Determinants of Marital Dissolution: A Survival Analysis Approach

A. Sanizah, F. Hasfariza, S. Norin Rahayu and A. Nur Niswah Naslina

***Abstract*—** The dissolution of marriage, also known as divorce, occurs when the bond of matrimony between a married couple is dissolved. Since the rate of divorce is on the rise, this study aims to identify potential determinants affecting the dissolution of marriage using survival analysis. We retrospectively studied 531 secondary data of the Muslim couples who filed for divorce in Selangor, Malaysia. The age at marriage of husband and wife, the presence of children, duration of marriage, couples' educational level and employment status, household income and counseling session were identified as potential determinants. The AIC (Akaike Information Criterion) were used to compare the efficiency of models between Weibull and Cox model. The Cox model had the best fit with respect to the lower AIC value. The survival result from the Cox model showed that age at marriage of husband and attending counseling session significantly affect marital dissolution.

***Keywords* —**Cox proportional hazards, divorce, Kaplan-Meier, survival time, Weibull

I. Introduction

MARRIAGE is a contract between a husband and wife and the contract should be made to work, but not when it becomes humanly impossible. Marriage is important as it gives the married couples happiness and prolongs their lives[1]. However, there are times when the marriage becomes almost impossible to continue. In the Islamic law, divorce is permissible when the relationship can no longer survive. Before proceeding with the divorce, couples are encouraged to go through counseling sessions organized by the Islamic Religious department in the hope for reconciliation. If it fails, they have the right to divorce as has been established in the Islamic Law. According to the court's procedures, those who applied for the divorce must have a follow up by the counseling report. If the husband fails to attend the counseling session for three times in a row, the wife has the right to request for the counseling report and she can proceed with the divorce to the *Syariah* court. However, there are cases where the husband or wife did not attend the counseling session because they managed to solve their marital problems themselves.

Our sincere appreciation is directed to Universiti Teknologi MARA Malaysia, the Ministry of Higher Education, Malaysia and JAIS. The research is funded by *Dana Pembudayaan Penyelidikan* (RAGS), 600-RMI/RAGS 5/3 (15/2012).

Sanizah Ahmad, Hasfariza Farizad, Norin Rahayu Shamsuddin & Nur Niswah Naslina Binti Azid @ Maarof are from Faculty of Computer and Mathematical Science, Universiti Teknologi MARA, Malaysia. (e-mail: sanizah@tmsk.uitm.edu.my)

About 1.2 million Muslim marriages were registered in Malaysia between year 2008 and 2012. However, 210,326 Muslim divorces were recorded over the same period, with Selangor topping the list with 28,570 cases [2]. [4] stated that age at marriage is significantly correlated with divorce. This statement is supported by [5] and she reported that in Malaysia apparently, many divorce cases are among senior citizens where divorce happens among 5 out of 10 senior couples. It is also discovered that the presence of children will avoid the couples from being separated [6]-[7]. A recent study also found that marriage is likely to be dissolved during the first 5 years of marriage [8]. Study by [6] found that factors of being educated and employed also contribute to a higher risk of divorce. [9] also revealed that higher total family income, and those who did not attend counseling session were more likely to proceed with the divorce.

The objective of this study is to identify potential determinants affecting to the dissolution of marriage using survival analysis approach.

II. Methodology

This study obtained its secondary data from one of the Islamic Religious Department in Selangor, a state in Malaysia. The dataset consists of 531 cases reported by couples who filed for a divorce for the year 2012. The determinants are divided into three main categories; demographic variables, socioeconomic variables and treatment variable as suggested by [6].

Demographic variables consist of the age at marriage of the husband and wife, the presence of children and duration of marriage. The age at marriage of the husband and wife are divided into eight categories which are less than 24 years old, 25-29 years, 30-34 years, 35-39 years, 40-44 years, 45-49 years, 50-54 years, and more than 55 years old. The presence of children is presented by two categories which are yes and no, and the duration of marriage is presented through four categories, which are very short (less than one year), short (one to five years), medium (five to ten years), and long (longer than ten years).

