 Shah Alam, Selangor, Malaysia.
Email: azura@tmsk.uitm.edu.my

1. INTRODUCTION

Coronary heart disease (CHD) is a disease in which plague in the form of waxy substance builds up inside the coronary arteries, whereby oxygen-rich blood will be supplied to heart muscle by these arteries [1]. The plague can break open (rupture) or hardened over time. If ones suffer with ruptured plague, the blood clot will be formed on its surfaces and this can disturb the blood flow through coronary artery. Nevertheless, the hardened plague that is built over time can narrow the artery system and blood flow will be reduced or blocked. If the flow of oxygen-rich blood to heart muscle is not sufficient, angina (chest pain) or heart attack can occur. Ones will feel like pressure or squeezing in the chest or may even feel like indigestion. The pain also can be felt at shoulders, arms, neck or back. Without quick treatments, CHD can lead to serious heart problem like heart failure, in which the heart fails to pump enough blood to meet the body needs. For almost half of century, coronary artery bypass grafting (CABG) has been regarded as the most effective revascularization treatment and become a standard care for CHD patients [2].

However, the role of CABG has been challenged over the last two decades by percutaneous coronary intervention (PCI) when it started to be introduced in 1977. Due to the increasing number of heart disease globally, several studies has been conducted that focus on the diagnostic system [3-5]. As up to date numerous studies have been conducted for the purpose of comparing the effectiveness of PCI with CABG

[6-8]. However, lacks of studies have been conducted to determine which risk factor is associated with the use of suitable procedure in order to treat CHD. Therefore, the purpose of this study is to identify the important risk factor for the selection of CABG or drug eluting stents (DES) procedure among young patients specifically less than 45 years old.

1.1.   Factors related to Selection of Procedure

Generally, the selection of suitable procedures (CABG or DES) to cure CHD depends on lifestyle profile, the types of comorbodities (such as diabetes melitus, chronic obstructive pulmonary disease, atrial fibrillation, renal impairment) and the number of vessel disease suffered by patients. For instance, Mavromatis *et al.* [9] in their study stated that coronary CABG is found to be more effective over PCI in improving the mortality in patients with diabetes and three-vessel coronary artery disease who were not at high surgical risk. The finding is similar with a study conducted by Nystrom et al. [10] where CABG as a preferred strategy for patients with Type 1 Diabetes in need of multivessel revascularization.

The safety and efficacy of DES in older patients with chronic kidney disease (CKD) is measured by Tsai *et al.* [11] and it was found that that there was a significant reduction in mortality, myocardial infarction and revascularization. The recent study by Lu *et al.* [12] also reported that the condition of CKD patients are significantly improved with DES, however large-sized randomized controlled trials were necessary to determine the real effect on CKD patients and whether efficacy differs by type of DES.

With regards to atrial fibrillation disease, Ruiz-Nodar *et al.* [13] had revealed the effectiveness of DES was not in the long run as many problems arose such as high rate major bleeding, higher rate of stent thrombosis and higher risk of thrombo-ebolism. CABG in other hand also do not show the long term effectiveness when atrial fibrillation remained the most common disease after open surgery and occurred in up to 50% of patients [14].

The development of cardiovascular disease is well associated with obesity when majority of high BMI patients are among the population of CHD. In a study by Johnson *et al.* [15], overweight and obese patients were benefited with CABG when they had a survival advantage over underweight, normal weight and morbidly obese patients. Psychologically, low BMI patients may not tolerate with the effects of weight loss as compared to obese individuals possibly may lead towards high mortality [16].

2.   RESEARCH METHOD

2.1.   Sample

This study used a secondary data which was obtained from Division of Cardiothoracic, Hospital Universiti Kebangsaan Malaysia (HUKM) and Institut Jantung Malaysia (IJN). It covered a total of 315 male patients' record from January 2007 until December 2010. The study sample involved only patients aged 45 years old and below who undergone CABG or DES procedure. Data comprised of demographic profiles, lifestyle profile, co-morbid condition, vessel involvement. The outcomes of procedure were either CABG or DES.

