

Rheumatics Heart Disease Using Stratified Cox Proportional Hazard Model with Time-Varying Covariate Effect

¹Nurhasniza Idham Abu Hasan, ¹ Nor Azura Md. Ghani, ¹Norazan Mohamed Ramli,
²Khairul Asri Mohd Ghani and ³Khairul Izan Mohd Ghani

¹Center for Statistical and Decision Sciences Studies,
Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA,
40450 Shah Alam, Selangor, Malaysia

²Department of Surgery, Faculty of Medicine and Health Sciences, Universiti Putra Malaysia,
43400 Serdang, Selangor, Malaysia

³Department of Surgery, Clinical Block, Universiti Kebangsaan Malaysia Medical Centre,
Jalan Yaakob Latif, Bandar Tun Razak, 56000 Cheras, Kuala Lumpur, Malaysia

Abstract: The most common way of analyzing prognostic factors in clinical data is by using Cox Proportional Hazard (PH) model. It is a well-recognized statistical technique for exploring the relationship between the survival of patient and several explanatory variables. The proportionality of the hazards is a critical assumption in the PH analysis and implies that the influence of covariates effect remains similar over time. However, this assumption is often violated. Therefore, different model should be used to deal with non-proportionality of hazards assumption. The aim of this study was to compare the Cox PH Model with non-PH Model as well as to identify the risk of death among Rheumatic Heart Disease (RHD) patients. In this study, we used secondary data in which a retrospective cohort study of 721 RHD patients that were obtained from University Kebangsaan Malaysia (UKM) and National Heart Institute (Institut Jantung Negara, IJN), Malaysia. Both Stratified Cox PH with and without non-PH covariate interaction were performed as non-PH Models. By using Akaike's Information Criterion (AIC) and Deviance, the efficiency of the model performance were compared and then, the most suitable model was determined. Based on these values, the stratified Cox PH Model with no-interaction is the best model for RHD dataset. Five statistically significant prognostic factors that contribute to the risk of death among RHD patients were identified, namely those who diagnosed as emergency status (Opstatus), performed with mitral valve repair alone have Hypertension (HPT), redo operation (Opepisode) and had longer Coronary Pulmonary Bypass (CPB). The non-PH Models fit better than Cox PH Model with respect to the lowest for both deviance and AIC values. This stratified Cox PH Model serves as an alternative approach that can cope with the non-PH situation.

Key words: Cox proportional hazard model, non-proportional hazard, rheumatic heart disease, stratified Cox proportional hazard model, time-dependent covariate effect, situation

INTRODUCTION

Survival analysis is a statistical tool that specifically designed for survival data in which measures duration of time taken for each subject experiences an event of interest. Moreover, survival data has some unique features that are difficult to handle with other statistical methods: censoring and time-dependent covariates (Ata and Sozer, 2007). Multiple regression models for survival data has widely used based on Cox Proportional Hazard (PH) Model which strictly depends on upon

proportionality hazard assumption assumes that the underlying hazard function for any two levels of some covariates are proportional over time or imply that the effect of given covariate does not change over time (Ata and Sozer, 2007).

Although, Cox PH Model is the most popular model to analyze the effects of covariates on survival time but in most medical studies, the assumptions of PH are often violated. This situation has occurred when the relationship between independent variable and dependent variable changes over time rather than remain

Corresponding Author: Nurhasniza Idham Abu Hasan, Center for Statistical and Decision Sciences Studies,
Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, 40450 Shah Alam,
Selangor, Malaysia

constant. Then, the Hazards Rates (HR) are no longer proportional. It is important to note here that when time-dependent variable is introduced into the model, the HR will not remain steady, i.e, they do not have parallel lines. For such situations, the effect of HR for any two level of covariates vary over time and then cross. Therefore, if HR varies with time, then the assumption of PH does not assume proportionality known as the non-PH Model to investigate the effects of covariates on survival time.

The identification of risk or prognostic factors related to the development of the disease has become equally important in recent years. Therefore, the aim of this study was to compare the Cox PH Model with the non-PH Model as well as to identify the risk of death among Rheumatic Heart Disease (RHD) patients.

MATERIALS AND METHODS

The data was a retrospective cohort study based on 721 RHD patients that were registered at University Kebangsaan Malaysia (UKM) and National Heart Institute (Institut Jantung Negara, IJN), Malaysia. In this study, we used secondary data which was collected from patient follow-up records. Information for this study was extracted from documents of all RHD cases registered from 1st January, 1992 to 31st December 2011.

