Pertanika J. Sci. & Technol. 28 (4): 1163 - 1177 (2020)

SCIENCE & TECHNOLOGY

Journal homepage: http://www.pertanika.upm.edu.my/

Classification of Familial Hypercholesterolaemia Using Ordinal Logistic Regression

Muhammad Hamizan Jamaludin[1], Yap Bee Wah[2]*, Hapizah Mohd Nawawi[3], Chua Yung-An[3], Marshima Mohd Rosli[1] and Muthukkaruppan Annamalai[1]

[1]*Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, 40450 UiTM, Shah Alam, Selangor, Malaysia*
[2]*Advanced Analytics Engineering Centre, Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, 40450 UiTM, Shah Alam, Selangor, Malaysia*
[3]*Institute for Pathology, Laboratory and Forensic Medicine (I-PPerForM), University Teknologi MARA, 47000 UiTM, Sungai Buloh, Selangor, Malaysia*

ABSTRACT

Familial hypercholesterolaemia (FH) is a genetic disease that causes the elevation of low-density lipoprotein cholesterol (LDL-C), which subsequently leads to premature coronary heart disease (CHD). Features which have been reported to be associated with FH include lipids level, tendon xanthomata, and history of CHD. The Ordinal Logistic Regression model using the classification of FH patients with the Dutch Lipid Clinic Network Criteria (DLCN) as the dependent variable (where 1=Possible, 2=Probable, 3=Definite) was developed and evaluated for different types of link functions. The FH patients (n = 449) were recruited from health screening programmes conducted in hospitals and clinics in Malaysia from 2010 to 2018. Results indicate there is a significant association between FH categories with demographic factors (ethnicity and smoking) and physical symptoms (corneal arcus and xanthomata). The Ordinal Logistic Regression using Cauchit link function has lower Akaike Information Criterion (AIC) value, higher Nagelkerke's R-Square and classification accuracy compared to Probit and Logit

ARTICLE INFO

Article history:
Received: 6 March 2020
Accepted: 12 June 2020
Published: 21 October 2020

DOI: https://doi.org/10.47836/pjst.28.4.03

E-mail addresses:
mizanjamaludin93@gmail.com (Muhammad Hamizan Jamaludin)
beewah@tmsk.uitm.edu.my (Yap Bee Wah)
hapizah.nawawi@gmail.com (Hapizah Mohd Nawawi)
yungan.chua@gmail.com (Chua Yung-An)
marshima@tmsk.uitm.edu.my (Marshima Mohd Rosli)
mk@tmsk.uitm.edu.my (Muthukkaruppan Annamalai)
*Corresponding author

ISSN: 0128-7680
e-ISSN: 2231-8526

link function, diastolic blood pressure, corneal arcus and xanthomata were found to be significant covariates of FH.

Keywords: Classification, dutch lipid clinic network criteria, familial hypercholesterolaemia, ordinal logistic regression

INTRODUCTION

Familial hypercholesterolemia (FH) is an inherited dyslipidaemia that causes abnormal elevation of low-density lipoprotein cholesterol (LDL-C) since birth and leads to premature cardiovascular diseases as early as in the second decade of life (Mundal et al., 2018). It is an autosomal dominant genetic disorder that usually affects LDL receptor (*LDLR*), apolipoprotein B (*APOB*) and proprotein convertase subtilisin/kexin type 9 (*PCSK9*) genes. The genetic defects impair the liver capability in metabolising or removing excess LDL-C, resulting in very high plasma LDL-C levels which can lead to atherosclerotic cardiovascular diseases (ASCVD), such as coronary heart disease (CHD), strokes, and peripheral vascular diseases (Wong et al., 2016). With the global prevalence of 1:250 (Akioyamen et al., 2017), FH is not a rare disease and will significantly contribute to the global premature ASCVD-related mortality if not treated.

There is a need to improve the identification of FH in general population. Using a validated tool developed from data available in primary care records may provide a high potential solution to the problem of under diagnosing diagnostic criteria of FH (Weng et al., 2018). Clinicians should give priority to the individuals or patients with highest chances of suffering FH to be assessed clinically using widely known diagnostic toolssuch as Simon-Broome (SB) or Dutch Lipid Clinic Network Criteria (DLCN)(Varghese, 2014).