2.2.   Binary Logistic Regression

Binary logistic regression model was used in order to model procedure (CABG vs DES) based on associated risk factors. Logistic regression allows one to predict a discrete outcome such as group membership from a set of variables that may be continuous, discrete, dichotomous or a mix of any of these. Generally, binary logistic regression is used when the dependent or response variable is a dichotomous variable such success or failure. The logistic regression model describes the probability of occurrence of the outcome of the dependent variable [17, 18]. The model is designed to describe the probability which is between 0 and 1, whereas in theory continuous variables can take any value between plus or minus infinity [19]. This means that we cannot assume normality for a proportion and we must recognize that proportion have a binomial distribution. The mean and variance of the binomial distribution are not independent. The mean is denoted by π and the variance is denoted by $\pi*(1-\pi)/n$ where *n* is the number of observation and π is the probability of the event occurring in any one trial.

Let:

$$\pi_i = \Pr\,(Y=1|\,X = x_i)$$

The logistic regression model:

$$\log\left(\frac{\pi_i}{1-\pi_i}\right)=\beta_0 + \beta_1 * gender + \beta_2 * race + \beta_3 * smoking + \beta_4 * obese + \beta_5 *$$
$$dyslipidemia+\beta_6 * Diabetes + \beta_7 * Atrial\ Fibrilliation + \beta_8 * Hypertension + \beta_9 *$$
$$Unstable\ Angina + \beta_{10} * Dyslipidmimia + \beta_{11} * Left\ Anterior\ Descending + \beta_{12} *$$
$$Left\ Circumflex + \beta_{13} * Right\ Coronary\ Artery$$

where

π_i	= Probability of having CABG as heart treatment option
$1-\pi_i$	= Probability of having DES as heart treatment option
x_i	= Risk factors of complication during surgery
β_k	= Coefficient of parameter

3. RESULTS AND DISCUSSION
3.1. Respondent Profile

In general, background of male patients are studied which include demographic profile, lifestyle profile, the other disease that they had at the same time (co-morbid condition), and the heart vessel that was affected from heart disease (vessel involvement). As regard to ethnicity, the disease being most common among the Indian population (Table 1). Based on the estimated population of Malaysia [20], coronary heart disease incidence among the Malay and Chinese populations are still considerably lower by 10.9% and 12.2%, respectively. However, for the Indian population the incidence of coronary heart disease was high at 23.8%. Further, the distribution of smokers and non-smokers was quite balanced with 50.8% and 49.2%, respectively. However, only 13.3% male patients face with obesity issue while 86.7% patients do not have the particular issue.

Table 1. Population in Malaysia 2010 According to Ethnicity

Ethnic Group	*Estimated Population	Patients Diagnosed with CHD	Difference
Malay	67.4%	56.5%	-10.9%
Chinese	24.6%	12.4%	-12.2%
Indian	7.3%	31.1%	23.8%
Others	0.7%	-	-

*Source: Department of Statistics Malaysia, 2010

3.2. Test of Association

The two-way association was attempted to measure the association between variables of each category (lifestyle profile, co-morbid condition and vessel involvement) with the selection of procedure. The odds ratio for variables that indicate significant association was then measured to quantify how strong the absence or presence of risk factors is associated with the presence or absence of procedure (CABG or DES). With reference to lifestyle profile, both variables (smoking and obesity) had a significant association with the procedure. Among all variables which fall under co-morbid condition category, only dyslipidemia and diabetes had an association with the selection procedure. In other hand, vessel involvement category shows association with the procedure when p-value for left anterior descending (LAD), left circumflex (LCX) and right coronary artery (RCA) were less than 0.05 which indicated significant results. With reference to odd ratio estimates, patient who smoked, obese or having dyslipidemia was more likely to be in DES group meanwhile diabetes, LAD, LCX and RCA belong to CABG procedure.