On the other hand, socioeconomic variables consists of the educational level of the husband and educational level of wife (primary school, secondary school, diploma, bachelor, and master), employment status of the husband and wife (employed, unemployed) and household income (less than RM2000, RM2000-RM3000, RM3000-RM4000, RM4000-RM5000, and above RM5000).

The treatment variable included in this study is the counseling session attended (yes, no) while the outcome variable is the survival time (in days). The survival time is calculated from the day the person or couple filed the case at the Islamic Religious Department until the day they made the decision on continuing with their divorcement or canceling it. R Programming was used to analyze the data.

A. Survival Analysis

Survival analysis is the study of durations between events such as duration of marriage [10] and unemployment duration [11]. This method is used to analyze data in which the time until the event is of interest. It deals with censored data in which the survival times are unknown. The censored data occurs when patients are still alive at the end of the study, withdraw before the end of the study or are lost to follow up. This is also known as right censoring. At times, analyses might also include left censoring due to its initial time at risk is unknown [12].

B. Univariate Analysis

It is highly recommended to look for Kaplan Meier curves for all categorical predictors, which will provide the shape of the survival function for each group and give some idea whether or not the groups are proportional (i.e. the survival functions are approximately parallel). Next, log-rank test will be carried out in order to determine whether the variables should be included in the Cox PH model [13]. It is also used to see the differences among the groups [14].

C. Cox Proportional Hazards (PH) Model

The Cox PH model is mainly used to assess the relationship of predictor variables such as age, gender, and the type of treatment to the survival time, T [15]. The proportional hazard relationship is:

$$\lambda(t;z) = \lambda_0(t)\cdot\exp(z\cdot\beta) \quad (1)$$

where $\lambda_0(t)$ is an arbitrary unspecified baseline hazard function for continuous time t and β is a vector of parameters. According to [14], tests and graphical diagnostics for proportional hazards may be based on the scaled Schoenfeld residual.

D. Weibull Model

Weibull model is a very popular as a failure time distribution because constant hazard cannot be assumed. It is also one of the most frequently used parametric models in failure time data analysis [16]. The proportional hazards model is:

$$\lambda(t;z) = \gamma\lambda(z)t^{\gamma-1} \quad (2)$$

A graphical check on the Weibull assumption can be made by a plot of ln[-lnS(t)] versus ln t , where S(t) is a survival estimate obtained from the Kaplan-Meier method.

E. Criterion Selection

The Akaike Information Criterion (AIC) was used to compare the efficiency of the models between Cox and Weibull model [17]-[19]. It is a measure of the goodness of fit of the model estimated and a practical way of showing the complexity of an estimated model against how well the model fits the data. The smaller the AIC the better it indicates the likelihood.

III. Results

A. Descriptive Analysis

According to Table 1, censored data are defined as 0 where the survival time is from the date the couple report the case until the case is close or the couple decides not to divorce or the couple delays the cases. Uncensored data are defined as status 1 indicating the survival time from the reported date until the date they decide to divorce. For example, 199+ means that the couple did not divorce within 199 days and was then unavailable for further study. This data are also classified as right censoring. A summary statistics is obtained from the data. On average, the survival time of marriage was 114.9 days. The minimum survival time was 1 day, which means the day the couples reported to file the case to the Islamic Religious Department was the same as the day they attended the counseling session and decided to file for divorce to the *Syariah* court. Meanwhile, the maximum survival time was 615 days. The standard deviation is 93.99922 days, indicating that the survival time varied a lot from one another.

Table I: Table of survival time with censoring data

Survival Time	Censor (0=Delay/Did not divorce/Case close, 1=Divorce)
199+	0
197+	0
207+	0
22+	0
258+	0
16	1
548+	0
615+	0
204	1
257+	0
209+	0
259+	0
208+	0
243+	0
10+	0
.	.
.	.
.	.