A good classification model will be useful to classify FH patients in Malaysia using some significant clinical and genetic factors related to FH. Several studies have used chi-square test of association to identify the factors (age, gender, ethnic and medical history) that are associated with FH (Abdul-Razak et al., 2017; Khoo et al., 2000). Lye et al. (2013) applied the logistic regression to determine the variations in *LDLR, APOB*, PCSK9 and other lipid-related genes inclinically diagnosed FH patients. Shi et al.(2014) also applied logistic regression and found work that needed physical skills and living in rural areas were significantly associated with FH. Most studies applied binary logistic regression model where the dependent variable is a dichotomous variable (FH or non-FH) (Norusis, 2005; Warner, 2008). However, the Dutch Lipid Clinic Network Criteria has four categories (Definite, Probable, Possible and Unlikely FH), while the Simon-Broome has three categories (Definite, Possible and Unlikely FH). All of the categories are sorted in ranks which perceives the Definite FH as the highest rank followed by "Probable FH" and the lowest rank is the Unlikely FH.

Multinomial Logistic Regression is useful when the dependent variable of study has more than two categories. However, when the categories of the dependent variable are ordinal in nature, the Ordinal Logistic Regression (OLR) model should be used because the model takes into consideration the ordinal nature of the data. The OLR model involves the proportional odds assumption or parallel lines assumption which allows a set of parallel line for each category of the dependent variable. If the test of parallel line is accepted, we can determine the effect of the covariate on the odds of moving to a higher-order category of the dependent variable. There are five different link functions that relates the expected value of the response to the linear predictors. The aim of this study was to identify the most suitable Ordinal logistic regression model using different link functions for classification of FH using DLCN for FH patients in Malaysia. The OLR models were evaluated based on AIC (Akaike Information Criterion), Nagelkerke's R-Square and classification accuracy.

MATERIALS AND METHODS

Data Source

Volunteers (n= 5171) were recruited from community health screening programmes and Specialist Lipid Clinics in Malaysia from 2010 to 2018. Familial hypercholesterolaemia was diagnosed using DLCN. Individuals with DLCN category of Possible, Probable or Definite FH were considered as FH patients. Irrespective of the DLCN categories, we excluded individuals with LDL-C<4.0 mmol/L. After excluding non-FH individuals, a total of 449 patients were recruited for this study.

Theoretical Framework

The demographic factos are age, gender, ethnicity, smoking status, and history of coronary artery disease (CHD). Other factors data consist of systolic blood pressure, diastolic blood pressure, total cholesterol (TC), triglyceride (TG), LDL-C, high-density lipoprotein (HDL), xanthomata, and corneal arcus. The theoretical framework proposed in this study is presented in Figure 1 where the dependent variable is FH Category (1=Possible, 2=Probable, 3=Definite).

Descriptive Analysis

Based on the frequency distribution of patients shown in Table 1, the distribution of gender is slightly higher for female (240 or 53.5% female , 209 or 46.5% male). About 78% were Malay (78%) more than half of the patients do not smoke (64.6%). About 78% did not suffer from hypertension and 72.4% (n = 325) were without diabetes while 80% did not have CHD. About 31.8% (or 143) patients did not have signs of corneal arcus and 30.3% did not suffer from xanthomata. As shown in Figure 2, 65% (n = 290) were diagnosed as Possible, 11% (n = 48) were Probable, while 25% (111) were Definite FH.

