Cox Proportional Hazards Model for Identification of the Prognostic Factors in the Survival of Acute Liver Failure Patients in India

Shankar Prasad Khanal[1*], V. Sreenivas[2] & S. K. Acharya[3]

Submitted: 25 August 2018; Accepted: 06 September 2018

Published online: 26 September 2018 **DOI: http://dx.doi.org/10.3126/njs.v2i0.21155**

ABSTRACT

Background: Acute Liver Failure (ALF) is a kind of dangerous rare liver injury among all liver diseases. Different statistical methods such as Logistic regression, Kaplan-Meier estimate of survival function followed by Log-rank test and semi-parametric approaches of survival analysis has been applied in order to identify the significant risk factors of ALF patients. In most of the studies, regression models used in this setup has not been evaluated by model assumptions and their goodness of fit tests.
Objective: To apply appropriate survival analysis technique to identify the prognostic factors in the survival of ALF patients, to develop prognostic index, and to predict survival probability for different scenario.
Materials and Methods: The study is based on the retrospective cohort study design with altogether 1099 ALF patients taken from the liver clinic, All India Institute of Medical Sciences, New Delhi India. Cox regression has been considered as the suitable model for handling this time to event data, and the assumptions of the model, goodness of fit of the model was assessed and survival probabilities were predicted.
Results: This study has identified six prognostic factors namely age, prothrombin time, cerebral edema, total serum bilirubin, serum creatinine and etiology for ALF patients. The hazards of mortality [HR: 2.38; 95% C.I.: (1.99, 2.85), $p < 0.001$] is the highest for cerebral edema among all these prognostic factors. Nearly 9%, 26%, 39%, 50%, 59% and 63% of ALF patients with a *PI* of 1, 3, 5, 7, 9 and 10 respectively die by 3 days of hospital stay.
Conclusion: The developed Cox Proportional Hazards model with six prognostic factors has satisfied the model assumptions and goodness of fit tests. The risk score and the predicted survival probabilities will be immensely helpful to the hepatologists to make a quick decision regarding the likely prognosis of a patient at admission and helpful in triaging the ALF patients for liver transplant.

Keywords: Acute liver failure, Cox-Snail residuals, hazard ratio, prognostic index, proportionality of hazards, Schoenfeld residuals.

Address correspondence to the author: Central Department of Statistics, Tribhuvan University, Kirtipur, Kathmandu, Nepal. Email: drshankarcds@gmail.com[1*] (Corresponding author email); Department of Biostatistics, All India Institute of Medical Sciences (AIIMS), New Delhi India, sreevishnubhatla@gmail.com[2]; Department of Gastroenterology, All India Institute of Medical Sciences (AIIMS), New Delhi India, subratacharya@gmail.com

INTRODUCTION

Acute Liver Failure (ALF) is characterized by severe and sudden liver cell dysfunction leading to coagulopathy and hepatic encephalopathy in previously healthy persons with no known underlying liver disease (Trey & Davidson, 1970). Existing literature shows that very few studies have been carried out on ALF from the Indian subcontinent. Because of rarity of ALF and the unavailability of data, there has not been a comprehensive description of ALF from the developing world. Majority of the reports on ALF were from the United Kingdom, United States of America, France and Japan. Generally Western physicians assume that what is true about a disease in the Western hemisphere must be universal; that is if a certain disease is described in Europe, the United Kingdom, or North America, that description is applied to the rest of the world. But unfortunately this has not been true always (Acharya et al., 1996; Lee & Sorrell, 1996). There may be different disease characteristics for the same disease in different geographic regions and the real characteristics of ALF of developing countries like India must be addressed separately. For instance, overdose of drugs is one of the major causes of ALF found in the United Kingdom whereas hepatitis viruses are found major causes in India. Because of differences in etiologies of ALF, the prognostic factors might be different. Naturally, management strategy to be implemented for ALF patients also depends on etiology of patients. Therefore, there is a strong need to identify prognostic factors of ALF patients especially in India where hepatitis virus is the major etiology.