Fig. 1 illustrates the Kaplan-Meier (KM) graph of the survival times for married couples ensuring for censored or uncensored data in the study. The vertical dashes represent the censored items. The curve shows higher survival rate within the first 83 days of the study and the survival curve decreases gradually as the number of days increase. The curve decreases even faster after 83 days reaching its minimum time on day one.

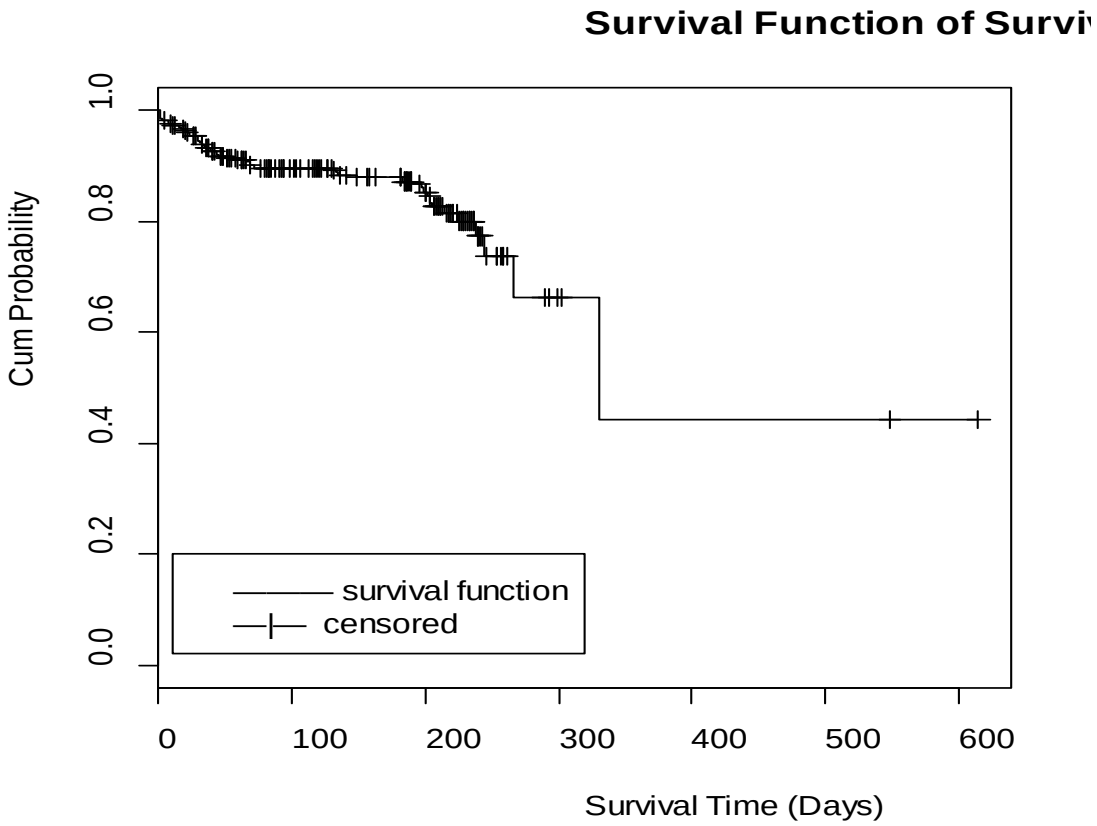

Fig. 1 Survival function of survival time of marriage

Fig. 2(a) shows the survival distributions for each factor in the demographic variables. The large diverge in the Kaplan-Meier curves only occur in the age at marriage of the husband. According to Figure 2(b), the curves for the education level of the husband, education level of the wife, and employment status of the husband show the curve increasingly diverge. There is a similar finding in the curve for the only treatment variable in this study, counseling session, as shown in Figure 2(c). Since these five variables give large impact to the survival time of marriage due to the divergence in the Kaplan-Meier curves, they may be included in the multivariate analysis. However, the final decision depends on the results of the log-rank test, which is tabulated in Table II. According to [13], the hypotheses being tested for each variable are:

H_0: There is no difference between survival curves
H_1: There is a difference between survival curves

The variables are included in the study if the p-value from the log-rank test is less than 0.05, which at 5 percent significant levels. From the table, only three variables are significant, the age at marriage of husband, education level of wife and attending counseling session and the p-values are 0.036, 0.0367 and 1.14e-07, respectively, which are less than 0.05 at 5 percent significant level. This value indicates that these three variables are the significant factors in giving more impact to the overall survival. Hence, this means that the survival times differ significantly between groups for the age at marriage of the husband, the education level of the wife, and attending counseling session. Therefore, only these 3 variables are considered as statistically significant and should be included in the multivariate model.

Table II. Log-rank Test Results

Factor	p-value
Demographic variables	
Age at marriage (husband)	0.036*
Age at marriage (wife)	0.41
Presence of children	0.422
Duration of marriage	0.292
Socioeconomic variables	
Education level of husband	0.642
Education level of wife	0.0367*
Employment status (husband)	0.934
Employment status (wife)	0.549
Household income	0.707
Treatment variable	
Attend Counseling Session	1.14e-07*

B. Multivariate Analysis

Before using the Cox PH model, it is necessary to check the underlying assumption of PH. Schoenfeld residuals show that all the potential determinants met the assumptions (not shown). Table III shows the results of AIC values and significant factors for the Cox PH and Weibull Model. It shows that both Cox PH model (Likelihood ratio test (10 d.f.) = 40.79, Wald test (10 d.f.)=40.04 and Weibull model (Chi-Square (10 d.f.)=36.44) are significant as a whole since their p-values are less than 0.05 at 5 percent significant levels. It also shows that the significant variables in the Cox PH model are the age at marriage of the husband (p-value=0.0429<0.05) and attending counseling session (p-value=3.49e-07<0.05). However, only attending counseling session is significant in the Weibull model since the p-value is 1.59e-06 which is less than 0.05 at 5 percent significant level. The results in Table 3 display that the Cox PH model has a lower AIC value (AIC=689.3905) compared to the Weibull model (AIC=979.8046). Therefore, this indicates that the Cox PH model produced better estimates compared to the Weibull model.

Table III. AIC values and significant factors for the Cox Proportional Hazards and Weibull Model

	Hazard Ratio	
Factors	**Cox** (AIC=689.3905) (Likelihood ratio test=40.79*) (Wald test=40.04*)	**Weibull** (AIC=979.8046) (Chi-Square=36.44*)
Demographic variables		
Age at marriage (husband)	0.2652*	-0.3284
Age at marriage (wife)	8.271e-05	-0.0129
Presence of children	-0.1424	0.1545
Duration of marriage	-0.2625	0.3238
Socioeconomic variables		
Education level of husband	0.02808	0.0214
Education level of wife	-0.2684	0.2815
Employment status (husband)	-0.1107	0.2119
Employment status (wife)	0.3941	-0.5713
Household income	0.09087	-0.1089
Treatment variable		
Attend Counseling session	1.479*	1.5958*

*Significant at 5% level

The global test in Table IV for the Cox PH model as a whole is not statistically significant (p-value>0.05). Therefore, there is a strong evidence of proportional hazards for the model as a whole. Hence, this supports the proportional hazards assumptions.

Table V shows the results of the model with main effects and model with the interaction effects. It shows that the age at marriage in Model 1 (p-value=0.00593) and attending the counseling session in Model 2 (p-value=5.91e-07) are significant since both p-values are less than 0.05 at 5 percent significant levels. Moreover, Model 3 which includes the age at marriage of husband and attending counseling session as their predictor variables, shows that both variables are significant since both p-values are lower than 0.05. However, in Model 4, the interaction between the age at marriage of husband and attending the counseling session as well as the main effect are not significant (p-value>0.05). It means that the interaction does not give impact to the model.