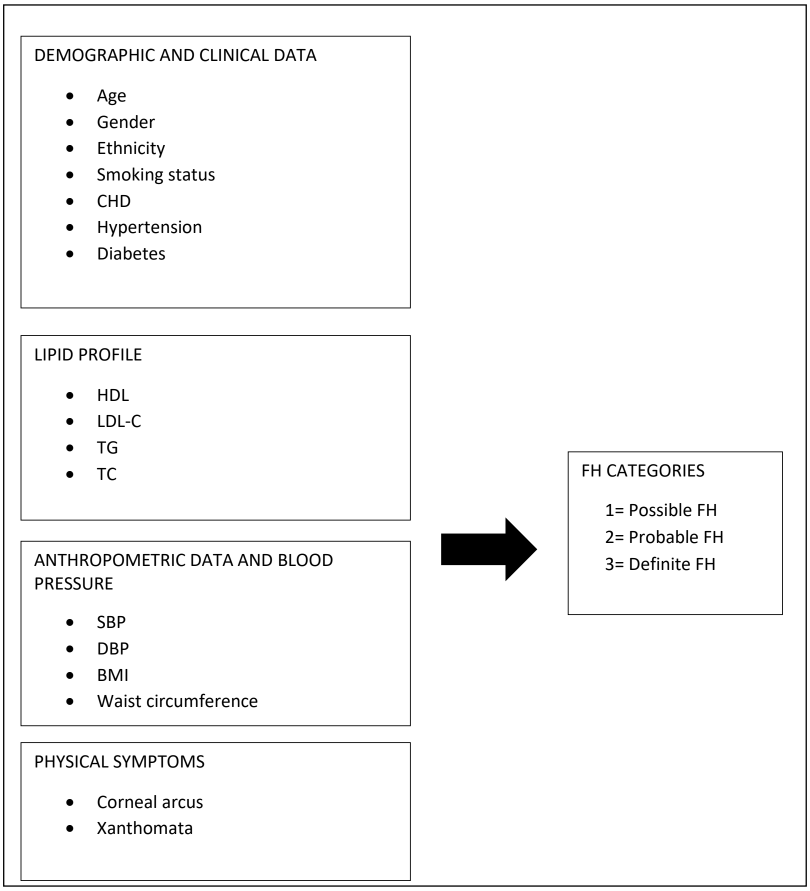

Figure 1. Theoretical framework

Notes. CHD = Coronary heart disease (CHD); HDL = High-density lipoprotein; LDL-C = Low-density lipoprotein cholesterol; TG = Triglyceride; TC = Total cholesterol; SBP = Systolic blood pressure; DBP = Diastolic blood pressure; BMI = Body mass index.

Table 1

Frequency distribution for demographic and clinical data, and physical symptoms variables

Type of Variable	Variable	Description of Variable	Frequency	Percentage (%)
Demographic and clinical data	Gender	Male	209	46.5
		Female	240	53.5

Table 1 *(Continued)*

Type of Variable	Variable	Description of Variable	Frequency	Percentage (%)
	Ethnicity	Malay	350	78.0
		Non-Malay	99	22.0
	Smoking	No	290	64.6
		Yes	144	32.1
	Hypertension	No	349	77.7
		Yes	77	17.1
	Diabetes	No	325	72.4
		Yes	25	5.6
	Coronary heartdisease (CHD)	No	401	89.3
		Yes	32	7.1
Physical symptoms	Corneal arcus (CA)	No	143	31.8
		Yes	84	18.7
	Xanthomata	No	136	30.3
		Yes	93	20.7

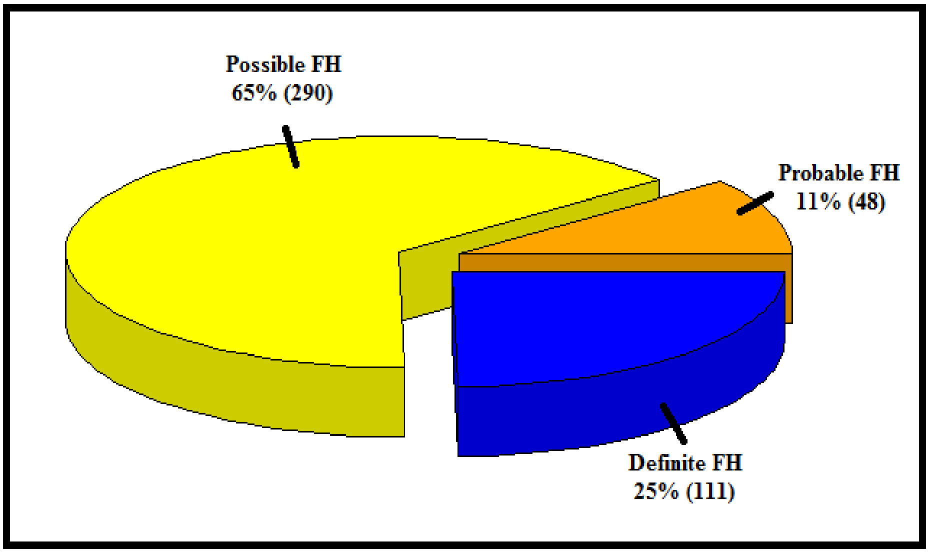

Figure 2. Distribution of FH category based on DLCN

Ordinal Logistic Regression

The ordinal logistic regression is appropriate to use when the dependent variable has ordinal categories. In this study, the dependent variable consists of three categories of FH. The ordinal categorical variables can have ordering from low to high or from high to low (Agresti, 2010). The model assumptions include no existence of multicollinearity and proportional odds to predict ordinal outcomes.