Identification of prognostic factors is generally done through specific statistical techniques. When the nature of data is time to event, different survival techniques are useful to identify important factors and to quantify their effects associated with survival of patients. Kaplan & Meier method (Kaplan & Meier, 1958), a nonparametric technique is one which is very commonly used in medical research. Non-parametric techniques have been extensively used in clinical research such as clinical trials, cancer research etc. For the past few decades, clinical trials of cancer therapy have relied almost exclusively on non-parametric statistical methods, such as log-rank test (Peto et al.; 1976, 1977). Nonparametric methods nevertheless suffer a serious limitation. These are sensitive to differences in survival between treatment groups, but do not give insight into the mechanisms by which therapy enhances survival (Gamel & Vogel, 1997). The other specific models in survival analysis are semi parametric (Cox proportional hazards model), parametric hazard models (Weibull, Exponential and Gompertz, etc) and parametric survival (Weibull, Exponential, Lognormal, Log-Logistic, etc) models. The most commonly used model among the existing methods is the Cox Proportional Hazards (CPH) model. It is preferred because estimation and inference about the parameters of interest are possible without assuming any form of the baseline hazard function. This standard model is considered as a robust method in survival analysis.

There are many applications of CPH model in the analysis of liver disease data including acute liver failure some of them are: Schlichting et al. (1983); Christensen et al. (1985); Christensen et al. (1986); Ginés et al. (1987); Acharya et al. (1996); Bird et al. (1998); EI-Serag & Everhart (2002); Ludvigsson, Elfstrom, Broome, Ekbom, and Montgomery (2007); Miyake et al. (2007); Boin et al. (2008), etc. ALF data has also been modeled by using Log-Logistic and Lognormal Accelerated Failure Time(AFT) model (Khanal, Sreenivas, & Acharya, 2014) for such models there is not necessary to satisfy the proportionality of hazards assumptions as it necessarily demands in the case of CPH model. Further, there may be the possibility to select different independent variables even from the same data set because of the differences in the model specifications. In this context this study is an attempt to apply CPH model to identify the prognostic factors of ALF patients based on the hospital data of All India Institute of Medical Sciences (AIIMS) by using CPH model and attempt has been made to check all the assumptions, goodness of fit of the model including its predictive probabilities etc.

MATERIALS AND METHODS

This study is based on retrospective cohort study design. The details of the statistical analysis, types of data, data management, evaluation of the assumptions of the model, goodness of fit of the model and the methods to predict survival probability for different scenario are presented in the subsequent units of the paper.

Data and its management

Altogether 1099 ALF patients' data was obtained from the liver clinic, AIIMS, New Delhi India, a government financed referral center in northern India for the period of 25 May 1986 to 31 December 2005. Each patients' details were cross verified with original case records for any coding and punching errors, which were rectified wherever necessary.

Response/outcome variable

The outcome variable is mortality of patients (coded 0 for survived and 1 for died) along with survival time (duration between date of hospital admission with ALF and date of death or discharge from the hospital after recovery).

Explanatory variables

Taking into account the theoretical considerations and the previous relevant studies, a set of demographic and clinical covariates were selected as explanatory variables. They are age of the patient (in years), gender (male/female), pregnancy status (male/non-pregnant female/ pregnant female), prothrombin time (in seconds), pre-encephalopathy period (in days), icterus to encephalopathy interval (in days), hepatic encephalopathy grade (I to IV), cerebral edema (absent/present), etiology (hepatitis E virus/ non-E), serum bilirubin (mg/dl), serum creatinine (mg%), urea (mg%), albumin (gm%),

Aspartate aminotransferase (AST) (measured in IU), and Alanine transaminase (ALT) (measured in IU). Before starting the analysis, all variables were checked for simple range checks for implausible values etc. After exploratory analysis, considering the numbers in each category, literature review of similar work on ALF, an expert hepatologists' suggestions and also the simplicity of the model developed for ease of interpretation, the explanatory variables were categorized taking clinically meaningful cut off values.