Table IV. Global test of Cox Regression model

Item	Chi-Square	p-value
Model (GLOBAL)	9.9418	0.2028

Table V. Results of multivariate Cox PH for four models

Model	Factors	β	p-value	AIC
1	Age at marriage (husband)	0.3280	0.0231*	704.9315
2	Attend Counseling session	-1.4147	5.91e-07*	687.4218
3	Age at marriage (husband)	0.3379	0.0219*	681.5687
	Attend Counseling session	-1.4223	5.27e-07*	
4	Age at marriage (husband)	0.29979	0.165	683.5657
	Attend Counseling session	-1.68527	0.137	
	(Age at marriage (husband))*(Attend Counseling session)	0.07135	0.810	

Furthermore, Model 3 has the lowest AIC value which is 681.5687, compared to the other model. Since the Model 3 has the lowest AIC value and all variables in its model are significant, therefore, the Model 3 is chosen as the final and best model. Model 3 is chosen to be the best model since it has the lowest AIC values compared to other models and all the variables included in Model 3 are significant. A positive sign indicates that the hazard (risk of divorce) is higher while a negative sign indicates that the hazard (risk of divorce) is lower.

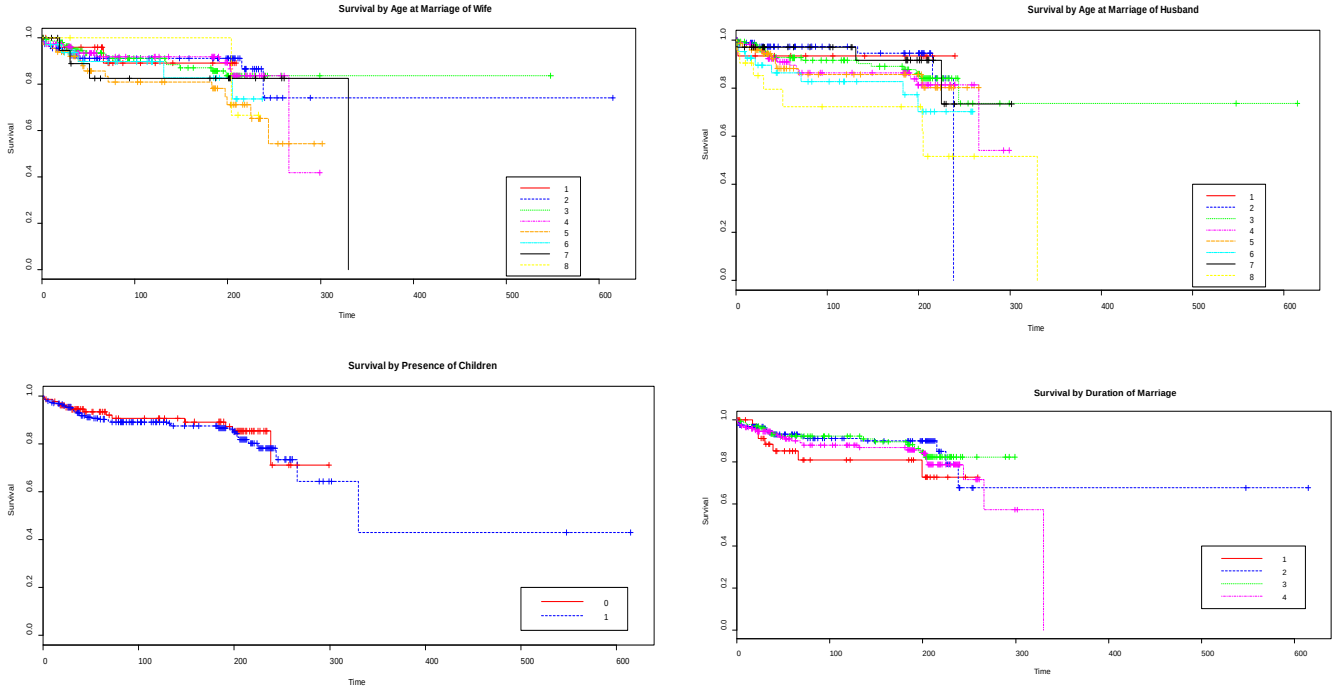

Fig. 2(a) Kaplan-Meier curves for different groups of variables (Demographic variables)