General Linear Model for Ordinal Data

The general linear model for ordinal data is based on the assumption that the dependent variableis in jth ordered categories according to link functions chosen (Hilbe, 2009; Norusis, 2005) (Equation 1):

$$link(\gamma_j) = \frac{\theta_j - [\beta_1 x_1 + \beta_2 x_2 + ... + \beta_k x_k]}{\exp(\tau_1 z_1 + \tau_2 z_2 + ... + \tau_m z_m)} \quad [1]$$

where γ_j is the cumulative probability for the jth category, θ_j is the threshold for the jth category, $\beta_1, \beta_2, ..., \beta_k$ are the regression coefficients, $x_1, x_2, ..., x_k$ are the predictor variables, k is the number of predictors, $\tau_1, \tau_2, ..., \tau_m$ are the coefficients for the scale component, $z_1, z_2, ..., z_m$ are m predictor variables chosen from the same set variables as the x's.

Build-in Link Functions

The ordinal regression model assumes that the slopes of several response levels are proportional (Hilbe, 2009). The different link functions for ordinal logistic regression is shown in Table 2. This study uses logit, probit and cauchit link only as complementary log-log link can only be applied if the chances of event is very small or very large (Allison, 1999). Negative log-log link can explain the data when the cumulative probability for lower group is high and it slowly approaches 1 (Norusis, 2005).

The cumulative probability of Casei for category j can be obtained as follows (Kutner et al., 2005) (Equation 2):

$$P(Y_i \leq j) = \frac{\exp(\alpha_i + (\beta_1 X_1 + \beta_2 X_2 + ... + \beta_k X_k)}{1 + \exp(\alpha_i + (\beta_1 X_1 + \beta_2 X_2 + ... + \beta_k X_k)} \quad [2]$$

RESULTS

Results in Table 3 showed that a higher percentage of non-Malays were found to be Definite FH. Among patients with corneal arcus, 53.6% were Definite FH while 95.7%

Table 2
Link Function for Ordinal Logistic Regression model

No	Binomial Link	Function	Typical application
1	Logit	$h\left(\frac{\gamma}{1-\gamma}\right)$	Evenly distributed categories
2	Complementary log-log	$\ln(-\ln(1-\gamma)$	Higher categories more probable
3	Negative log-log	$-\ln(-\ln(\gamma)$	Lower categories more probable
4	Probit	$\Phi^{-1}(\gamma)$	Analyses with explicit normally distributed latent variable
5	Cauchit	$\tan(\pi(\gamma-0.5))$	Outcome with many extreme values

Note. γ is probability of event *Y*.

with xanthomata are Definite FH. The results of Chi Square and Fisher's Exact Test in Table 4 showed there was a significant association between FH category (using DLCN) with smoking, corneal arcus and xanthomata at 5% significance level. Ethicity is significant only at 10% significance level.

The ordinal regression model analysis was carried out to determine the best classification model and to identify the demographic, lipid, medical and physical symptoms factors that were significantly associated with the dependent variable (FH Category). The models were evaluated using Akaike Information Criterion (AIC), NagelkerkeR-Square and classification accuracy. Table 5 presents the value for AIC, Nagelkerke R-Square and accuracy value for OLR (Logit), OLR (Probit) and OLR (Cauchit). Results in Table 5 showed that OLR (Cauchit) was the best model as it has the lowest AIC, highest NagelkerkeR-Square and classification accuracy value compared to Probit and Logit link function.

Next, we performed data imputation for missing values using mean imputation and treating of extreme values using the "Winsorisation process". Winsorization set all outliers (extreme values) to a specified percentile of the data. We performed 90% winsorization where all data below the 5th percentile were set to the 5th percentile, and data above the 95th percentile were set to the 95th percentile. The results for the treated dataset were compared with the original one to check whether there were any insights that could be obtained from untreated and treated datasets. The model compasrison results are presented in Table 6.