The aim of the study was to determine the risk of death among RHD patients. In the study, the following information were included: Hypertension (HPT) (1 = yes, 0 = No), Coronary Pulmonary Bypass (CPB) in minutes, Mitral valve repair procedures (1 = Combination repair with other procedures, 0 = repair alone), length of days of Hospital stay (IHOSP) in days, Intra-Operative status (Opepisode) (1 = elective, 2 = emergency, 3 = urgent), Post-Operative status (Opstatus) (1 = first operation, 0 = redo).

Statistical analysis: Tolosie and Sharma (2014) suggested that start the variable selection procedures by including all variables that are significant in the univariable analysis at the 20-25% level and also any other variables which are presumed to be important to fit the initial multivariable model.

The multivariate analysis was carried out: Model 1: Cox PH Model (ignoring the PH assumption) and Model 2: non-PH Models namely stratified Cox PH with and without interaction of non-PH covariate models were performed (consideration for the variables do not satisfy the PH assumption). Then, all the models were compared and the best model was determined based on the lowest value for both AIC and deviance. Statistical significance was defined with the corresponding $p < 0.05$.

Assessment of proportional hazards assumption: There are a few assessments for checking the PH assumption such as; Plotting Kaplan-Meier curves for each level of the categorical variable. The PH assumption was met, if the Kaplan-Meier survival curves versus time or $\log(-\log(S(t)))$ versus $\log(t)$ shows that they do not cross or parallel lines.

Statistical tests based on the Scaled Schoenfeld residuals (Grambsch and Therneau, 1994). There is strong evidence for proportionality as shown by insignificant global test statistics ($p > 0.05$). The idea behind this test is if the PH assumption is fulfilled, then the residuals should not be correlated with survival time (or ranked survival time).

Non-proportional hazards model: In the violation PH assumption, the extension of Cox PH Model was required to take into account the non-PH situation. This done by using stratified Cox PH Model which come from the modification of Cox PH Model. There are two models under the concept of the stratified Cox PH Model (Kleinbaum and Klein, 2006) namely stratified Cox PH with and without interaction of non-PH covariate models. Using these models, a covariate that not satisfies PH assumption is not included in the model while, the covariates that satisfy the PH assumption was included in the model.

RESULTS AND DISCUSSION

Model 1; Cox proportional hazard model: In the first model, Cox PH Model was considered to model RHD data by ignoring the verification of PH assumption. To identify the risk of death that influences the survival time among RHD patients, the insignificant variables are dropped one by one until the final model observed which consists of all the significant variables at $p < 0.05$.

By applying forward stepwise method, the full multivariate Cox PH Model (Table 1) contains six statistically significant prognostic factors namely HPT ($p = 0.024$), Intra-Operative status ($p = 0.003$), Mitral valve repair procedures ($p = 0.039$), CPB ($p = 0.000$), HOSP ($p = 0.030$) and post-operative status ($p = 0.004$).

Assessing non-proportionality; Statistical testing strategies: Table 2 presents schoenfield residual to test the PH assumption. The global test reveals that there is strong evidence that the PH assumption has been violated by this data ($p < 0.05$). This results confirm that the HOSP ($\rho = -0.471$, $p = 0.000$) do not satisfy the PH assumption. Therefore, it can be concluded that the effect of HOSP change over time.