3.3. Binary Logistic Regression

The best model development starts with measuring its performance using several statistical tests after confirming the dataset has no indication of outliers and multicollinearity. Table 2 shows the Omnibus Tests of Model Coefficient to measure the performance of logistic regression model. Based on the analysis, it was concluded that at least one of predictors was significant in order to predict the suitable procedure of treating coronary heart disease when the significant value is less than 0.10

Table 2. Omnibus Test of Model Coefficient

	Chi-square	Degree of freedom	Sig.
Model	250.102	10	0.000

The proportion of variation in the response variable by predictors is determined by Cox & Snell R-Square and Nagelkerke R-Square in Table 3. The model was better when the value was closer to one. Therefore, the selection of procedure is explained by predictors about 54.7% to 73.2%.

Table 3. Variation of Response

Step	1.
-2 Log likelihood	182.684
Cox & Snell R Square	0.584
Nagelkerke R Square	0.734

The Hosmer and Lemeshow goodness-of-fit statistic assessed how well the logistic regression model fits the data. Based on the analysis as displayed on Table 4, it was concluded that the model fits the data very well since the significant value was greater than 0.05.

Table 4. Hosmer and Lemeshow Test

Chi-square	Df	Sig.
3.20	8	0.921

Classification table in Table 5 showed the percentage of correct classification of dependent variable (CABG vs DES) as well as measuring the error rate. The performance of the model also can be predicted by error rate by measuring the proportion of misclassified observation. As illustrated in the same table, the data was correctly classified by 87.3% and the error rate was 12.7%. Hence, it showed that this model could classify accurately the observed and predicted of procedure.

Table 5. Classification Table

	Observed	Predicted		
		Procedure		% correct
		CABG	DES	
Procedure	CABG	121	19	86.4
	DES	21	54	88.0
Overall Percentage		45.1	54.9	87.3

Table 6 shows the parameter estimates, standard error, Wald statistics, degree of freedom, significant value and odds ratio (Exp B) of full model (included all predictors). Among variables that have been evaluated, there were seven significant variables in predicting the suitable procedure for treating CHD which included smoking, obesity, dyslipidemia, diabetes, LAD, LCX and RCA. The final estimated logistic regression as follows.

Table 6. Variables of Estimated Model

Variables	B	S.E.	Wald	df	Sig.	Exp(B)
Smoking(1)	-1.942	0.420	21.377	1	0.000	0.143
Dyslipidemia(1)	-2.248	0.429	27.428	1	0.000	0.106
Pulmonary Hypertension(1)	0.137	0.409	.112	1	0.738	1.147
Diabetes_Melitus(1)	0.854	0.408	4.369	1	0.037	2.348
Obesity(1)	-2.294	0.655	12.275	1	0.000	0.101
Unstable Angina(1)	0.084	0.756	0.012	1	0.912	1.087
Atrial Fibrilliation(1)	0.992	1.107	0.803	1	0.370	2.696
Left Anterior Desceding(1)	3.421	0.518	43.575	1	0.000	30.585
Left Circumflex(1)	2.693	0.449	35.930	1	0.000	14.778
Right Coronary Artery(1)	1.791	0.440	16.535	1	0.000	5.993
Constant	-3.771	0.701	28.923	1	0.000	0.023

$$\log\left(\frac{\pi_i}{1-\pi_i}\right) = -3.771 - 1.942 * smoking - 2.248 * dyslipidemia + 0.137 * Diabetes - 0.294 * obesity + 3.421 * LAD + 2.693 * LCX + 1.791\, RCA$$

$$\pi_i = P(\pi(y) = 1)$$

In summary, with respective to lifestyle profile, smoking and obese patients were more likely to have DES with odds ratio of 6.99 and 9.90 respectively. In term of co-morbid condition, patients with dyslipidemia was 9.43 times more likely to have DES meanwhile patient with diabetes mellitus is 2.35 times more likely to have CABG. As regard to vessel involvement all risk factors (LAD, LCX and RCA) were more likely to have CABG at odd ratio of 30.6, 14.78 and 5.99 respectively.