The results in Table 6 showed the OLR model was significant ($p < 0.05$). Additionally, the model fitted the data with Pearson chi-squared of 0.510 (C1) and 0.912 (C2). C1

Table 3

Frequency distribution of patients based on demographic & clinical data, and physical symptoms variables with respect to FH category

Type of Variable	Variable	Description of Variable	DLCN Diagnosis, n(%)			Total (n)
			Possible FH	Probable FH	Definite FH	
Demographic and clinical data	Gender	Male	144(68.9)	21(10)	44(21.1)	209
		Female	146(60.8)	27(11.3)	67(27.9)	240
	Ethnicity	Malay	232(66.3)	40(11.4)	78(22.3)	350
		Non-Malay	58(58.6)	8(8.1)	33(33.3)	99
	Smoking	No	171(59)	37(12.8)	82(28.3)	290
		Yes	110(76.4)	10(6.9)	24(16.7)	144
	Hypertension	No	240(68.8)	34(9.7)	75(21.5)	349
		Yes	44(57.1)	10(13)	23(29.9)	77
	Diabetes	No	242(74.5)	25(7.7)	58(17.8)	325
		Yes	19(76)	4(16)	2(8)	25
	Coronary heartdisease (CHD)	No	259(64.6)	44(11)	98(24.4)	401
		Yes	16(50)	4(12.5)	12(37.5)	32
Physical Symptoms	Corneal arcus (CA)	No	72(50.3)	8(5.6)	63(44.1)	143
		Yes	4(4.8)	35(41.7)	45(53.6)	84
	Xanthomata	No	76(55.9)	39(28.7)	21(15.4)	136
		Yes	0(0)	4(4.3)	89(95.7)	93

Table 4

Chi-Square test of association between FH category and covariates

Test used	Variable	Test value	P-value (α=0.05)	Association
Chi-squared Test	Gender	3.405	0.182	Not Significant
	Ethnicity	5.327*	0.070	**Significant**
	Smoking	12.825***	0.002	**Significant**
	Hypertension	3.848	0.146	Not Significant
	Corneal arcus (CA)	70.203***	<0.001	**Significant**
	Xanthomata	143.510***	<0.001	**Significant**
Fisher's exact test	Diabetes	3.178	0.194	Not Significant
	Coronary heartdisease (CHD)	3.179	0.189	Not Significant

p<0.10: **p-value<0.05;*p-value<0.01*

Table 5

Model comparisons: Ordinal Logistic Regression with different link functions

FH Instrument	DLCN		
OLR (Link)	Nagelkerke R^2	AIC	Accuracy
OLR (Logit)	0.577	497.2866	84.54%
OLR (Probit)	0.753	504.5605	84.54%
OLR (Cauchit)	**0.797**	**465.0818**	**85.01%**

was Cauchit Model for untreated data while C2 was Cauchit Model for treated data. The assumption of parallelism can be accepted as test of the parallel lines is not significant (C1: p-value =0.603, C2: p-value=0.655). The Nagelkerke's pseudo R^2 is higher for model C1 (0.797) compared with C2 (0.401). The significant variables based on C1 model are Diastolic Blood Pressure (DBP), Corneal Arcus (CA) and Xanthomata for Cauchit model (C1). Meanwhile, the Cauchit model (C2) consists of 11 significant variable: Total Cholesterol (TC), Triglycerides (TG), Low Density Lipoprotein (LDL), Diastolic Blood Pressure (DBP), Body Mass Index (BMI), Age, Corneal Arcus (CA), Gender , Ethnicity, Hypertension and CHD.

Ordinal Logistic Regression using Cauchitlink function for untreated data (C1) has higher Nagelkerke R-Square (0.797) compared to the model for treated data C2 (0.401). The model classification accuracy for C1(85.01%) is also higher compared to C2 (73.05%). It can be concluded that Cauchit model for untreated data (C1) performed better in classifying FH category. The final model with significant covariates is shown in Table 7 and some findings are as follows:

a) Patients who has high diastolic blood pressure is more likely to be classified into Definite FH [odds-ratio: exp(0.015) = 1.0151].

b) Patient who has no Corneal Arcus is less likely to be classified as Definite FH [odds ratio:exp(-8.648) = 0.0002]

c) Patients who do not suffer from xanthomatais less likely to be classified as Definite FH [odds ratio: exp(-13.063) = 0.000002].