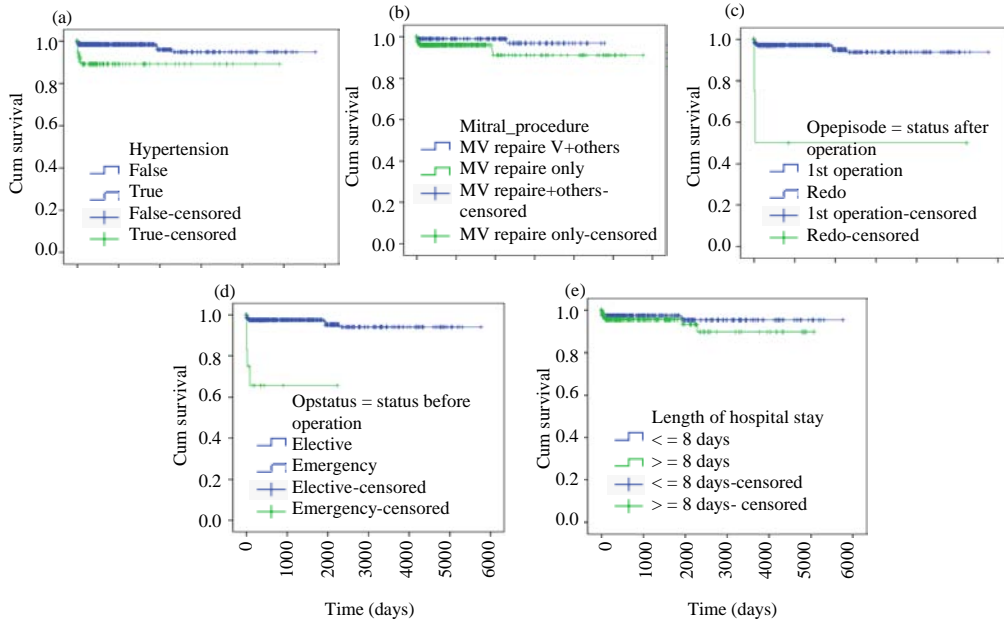


Fig. 1: Kaplan-Meier for: a) Hypertension; b) Mitral valve procedures; c) Post-Operative status; d) Intra-Operative status and e) Hospitalization (survival functions)

Table 1: Multivariate Cox PH Model

Variables	Regression coefficient (B)	HR	p-values
HPT			
(Yes vs. No)	1.031	2.803	0.024
CPB	0.011	1.012	0.000
Intra-Operative status			0.003
(Emergency vs. elective)	1.910	7.315	0.001
(Urgent vs. elective)	-9.488	0.000	0.986
Mitral valve procedures			
(Repair only vs. combination)	1.332	3.788	0.039
IHOSP	0.045	1.046	0.030
Post-operative status			
Redo vs. first operation	2.207	9.089	0.004

Table 2: Test for the proportionality assumption

Variables	Rho	p-values
HPT	-0.121	0.495
CPB	0.058	0.776
Intra-operative status		
(Emergency vs. elective)	-0.146	0.364
(Urgent vs. elective)	0.355	1.000
Mitral valve procedures	0.048	0.811
IHOSP	0.471	0.000
Post-operative status	-0.208	0.258
Global	NA	0.032

Assessing non-proportionality; Graphical strategy: The Kaplan-Meier survival curves appear to steadily drift apart for all categorical covariates as display in Fig. 1a-d. Again, this finding is also supported by the statistical testing strategy. We conclude that the PH assumption is fulfilled for all categorical covariates.

By examining the PH assumption for continuous variables, we should categorize these variables into a smaller number of categories, for example, two or three to

provide a reasonable balance of numbers (Kleinbaum and Klein, 2006). In this data, the HOSP variable needed to be categorized into; ≤ 8 days, more than 8 days. Again, after the HOSP variable is categorized into two different categories, then the Kaplan-Meier curve were plotted. There seems to be a slight downward trend from left to right for this variable (Fig. 1e).

Model 2; Stratified Cox PH Model: For the second model, the stratification Cox PH Model is used which requires stratification (categorized) over covariates that do not satisfy the PH assumption. There are two models under the concept of the stratified Cox PH Model namely stratified Cox PH with and without interaction of non-PH covariate models (Kleinbaum and Klein, 2006).

The Likelihood Ratio Test (LRT) is used to determine which model is performed better. Under the null hypothesis H_0 : the no-interaction model is correct and statistically appropriate.

Our test shows that the Stratified Cox PH Model without interaction of HOSP variable was adequate with a significant $p < 0.05$ level (Table 3). Then this non-PH model was compared via. Cox PH Model to determine the best model for non-PH situations.

Model comparison: By using AIC and deviance, the models obtained via. Cox PH and non-PH Models were compared and the most suitable model was determined in Table 4.