Statistical model

Since the ALF data considered in this study is of time to event, CPH model has been applied in order to identify the prognostic factors of ALF patients and to quantify the effects of these variables on outcome variable. The mathematical form of Cox regression model (Cox, 1972) for the i^{th} individual with a set of $x_1, x_2, \ldots\ldots, x_p$ explanatory variables is:

$$h_i(t) = h_0(t)\exp(\beta_1 x_{1i} + \beta_2 x_{2i} + \ldots\ldots + \beta_p x_{pi}) \quad \ldots\ldots\ldots\ldots\ldots \quad (1)$$

where $\beta_1, \beta_2, \ldots\ldots, \beta_p$ are unknown regression coefficients and $h_0(t)$, the baseline hazard function. Cox regression model assumes that the hazards of any two individuals are proportional over time i.e. the ratio between the hazards is the same at any time *t*. However, this does not preclude that the hazard may change over time. It only assumes that the changes in the hazard of any patient over time will always be proportional to changes in the hazard of any other subject. The β coefficients in the CPH model are estimated by maximum likelihood Estimation (MLE) method using the maximum likelihood function by using Newton-Raphson procedure (For detail explanation, please see Cox (1972, 1975)).

If an event time and a censored time are tied, it is assumed that the censoring occurs after the occurrence of the event. If there are only tied censored observations, the calculations of the likelihood function is not affected (Collett, 2003). In this ALF data being used for the present work, tied observations were found. Accordingly, all three methods of adjusting for the tied event times, namely Breslow and Crowley (1974); Efron (1977); Kalbfleisch and Prentice (2002) were attempted. As the three methods showed similar results, Breslow's approximation of likelihood function was applied in estimating the CPH model in this study.

Testing the significance of regression coefficients

The significance of the regression coefficient of each covariate in the CPH model can be assessed by three different tests namely partial likelihood ratio test, the Wald test, and the score test. Generally, the three tests lead to similar results and so is the conclusion about the significance of a regression coefficient (Hosmer & Lemeshow, 1999). In this study, the analysis has been based on Wald test. The

hazard ratio (*HR*), the exponential form of regression coefficient $\left[\exp(\hat{\beta})\right]$, is preferably reported because it is simple and easy to interpret as to the effect of the covariate on the outcome along with its confidence interval (C.I.) $\exp\left[\hat{\beta} \pm Z_{(1-\alpha/2)} SE(\hat{\beta})\right]$ where SE stands for standard error of the coefficient.

Model building procedure

Initially, CPH model has been applied for each explanatory variable, one at a time to identify the candidate variables for the multiple CPH model. Any variable whose univariate test indicates at least a moderate significance ($p \leq 0.10$) is considered as a candidate for multivariable CPH model. Stepwise forward selection procedure with removal probability 0.051 and entry probability 0.05 was applied to select variables for the final multivariable CPH model.

Testing the proportionality hazards assumption

The proportionality of hazards for each covariate in the final multivariable model was evaluated by three different methods, namely, plot of the log-cumulative hazard, Schoenfeld residuals based test, and adding an interaction term with time in the model. The proportionality of hazards has been tested graphically through the curves of the log-cumulative hazard function, $\log H_i(t) = \beta' x_i + \log H_0(t)$ with survival time for individuals with and without the exposure are parallel and equidistant vertically, throughout the follow-up time, if the proportional hazards model is to be satisfied. Another test for the proportionality of hazards was tested through Pearson's correlation coefficient (ρ) between scaled version of Scheoenfeld (Grambsch & Therneau, 1994) residuals for each covariate. Significant ($p < 0.05$) correlation at 5% level of significance between them indicates the violation of proportionality hazards assumption for that covariate (For detail formula and computation procedure for this, please see Schoenfeld(1982); Grambsch and Therneau(1994)). This test can be used as a global test for the final model which has been adopted in this study too. Another approach for testing the proportionality of hazards assumption adopted is by adding the interaction term between the exposure variable and time (or a function of time) separately for each predictor in the CPH model. Significant of interaction coefficient indicates the violation of PH assumption; otherwise the PH assumption can be taken as valid.