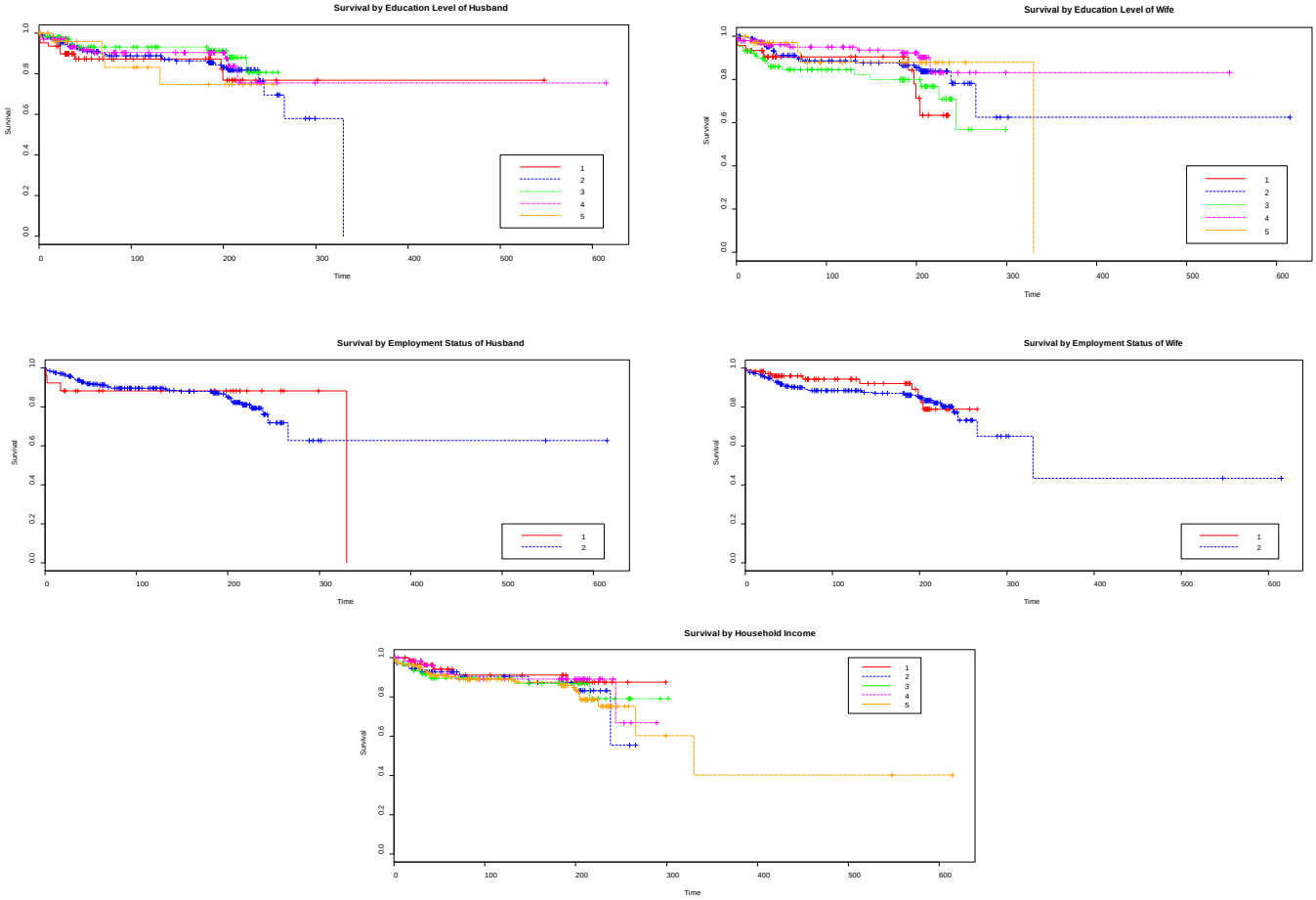

Fig. 2(b) Kaplan-Meier curves for different groups of variables (Socioeconomic variables)