4. CONCLUSION

This paper intended to add knowledge of the use of statistics in medical as it allows clinical researchers to draw reasonable inferences based on available information, thus to make a sound decision. Even though, medical practitioners would usually conduct clinical trial to determine the safe and effective procedure for patients [21-23], this paper would beneficial them as a guideline in statistical point of view to select a suitable procedure to cure CHD. CABG or DES is the treatment procedure that commonly used for CHD patients. Binary logistic regression model was used in order to model procedure (CABG vs DES) based on associated risk factors (lifestyle profile, co-morbid condition and vessel disease). From the findings, it clearly justified that binary logistic regression model was appropriate to find the risks factors of disease [24- 26], particularly in this study that associated with the selection of either CABG or DES. With the reference to lifestyle, smoking and obese patients were more likely to have DES. In term of co-morbid condition, patients with dyslipidemia were likely to have DES while diabetic patients were prone to have CABG. All risk factors that associated with vessel disease were more likely to have CABG as a heart treatment.

REFERENCES

[1] J. E. Dalen, et al., Department of Medicine, University of Arizona College of Medicine, *Am. J. Med.*, 2014.
[2] D. P. Taggart, "Stents or surgery in coronary artery disease in 2013," vol. 2, pp. 431-434, 2013.
[3] A. S. Aziz, et al., "Cardiac motions classification on sequential PSAX echocardiogram," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 12, pp. 1289-1296, 2018.
[4] M. K. Omer, et al., "Deep Neural Network for Heart Disease Medical Prescription Expert System," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 6, 2018.
[5] W. Wiharto, et al., "A tiered approach on dimensional reduction process for prediction of coronary heart disease," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 11, pp. 487-495, 2018.
[6] M. A. Hlatky, et al., "Comparing the Effectiveness of Coronary Artery Bypass Graft Surgery and Nonsurgical Catheter-Based Interventions for Coronary Artery Disease," *Ann. Intern. Med.*, vol. 158, pp. 727-734, 2013.
[7] D. Giacoppo, et al., "Percutaneous Coronary Intervention vs Coronary Artery Bypass Grafting in Patients With Left Main Coronary Artery Stenosis," *JAMA Cardiol.*, vol. 2, pp. 1079, 2017.
[8] U. Baber, et al., "Comparative efficacy of coronary artery bypass surgery vs. percutaneous coronary intervention in patients with diabetes andmultivessel coronary artery disease with or without chronic kidney disease," *Eur. Heart J.*, vol. 37, pp. 3440-3447, 2016.
[9] K. Mavromatis, et al., "Revascularization in Patients with Diabetes : PCI or CABG or None at All," 2015.
[10] T. Nystrom, et al., "PCI Versus CABG in Patients With Type 1 Diabetes and Multivessel Disease," no. 20150603, 2017.
[11] T. T. Tsai, et al., "Safety and Efficacy of Drug-Eluting Stents in Older Patients With Chronic Kidney Disease A Report From the Linked CathPCI Registry-CMS Claims Database," *JAC*, vol. 58, pp. 1859-1869, 2011.
[12] R. Lu, et al., "Comparison of Drug-Eluting and Bare Metal Stents in Patients With," 2016.
[13] A. Hurtado, et al., "Efficacy and safety of drug-eluting stent use in patients with atrial fibrillation," pp. 932-939, 2009.
[14] M. Golmohammadi, et al., "Incidence and Risk Factors for Atrial Fibrillation after First Coronary Artery Bypass Grafting in Urumiyeh Imam Khomeini Hospital from 2006 to 2008," pp. 86-90, 2008.
[15] A. P. Johnson, et al., "Body Mass Index, Outcomes, and Mortality Following Cardiac Surgery in Ontario, Canada," 2011.
[16] C. L. Jackson, et al., "Body-Mass Index and All-Cause Mortality in US Adults With and Without Diabetes.," *J. Gen. Intern. Med.*, vol. 29, pp. 25-33, 2014.
[17] M. P. Mamat, et al., "Forestry and Economic Research using SPSS," *Univ. Teknol. Mara (UiTM), For. Res. Inst. Malaysia Minist. Nat. Resour. Environ.*, 2010.