Table 6

Ordinal logistic regression using Cauchit link function for untreated and treated data

Dataset	Untreated dataset	Treated dataset
Model	Cauchit Final Model (C1)	Cauchit Final Model (C2)
Model fitting information, χ^2	Chi-squared = 247.396 p-value <0.000	Chi-squared = 180.071 p-value <0.000
Pearson, χ^2	Chi-squared = 460.639 p-value = 0.510	Chi-squared = 822.933 p-value = 0.912
Deviance, χ^2	Chi-squared = 175.012 p-value = 0.378	Chi-squared = 598.353 p-value = 1.000
Test of parallel line, χ^2	Chi-squared = 3.639 p-value = 0.603	Chi-square = 56.178 p-value = 0.655
Cox and Snell	0.697	0.330
Nagelkerke	0.797	0.401
McFadden	0.574	0.231
Variable significant	Diastolic blood pressure (DBP) Corneal arcus (CA) Xanthomata	TC TG LDL-C DBP BMI Age CA

Table 6 *(Continued)*

Dataset	Untreated dataset	Treated dataset
Model	Cauchit Final Model (C1)	Cauchit Final Model (C2)
Variable significant	Diastolic blood pressure (DBP) Corneal arcus (CA) Xanthomata	CA Gender (Male) Ethnicity(Malay) Hypertension CHD
Total number of patients(n)	208	449

Notes. C1 (model for untreated data); C2 (model for treated data); TC= Total cholesterol; TG = Triglycerides; LDL-C = Low-density lipoprotein cholesterol; DBP = Diastolic blood pressure; BMI = Body mass index; CHD = Coronary heart disease.

Table 7

Parameter Estimates for Final OLR Model

Variables	Cauchit link function				
	Estimate	Standard Error	Wald	Degree of Freedom	P-Value
Constant1 (Possible)	-18.385	5.489	11.217	1	0.001
Constant2 (Probable)	-12.969	4.084	10.085	1	0.001
Diastolic blood pressure	0.015	0.029	0.263	1	0.008
Corneal arcus (No vs. Yes)	-8.648	2.648	10.669	1	0.001
Xanthomata (No vs. Yes)	-13.063	3.217	16.492	1	<0.001

*Note.*Since there are 3 categories there are 2 constant values (or intercepts).

DISCUSSION

This study seeked to classify FH patients in Malaysia using significant clinical and genetic factors related to FH. We used 449 patient data which were collected through health screening events, hospitals and clinics in Malaysia from 2010 to 2018. Results showed ethnicity smoking, and physical symptoms (corneal arcus and xanthomata) had significant association with FH. This finding show consistency with prior findings that indicated xanthomata and corneal arcus were significant factors of FH (Khoo et al., 2000; Michaelides et al., 2004; Scicali et al., 2018). This indicates that patients who have physical symptoms such as corneal arcus and xanthomata more likely (have higher likelihood) to to be classified as FH patient. Smoking by itself is not the cause of FH, but may reduce HDL level and increase the risk of CHD (Gepner et al., 2011). Significant univariate association between smoking and FH and smoking was probably due to the increased scoring of DLCN among smokers due to history of CHD.

Next, ordinal logistic regression models were developed for the classification of FH patients using DLCN categories as the dependent variable. The ordinal logistic regression with Cauchit link function performed better (highest Nagelkerke R-square and smaller AIC) compared to models with logit and probit link functions. The ordinal logistic regression with Cauchit link function had an accuracy of 85%. The model was also evaluated by splitting data into training (70%, n_{train}=146) and testing sample (30%; $_{ntest}$=62). The accuracy was 88.3% and 83.9% for training and testing sample respectively. This shows consistency in the model classification accuracy. The significant factors were only diastolic blood pressure, corneal arcus and xanthomata. These findings indicate that patients who have high diastolic blood pressure are more likely to be classified into Definite FH, while patients who do not suffer from corneal arcus are less likely to be classified as Definite FH.

The overall findings of the study are that patients with associated clinical FH factors such as diastolic blood pressure, corneal arcus and xanthomata have higher chances to be classified as Definite FH. Therefore, there is a need to create health awareness in the community focusing on FH risk factors. Awareness of FH hereditary disease can be educated to the public in primary care or specialist lipid clinics. Upon clinical diagnosis, family cascade screening should be recommended to the FH patients. General practitioners or even Lipid Specialists should be sufficiently trained to deliver effective consultation and management to FH patients. FH patients should be advised to undergo therapeutic lifestyle changes, including adjusting their dietand involve in physical activities.Educated FH patients who are aware of the CHD risk tend to practise healthy lifestyle, such as cease smoking (Razali et al., 2019). The establishment of Familial Hypercholesterolaemia Support Group, a community-level voluntary organisation, may provide education to the FH patients and their family members on how to live with FH, and morally support the FH patients to practise healthy lifestyle. The support group is being implemented in Australia, Canada and some European Countries (Payne et al., 2018; Watts et al., 2012), but currently,it is still not practised in Malaysia.