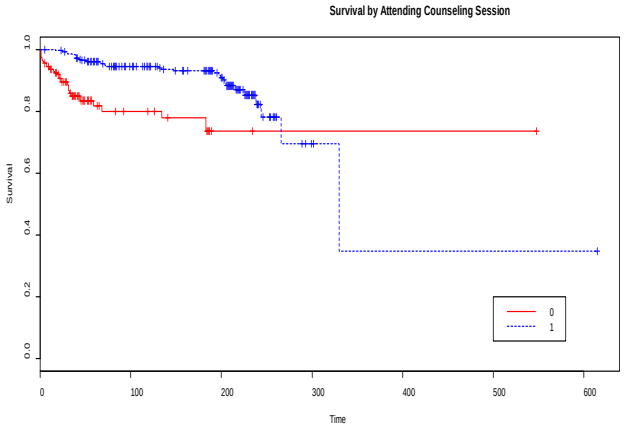

Fig. 2(c) Kaplan-Meier curves for different groups of variables (Treatment variable)

The proportional hazards relationship for this model is:

$$\lambda(t; z) = \lambda_0(t)\exp(0.19891_{\text{Age at Marriage of Husband}} - 1.44138_{\text{Attend Counseling Session}}) \quad (3)$$

where the baseline hazard function, unspecified. Based on Table VI, the estimate for β_1 had a positive value, meaning that the hazard ratio increases over time. It can be interpreted that for the covariate, age at marriage of the husband, the risk of divorce increases about 1.22 times as the age increases. The 95% confidence interval for the age at marriage of the husband did not include 1.000 (1.0647, 1.3981), suggesting that there is a difference in the survival of marriage between groups in this variable. On the other hand, the estimate for β_2 had a negative value indicating that the hazard ratio decreases over time. This means that the couples who attended the counseling session will reduce the risk of the divorcement by 0.23. The 95% confidence interval for attending a counseling session also did not include 1.000 (0.1351, 0.4144). This suggests that there is a difference in the survival of marriage between those who attended and who did not attend any counseling session.

Table VI. Parameter estimates of covariates, p-value and confidence interval for final model

Factors	β	exp(β) (Hazards ratio)	p-value	95% Confidence Interval
Age at marriage (husband)	0.19891	1.22007	0.0042	(1.0647,1.3981)
Attend Counseling session	-1.44138	0.23360	4.64e-07	(0.1351,0.4144)

IV. Conclusions

This study concludes that the Cox PH model is the best model in describing the divorce data compared to the Weibull model since it has the lowest AIC value. For both models, attending counseling session is found to be significant. However, in the Cox PH models, the age at marriage of the husband is also found to be significant. Since Cox PH model is the best model, hence, it can be concluded that the age at marriage of the husband and attending counseling session were included in the model for this study. Furthermore, this model also satisfied the proportional hazards assumptions.