Test of goodness of fit of the CPH model

In order to test the goodness of fit of the final Cox regression model, different tests such as Cox-Snell (CS) residuals plot, visual assessment of the observed versus predicted survival pattern, Grønnesby and Borgan test modified by May & Hosmer and R^2 type statistic were applied in this paper. The Cox-Snell (CS) residual (Cox & Snell, 1968) for the i^{th} individual (i = 1, 2, ……., n) is given by:

$$rC_i = \exp(\hat{\beta}' x_i)\hat{H}_0(t_i) \quad \text{.....(2)}$$

where, $\hat{H}_0(t_i)$ is an estimate of the baseline cumulative hazard function at observed survival time (t_i) for an individual. In contrast to Schoenfeld residuals, Cox-Snell residuals are only one for each subject. If the model fits the data well, then the Cox-Snell residuals should have a standard exponential distribution with hazard function equal to one, and thus the plot of cumulative hazard of the Cox-Snell residuals against Cox-Snell residuals should be a straight line passing through the origin with a slope of unity. While assessing the goodness of fit of the final CPH model through visual assessment of observed vs predicted survival pattern, Patients are divided into different groups with non-overlapping levels of risk. The risk score termed as the prognostic index (*PI*) which is a linear combination of regression coefficients and their corresponding covariates values was computed for i^{th} individual for $x_1, x_2, \ldots x_p$ covariates can be calculates as:

$$PI = \hat{\beta}' x_i = \hat{\beta}_1 x_{1i} + \hat{\beta}_2 x_{2i} + \ldots\ldots\ldots + \hat{\beta}_p x_{pi} \quad \text{..... (3)}$$

The *PI* values are ranked and the total subjects in the study are divided into certain number of strata in such a way that there is approximately equal number of events in each stratum. The survival probabilities at each time point are estimated in each stratum and are plotted along with the observed survival curve in that stratum. The observed survival pattern for each stratum is obtained from the usual Kaplan-Meier method. The predicted model survival function for i^{th} individual will be:

$$\hat{S}(t, x_i, \hat{\beta}') = \left[\hat{S}_0(t)\right]^{\exp(\hat{\beta}' x_i)} \quad \text{..... (4)}$$

where, $\hat{S}_0(t)$ is the estimated baseline survival function

A close observed and predicted survival curves in each stratum indicate that the model fits well. Grønnesby and Borgan (1996) proposed a goodness of fit test for testing the overall fit of a Cox regression model which is similar to Hosmer & Lemeshow test (Hosmer & Lemeshow, 1989) as used in the logistic regression. Finally May and Hosmer (1998), following a method applied by Tsiatis (1980), suggested a useful method to test the goodness of fit of the CPH model modifying Grønnesby and Borgan (1996) test, and this was applied in this study. In this method, data is divided into ten parts on the basis of ranked values of *PI*, and the observed and expected number of events is computed in each group. It is considered that the observed counts within each risk decile are approximated by a Poisson distribution with mean equal to the estimated expected number of counts. For large value of mean, Poisson distribution is approximated to Normal distribution. Hence the observed and expected counts are compared in each decile by using Z score calculated as follows for i^{th} decile:

$$Z_i = \frac{O_i - E_i}{\sqrt{E_i}} \quad \text{..... (5)}$$

where O_i and E_i are the observed and expected number of events.

Z score is calculated in each decile and two sided p-value is used to check whether the difference is significant. Since multiple comparisons are involved, the level of significance is adjusted using Bonferroni method (Hosmer & Lemeshow, 1999). If the difference between observed and expected number of events in each decile is not significant, it indicates that the model fits the data well. Another method used in this study for assessing the goodness of fit of the final CPH model, Hosmer and Lemeshow (1999) recommended a R^2 Type statistic based on the log-likelihood of the model, which is defined as follows:

$$R_p^{\,2} = 1 - \left\{ \exp \frac{2}{n} \left[(L_0 - L_p) \right] \right\} \quad \ldots\ldots\ldots\ldots\ldots\ldots (6)$$

where L_p is the log-likelihood for the fitted CPH model with p covariates and L_0, the log-likelihood for null model i.e. the model with no covariates. Higher the value of R^2 naturally indicates the better fit of the model. However, lesser the value of R^2 may not mean that the model is a poor fit.