Treatment of FH incurs large financial burden to the patient and nation. Direct medical costs may include preventive, diagnostic, and treatment services which is related to FH. Meanwhile, indirect costs relate to morbidity and mortality costs including productivity. Productivity measures include 'absenteeism' (costs due to employees being absent from work for health-related reasons) and ‘presentism’ (decreased productivity of employees while at work) as well as premature mortality and disability. The developed OLR model highlighted the major observable clinical symptoms that significantly associated with FH in Malaysian population , thus allowing the clinicians to confidently screen the FH patients based on the symptoms, and treat the primary outcome of FH, which is CHD death, by advocating practise of healthy lifestyle.Thus, it is important if the trend or major groups who suffers this disease can be determined so that, they can benefit through the effortstaken by healthcare stakeholders to lower the risk of premature CHD-related death.An efficient statistical model allows clinicians to assess the risk (probability) of the patient in each FH category. Thus, patients with higher risk of being Possible FH should be advised to further clinical tests and healthcare treatments.

The limitation of this study is due to some numbers of missing data, an inherent problem suffered by secondary community data. This study found that the model fitted better with complete cases (i.e without data imputation or removing outliers). Future studies can also consider the new discrimination procedure proposed by Hamid and Hamid et al. (2018, 2018a) or treatement of outliers with new location model through the integration of Winsorization and smoothing approach (Hamid, 2018b). A larger sample with complete cases should be employed for future studies to confirm the findings of this study. Future studies will compare the models using different FH instruments for classification of FH.

CONCLUSION

This study successfully developed an OLR model for classification of Malaysian patients using a sample of secondary data. Ordinal logistic regression models with Cauchit link function for complete case was found to have higher Nagelkerke R-Square and classification accuracy. The final model results revealed that FH was found to be associated with diastolic blood pressure, corneal arcus and xanthomata. Ordinal logistic regression model is more appropriate when the dependent variable has ordinal categories. Taking into account the ordinal nature of the data, the probability obtained will be more accurate in the classification of patients into the three possible categories. The model is more informative than just modelling FH as a binary Yes or No category. It is imperative that effective early detection and timely control of FH must be strengthened in Malaysia to reduce disease burden. Future research can validate the statistical model by involving a larger sample of patients. Research should be ongoing to develop an efficient prediction model to assist clinicians in classification of FH patients. Early and timely detection is important for prevention of premature cardiovascular disease. Health education, intervention and preventive strategies of FH are highly important and more population-based disease screening for early detection and treatment should be carried out.

ACKNOWLEDGEMENT

This study was funded by UiTM MITRA Grants [(600-IRMI/MYRA 5/3/MITRA (003/2017)-1 and 600-IRMI/MYRA 5/3/MITRA (003/2017)-3].

REFERENCES

Abdul-Razak, S., Rahmat, R., Kasim, A. M., Rahman, T. A., Muid, S., Nasir, N. M., … & Ghani, R. A. (2017). Diagnostic performance of various familial hypercholesterolaemia diagnostic criteria compared to Dutch lipid clinic criteria in an Asian population. *BMC Cardiovascular Disorders*, *17*(1), 1-8.

Agresti, A. (2010). *Analysis of ordinal categorical data* (2nd Ed.).Hoboken, New Jersey; John Wiley & Sons.

Akioyamen, L. E., Genest, J., Shan, S. D., Reel, R. L., Albaum, J. M., Chu, A., & Tu, J. V. (2017). Estimating the prevalence of heterozygous familial hypercholesterolaemia: A systematic review and meta-analysis. *BMJ Open*, *7*(9), 1-13.

Allison, P. D. (1999). *Logistic regression using the SAS system: Theory and application*. Cary, NC: SAS Institute. Inc. and John Wiley and Sons.

Gepner, A. D., Piper, M. E., Johnson, H. M., Fiore, M. C., Baker, T. B., & Stein, J. H. (2011). Effects of smoking and smoking cessation on lipids and lipoproteins: outcomes from a randomized clinical trial. *American Heart Journal*, *161*(1), 145-151.

Hamid, H. (2018a). New location model based on automatic trimming and smoothing approaches. *Journal of Computational and Theoretical Nanoscience, 15*(2), 493-499.

Hamid, H. (2018b). Winsorized and smoothed estimation of the location model in mixed variables discrimination. *Applied Mathematics and Information Sciences, 12*(1), 133-138.