References

[1] M. Gallagher, "The case for marriage," *Institute for American values*, vol. 1, no. 6, 2001.
[2] S.C.H., Chlen and M.S. Mustaffa, "Divorce in Malaysia," *Seminar Kauseling Keluarga*, pp. 23-28, 2008.
[3] Bernama"Divorce cases increases in the last 10 years", *Sinar Harian*. . (2012, June 11). Retrieved from http://www.sinarharian.com.my/nasional/kes-cerai-meningkat-dalam-10-tahun-lalu-1.55706?localLinksEnabled=false
[4] T.R. Balakrishnan, K.V. Rao, E. Lapierre-Adamcy and K.J. Krotki, "A hazard model analysis of the covariates of marriage dissolution in Canada," *Demography*, vol. 24, no. 3, pp. 395-406, 1987.
[5] 'Kes cerai pada usia emas', *Utusan Malaysia*, 2010
[6] F.P. Menard, "What makes it fall apart? Determinants of the dissolution of marriages and the common-law unions in Canada," McGill Sociological Review, vol. 2, no. 4, pp. 59-76, 2011.
[7] J. Vazquez, 'Using survival analysis methods to study Santa Barbara County divorces', California Polytechnic State University, 2011.
[8] N.K. Frempong, C. Osei-Mensah, D. Asamoah Owusu, and E. Okyere, "Survival analysis on marriage and divorce in the Metropolis," *Canadian Journal on Computing in Mathematics, Natural Sciences, Engineering and Medicine*, vo. 3, no. 5, pp. 159-163, 2012.
[9] S. Ahmad, and M.R.A. Bakar, "Predictors of the risk of divorce in Muslim marriages," paper presented at the Proceedings of ICREM, 2003.
[10] A. Sanizah, S. Norin Rahayu and F. Hasfarizah, "Determinants of marital dissolution: A cox regression model", *Proceeding of the 16th WSEAS International Conference on Mathematical and Computational Methods in Science and Engineering (MACMESE'14)*, 2014.
[11] V. Ciuca and M. Matei, "Survival analysis for the unemployment duration", Proceeding of the 5th *WSEAS In ternational Conference on Economy and Management Transformation (Volume 1)*, 2010.
[12] S. Ahmad *Analysis of muslim marriages using survival data-A case study.* Unpublished dissertation Degree of Master of Science (Applied Statistics), University Putra Malaysia, 2002.
[13] D.G. Kleinbaum, 'Survival analysis' New York: *Springer*, 2nd (edn.) 1996.
[14] J. Fox, 'Cox proportional-hazards regression for survival data', Appendix to an R and S-PLUS companion to applied regression (Sage Publication,) 2002.
[15] S.S. Halli and K.V. Rao, 'Demographic models', in Land, K.C. (Ed.): *Advanced techniques of population analysis*, Plenum Press, pp. 224, 1992.
[16] E. A. Hayat, A. Suner, B. Uyar, O. Dursun, M.N. Orman, and G. itapcioglu, 'Comparison of five survival models: breast cancer registry data from Ege University cancer research center', *Biostatistics*, vol. 30, no. 5, 2010.
[17] R. Ravangard, M. Arab, A. Rashidian, A. Akbarisari, A. Zare, and H. Zeraati, 'Comparison of the results of cox proportional hazards model and parametric models in the study of length of stay in a tertiary teaching hospital in Tehran, Iran', *Acta Medica Iranica*, vol. 49, no. 10, pp. 650-658, 2011.
[18] H.P. Zhu, X. Xia, C.H. Yu, A. Adnan, S.F. Liu, and Y.K. Du, 'Application of Weibull model for survival of patients with gastric cancer,' *BMC Gastroenterology*, vol. 11, no. 1, 2011.
[19] Y.H. Chan, 'Biostatistics 203. Survival analysis', Basic statistics for doctors, vol. 45, no. 6, pp. 249, 2004.

Sanizah Ahmad is a senior lecturer at the Faculty of Computer & Mathematical Sciences, Universiti Teknologi MARA Shah Alam, Malaysia. She is also a Life member of the Malaysia Institute of Statistics. Her research interests are survival analysis, robust methods and logistic regression.

Hasfariza Farizad is a post graduate student from the Faculty of Computer & Mathematical Sciences, Universiti Teknologi MARA Shah Alam.

Norin Rahayu Shamsuddin is a senior lecturer at the Faculty of Computer & Mathematical Sciences, Universiti Teknologi MARA Kedah Campus, Malaysia. She is also a Life member of the Malaysia Institute of Statistics and has published a number of articles related to medical area.

Nur Niswah Naslina Azid @ Maarof is a lecturer at the Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, Malaysia. She has published articles related to Health Related Quality of Life (HRQoL).