Table 3: ANOVA table

Model	Log-likelihood	χ^2	df	$p > \chi^2$
1	-89.541	-	-	-
2	-87.354	4.3742	12	0.9757

Log-likelihood for no-interaction model = -87.354; Log-likelihood for interaction model = -89.541

Table 4: 2-logL and AIC values of the Cox and stratified model

Parameters	Cox PH Model	Stratified Cox PH Model
AIC	248.77	117.38
-2LogL	241.43	101.22

Table 5: Final model using stratified Cox PH with no-interaction model by stratified HOSP

Variables	Regression coefficient (B)	HR	p-values
HPT (Yes vs. No)	1.507	4.514	0.002
CPB	0.009	1.009	0.000
Intra-operative status (Emergency vs. elective)	1.666	5.291	0.006
(Urgent vs. elective)	-8.443	0.000	0.986
Mitral valve procedures (Repair only vs. combination)	1.329	3.776	0.038
post-operative status (Redo vs. first operation)	2.904	18.245	0.001

HOSP, recode length of hospital stay

The results revealed that the non-PH Models were the most appropriate one under the non-PH situation. This indicates that stratified Cox PH Model without interaction of HOSP fits much better than Cox PH Model in the presence of non-PH situations.

Table 5 shows that when the variable HOSP is controlled, it has been found that the RHD patient with HPT has significantly higher risk of death compared to those patients who do not have HPT among RHD patients. For those who were diagnosed as emergency status had significantly higher risk of death compared to those diagnosed as elective (HR = 5.291, p = 0.006). Patients with mitral valve surgery repair alone had 4 times higher risk of death than those with a combination of mitral valve repair (HR = 3.776, p = 0.038). Having experienced surgery are 18 times more likely to face the risk of death compared to those who do not undergo the surgery (HR = 18.245, p = 0.001). Each minute increases in CPB, it will increase the risk of death by 1% while all other variables are held constant (HR = 1.009, p = 0.000).

In most medical studies, the non-PH is very common in modelling multiple covariates using Cox PH Model and detected the presence of time-varying covariate effect. This situation occurs when the effect of the covariate varies over time. Wilson states in his study that fitting the Cox PH Model for non-PH data would lead to incorrect conclusions. Moreover, it may entail serious bias and loss of power when making estimates and inferences about the effects of a given covariates on the risk of death (Ata and Sozer, 2007; Kleinbaum and Klein, 2006). Modelling

survival data that ignores the non-PH situation can lead to incorrect results. Therefore, when the PH assumption does not hold, the non-PH Models could be used as substitution for Cox PH Model and could lead to acceptable conclusions. This is important as an alternative approach which can cope with the non-PH situation.

The assessment of the PH assumption become the central theme in survival analysis and discussed in several statistical textbooks (Kleinbaum and Klein, 2006) as well as in the general statistical literature (Ata and Sozer, 2007; Pelissier *et al.*, 2008). More importantly, this assumption does not seem to be systematically assessed (Pelissier *et al.*, 2008; Altman *et al.*, 1995). Therefore, one should first check the PH assumption before applying the Cox PH Model to survival data. The stratified Cox PH Model was used to allow for time-varying effects in the model. This model comes from the extension of the Cox PH model (Vassileva *et al.*, 2013). The present study confirms that under non-PH situations, non-PH Model such as stratified Cox PH Model is better than the Cox PH Model. This finding is similar to the other studies (Ata and Sozer, 2007; Ahmed *et al.*, 2007; Sayehmiri *et al.*, 2008; Bellera *et al.*, 2010).

The present study also indicates that there is evidence of marked departures from PH assumption for IHOSP. This variable can be considered as time-varying covariates, since, the covariate varies over time. The incorporation of time-varying covariates effect is one of the strengths of stratified Cox PH Model over Cox PH Model that takes into account the subjects whose the effect for the particular covariate vary over time. In the stratified Cox PH Model, the effects of covariate depend on strata where constant within strata but non-constant HR between strata. From that it is clear that, through a single value estimated HR derived from Cox PH Model would influence other prognostic factors and these findings will always mislead.

In a previous study, Cox PH Models were used to find the risk of death among RHD after undergoing mitral valve repair (Gao *et al.*, 2009; Yakub *et al.*, 2013; Chan *et al.*, 2011; Vassileva *et al.*, 2013). A recent study indicates the effect of that covariate diminishes over time known as time-varying covariate. Inherently, the Cox PH model did not show any significant effect on this covariates but a graphical and numerical test provided that there was strong evidence of this variables had a non-constant effect over time. Since, PH assumption was not met for this variables, then the stratified Cox PH Model was carried out to stratify the HOSP variable. Based on stratified Cox PH Model without interaction,

there were five statistically significant risks of death among RHD patients were identified namely for those who diagnosed as emergency status, performed with mitral valve repair alone have HPT, redo operation and had longer CPB. Our findings were also confirmed by other previous studies (Yakub *et al.*, 2013; Chan *et al.*, 2011; Vassileva *et al.*, 2013).