Hamid, H., Ngu, P. A. H., & Alipiah, F. M. (2018). New smoothed location models integrated with PCA and two types of MCA for handling large number of mixed continuous and binary variables. *Pertanika Journal of Science and Technology, 26*(1), 247-260.

Hilbe, J. M. (2009). *Logistic regression models*. Boca Raton, Florida: Chapman and hall/CRC.

Khoo, K., Van Acker, P., Defesche, J., Tan, H., Van de Kerkhof, L., Heijnen-van Eijk, S., … & Deslypere, J. (2000). Low-density lipoprotein receptor gene mutations in a Southeast Asian population with familial hypercholesterolemia. *Clinical Genetics*, *58*(2), 98-105.

Kutner, M. H., Nachtsheim, C. J., Neter, J., & Li, W. (2005). *Applied linear statistical models* (Vol. 5). New York, NY: McGraw-Hill Irwin.

Lye, S. H., Chahil, J. K., Bagali, P., Alex, L., Vadivelu, J., Ahmad, W. A. W., … & Mohamed, R. (2013). Genetic polymorphisms in LDLR, APOB, PCSK9 and other lipid related genes associated with familial hypercholesterolemia in Malaysia. *PloS One*, 8(4), 1-8.

Michaelides, A. P., Fourlas, C. A., Pitsavos, C., Andrikopoulos, G. K., Skoumas, I., Kartalis, A., … & Stefanadis, C. I. (2004). Exercise testing in asymptomatic patients with heterozygous familial hypercholesterolaemia. *Coronary Artery Disease*, *15*(8), 461-465.

Mundal, L. J., Igland, J., Veierød, M. B., Holven, K. B., Ose, L., Selmer, R. M., … & Leren, T. P. (2018). Impact of age on excess risk of coronary heart disease in patients with familial hypercholesterolaemia. *Heart*, *104*(19), 1600-1607.

Norusis, M. J. (2005). Ordinal regression. In M. J. Norusis (Ed.), *SPSS 13.0 Advanced Statistical Procedures Companion* (pp. 69-89). New Jersey, USA: Prentice Hall.

Payne, J., Williams, S., Maxwell, D., Pariente, M. T., Olivares, R. A., ten Haaf, M. J., … & Wong-Rieger, H. (2018). Familial hypercholesterolaemia patient support groups and advocacy: A multinational perspective. *Atherosclerosis*, *277*, 377-382.

Razali, S., Ismail, Z., Chua, Y. A., & Nawawi, H. M. (2019). Quality of life and its contributing factors in patients with familial hypercholesterolaemia in Malaysia. *Environment-Behaviour Proceedings Journal*, *4*(10), 86-92.

Scicali, R., Di Pino, A., Platania, R., Purrazzo, G., Ferrara, V., Giannone, A., … & Farruggia, E. (2018). Detecting familial hypercholesterolemia by serum lipid profile screening in a hospital setting: Clinical, genetic and atherosclerotic burden profile. *Nutrition, Metabolism and Cardiovascular Diseases*, *28*(1), 35-43.

Shi, Z., Yuan, B., Zhao, D., Taylor, A. W., Lin, J., & Watts, G. F. (2014). Familial hypercholesterolemia in China: Prevalence and evidence of underdetection and undertreatment in a community population. *International Journal of Cardiology*, *174*(3), 834-836.

Varghese, M. J. (2014). Familial hypercholesterolemia: A review. *Annals of Pediatric Cardiology*, *7*(2), 107-117.

Warner, P. (2008). Ordinal logistic regression. *Journal of Family Planning and Reproductive Health Care, 34*(3), 169-170.

Watts, G. F., Sullivan, D. R., van Bockxmeer, F. M., Poplawski, N., Hamilton-Craig, I., Clifton, P. M., … & Burnett, J. R. (2012). A model of care for familial hypercholesterolaemia: Key role for clinical biochemistry. *The Clinical Biochemist Reviews*, *33*(1), 25-31.

Weng, S., Kai, J., Tranter, J., Leonardi-Bee, J., & Qureshi, N. (2018). Improving identification and management of familial hypercholesterolaemia in primary care: Pre-and post-intervention study. *Atherosclerosis*, *274*, 54-60.

Wong, B., Kruse, G., Kutikova, L., Ray, K. K., Mata, P., & Bruckert, E. (2016). Cardiovascular disease risk associated with familial hypercholesterolemia: A systematic review of the literature. *Clinical Therapeutics*, *38*(7), 1696-1709.