[18] N. A. Md Ghani, et al., "Comparison between Probit Link and Cloglog Link Functions for the Identification of Pathological Staging in Colorectal Cancer Data," vol. 7, 2013.
[19] D. G. Kleinbaum, et al., "Logistic Regression: A Self- Learning Text.," *New York Springer Bus Media LCC*, 2010.
[20] "Population Distribution and Basic Demographic Characteristic Report 2010," Department of Statistics Malaysia, 2010.
[21] H. S. Alamri, et al., "Lessons from the SYNTAX trial," *J. Saudi Hear. Assoc.*, vol. 22, pp. 35-41, 2010.
[22] L. Gao, et al., "Percutaneous coronary intervention using drug-eluting stents versus coronary artery bypass graft surgery in left main coronary artery disease an updated meta-analysis of randomized clinical trials," *Oncotarget*, vol. 8, pp. 66449-66457, 2017.
[23] A. H. Gershlick, et al., "Outcomes After Left Main Percutaneous Coronary Intervention Versus Coronary Artery Bypass Grafting According to Lesion Site: Results From the EXCEL Trial," *JACC Cardiovasc. Interv.*, vol. 11, pp. 1224-1233, 2018.
[24] N. Mohamed and M. Abd, "Binary Logistic Regression to Identify the Risk Factors of Eye Glaucoma," vol. 4531, pp. 366-376.
[25] Z. Zheng, et al., "The Logistic Regression Analysis on Risk Factors of Hypertension among Peasants in East China & Its Results Validating," vol. 10, pp. 416-420, 2013.
[26] M. Kudakwashe and K. M. Yesuf, "Application of Binary Logistic Regression in Assessing Risk Factors Affecting the Prevalence of Toxoplasmosis," *Am. J. Appl. Math. Stat.*, vol. 2, pp. 357-363, 2014.

BIOGRAPHIES OF AUTHORS

Aufzalina Mohd Yusof is now a doctorate student in Center for Statistical Studies and Decision Sciences, Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, Malaysia under the supervision of Nor Azura Md. Ghani. Her research interest is in biostatistics and she is currently working on a research related to survival analysis specifically dealing with recurrent disease.
Email: aufzalina@gmail.com

Nor Azura Md.Ghani is an Associate Professor in Center for Statistical Studies and Decision Sciences, Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, Malaysia. Currently she is serving as a Head of Data Research Unit, Research Management Center, Institute Research Management & Innovation, Universiti Teknologi MARA, Malaysia. Since 2017, she serves as Vice Chair IEEE Computer Society Malaysia Chapter. Her expertise is big data, statistical pattern recognition and forensic statistics.
E-mail: azura@tmsk.uitm.edu.my

Assoc. Prof Dato' Dr. Khairul Asri Mohd Ghani is a Consultant Urologist Surgeon and Head Department of Surgery, Faculty of Medicine and Health Sciences UPM. He is also senior lecturer for Universiti Putra Malaysia. Dr Khairul Asri graduated in 2001 and he obtained Master Of Surgery (General Surgery) from UKM in 2009. He continued his higher urological training in Institute of Urology Hospital Kuala Lumpur and at Department of Urology Hospital Selayang. He passed the Conjoint Malaysian Board of Urology examination and FRCS (Urol) Glasgow examination in 2013. He subsequently completed his fellowship in advanced minimally invasive surgery at Uniklinikum Tubingen, Germany. He is currently Malaysia Urology Association (MUA) council member and board member of the Training/Credentialing committee for Urologist Surgeon in Malaysia .He has published in various international journals, presented at many urology conferences and health care programmed in Malaysia.
E-mail: drkhairulasri@gmail.com

Khairul Izan Mohd Ghani is a Consultant General Surgeon of Columbia Asia Petaling Jaya, Selangor. Before that he was a Clinical Lecturer and Specialist Surgeon at Department of Surgery, Faculty of Medicine, Universiti Kebangsaan Malaysia. He has extensive experience in the field of surgery and also published in various international journals.
E-mail: drkhairulizan@gmail.com