CONCLUSION

In this analysis, the non-PH Models fit better than Cox PH Model with respect to the lowest for both AIC and deviance values and the stratified Cox PH Model with no-interaction was determined as the best model. Therefore, when the PH assumption does not hold, the non-PH Models could be used as substitution for Cox PH Model and also could lead to acceptable conclusions. The risk of death among RHD were namely for those who diagnosed as emergency status, performed with mitral valve repair alone have HPT, redo operation and had longer CPB.

ACKNOWLEDGEMENTS

We would like to dedicate our appreciation and gratitude to University Kebangsaan Malaysia (UKM) and National Heart Institute (Institut Jantung Negara, IJN), Malaysia for providing RHD data. Special thanks also goes to Universiti Teknologi MARA (UiTM) for supporting this study under the Research University Grant No. 600-RMI/DANA 5/3/CIFI (65/2013).

REFERENCES

Ahmed, F.E., P.W. Vos and D. Holbert, 2007. Modeling survival in colon cancer: A methodological review. *Mol. Cancer*, 6: 6-15.

Altman, D.G., B.L. De Stavola, S.B. Love and K.A. Stepniowska, 1995. Review of survival analyses published in cancer journals. *Br. J. Cancer*, 72: 511-518.

Ata, N. and M.T. Sozer, 2007. Cox regression models with nonproportional hazards applied to lung cancer survival data. *Haceteppe J. Math. Stat.*, 36: 157-167.

Bellera, C.A., G. MacGrogan, M. Debled, C.T.D. Lara and V. Brouste *et al.*, 2010. Variables with time-varying effects and the Cox model: Some statistical concepts illustrated with a prognostic factor study in breast cancer. *BMC. Med. Res. Method.*, 10: 1-20.

Chan, V., M. Ruel and T.G. Mesana, 2011. Mitral valve replacement is a viable alternative to mitral valve repair for ischemic mitral regurgitation: A case-matched study. *Ann. Thoracic Surg.*, 92: 1358-1366.

Gao, C., M. Yang, Y. Wu, G. Wang and C. Xiao *et al.*, 2009. Hybrid coronary revascularization by endoscopic robotic coronary artery bypass grafting on beating heart and stent placement. *Ann. Thoracic Surg.*, 87: 737-741.

Grambsch, P.M. and T.M. Therneau, 1994. Proportional hazards tests and diagnostics based on weighted residuals. *Biometrika*, 81: 515-526.

Kleinbaum, D.G. and M. Klein, 2006. *Survival Analysis: A Self-Learning Text*. Springer Science & Business Media, Berlin, Germany.

Pelissier, S.M., S.G. Bourgade, F. Bonnetain and A. Kramar, 2008. Survival end point reporting in randomized cancer clinical trials: A review of major journals. *J. Clin. Oncol.*, 26: 3721-3726.

Sayehmiri, K., M.R. Eshraghian, K. Mohammad, K. Alimoghaddam and A.R. Foroushani *et al.*, 2008. Prognostic factors of survival time after hematopoietic stem cell transplant in acute lymphoblastic leukemia patients: Cox proportional hazard versus accelerated failure time models. *J. Exp. Clin. Cancer Res.*, 27: 1-9.

Tolosie, K. and M.K. Sharma, 2014. Application of Cox proportional hazards model in case of tuberculosis patients in selected Addis Ababa health centres, Ethiopia. *Tuberculosis Res. Treat.*, 2014: 1-11.

Vassileva, C.M., G. Mishkel, C. McNeely, T. Boley and S. Markwell *et al.*, 2013. Long-term survival of patients undergoing mitral valve repair and replacement: clinical perspective. *Circulation*, 127: 1870-1876.

Yakub, M.A., J. Dillon, P.S.K. Moorthy, K.K. Pau and M.N. Nordin, 2013. Is rheumatic aetiology a predictor of poor outcome in the current era of mitral valve repair? Contemporary long-term results of mitral valve repair in rheumatic heart disease. *Eur. J. Cardio Thoracic Surg.*, 44: 673-681.