All statistical analysis has been carried out using STATA 9.0

RESULTS

A total of 1099 ALF cases were registered during the period 1986 – 2005. Males constituted 43.5%. Among females, 381 (61.3%) were non-pregnant and 240 (38.7%) were pregnant. The average age of patients was 28.5 ± 11.85 years. Of the 1099 patients enrolled 647 (58.9%) died within the hospital. Median stay of the patients in the hospital was 6 days and the median survival time of patients was 7 days, the maximum being 30 days. By 3 days 25% of patients died and by 23 days 75% of patients died. Patients with cerebral edema survived on the average for 5 days while those without cerebral edema survived for 15 days ($p < 0.001$). Similarly, the median survival time of patients with prothrombin time ≥ 25 seconds was significantly lower than that of patients with prothrombin time < 25 seconds (5 vs 10 days, $p < 0.001$). Poor survival experience was observed among patients aged ≥ 40 years as compared to those aged < 40 years (median 5 vs 8 days, $p < 0.001$). Thus, out of the 15 predictor variables tested, 9 variables, namely age, pregnancy status, total serum bilirubin, cerebral edema, grade of hepatic encephalopathy, prothrombin time, serum creatinine, etiology and pre-encephalopathy period showed significant association with mortality ($p < 0.05$). In addition, one variable, namely AST showed marginal association with mortality ($p = 0.08$) in the Keplan-Meier (K – M) survival analysis followed by Log- rank test which is shown in Table1. Please note that some information of some of the variables' could not be available because of which for some of the variables the total 1099 value does not match.

Table 1. Results of univariate analysis (N = 1099).

Characteristic		Num	Alive Num. (%)	Dead Num (%)	Mortality 25%	Mortality 50%	Mortality 75%	*p* value*	*HR* (95% C.I.)	*p* value $
Overall		1099	452 (41.1)	647 (58.9)	3	7	23	-	-	-
Age (years)	< 40	928	415 (44.7)	513 (55.3)	3	8	23		1.00	
	≥ 40	171	37 (21.6)	134 (78.4)	3	5	9	< 0.001	1.50 (1.24, 1.81)	< 0.001
Sex	Male	478	193 (40.4)	285 (59.6)	3	7	22		1.00	
	Female	621	259 (41.7)	362 (58.3)	3	8	23	0.26	0.93 (0.80, 1.09)	0.40
Pregnancy	Male	478	193 (40.4)	285 (59.6)	3	7	22		1.00	
	Not pregnant	381	147 (38.6)	234 (61.4)	3	6	21		1.02 (0.85, 1.21)	0.85
	Pregnant	240	112 (46.7)	128 (53.3)	4	9	26	0.04	0.81 (0.66, 1.00)	0.06
Total S. bilirubin (mg/dl)										
< 15		602	310 (51.5)	292 (48.5)	4	10	27		1.00	
≥ 15		497	142 (28.6)	355 (71.4)	3	5	12	< 0.001	1.65 (1.41, 1.92)	< 0.001
Cerebral Edema										
Absent		494	303 (61.3)	191 (38.7)	6	15	-		1.00	
Present		605	149 (24.6)	456 (75.4)	3	5	10	< 0.001	2.58 (2.18, 3.06)	< 0.001
Hepatic enc grade										
Gr. I & II		215	139 (64.6)	76 (35.4)	7	15	23		1.00	
Gr. III & IV		884	313 (35.4)	571 (64.6)	3	6	20	< 0.001	2.14 (1.69, 2.73)	< 0.001
Prothrombin time (seconds)										
< 25		578	297 (51.4)	281 (48.6)	4	10	-		1.00	
≥ 25		519	154 (29.7)	365 (70.3)	3	5	14	< 0.001	1.77 (1.51, 2.07)	< 0.001

*: Based on Log-rank test $: Based on Cox's regression Num: Number

Table 1. Results of univariate analysis (N =1099). (..Contd)

Characteristic		Num	Alive Num. (%)	Dead Num (%)	Mortality 25%	Mortality 50%	Mortality 75%	*p* value *	*HR* (95% C.I.)	*p* value $
Serum creatinine (mg%)										
≤ 1		767	340 (44.3)	427 (55.7)	4	9	23		1.00	
> 1		273	90 (33.0)	183 (67.0)	3	6	-	< 0.001	1.40 (1.17, 1.66)	< 0.001
Etiology	Hepatitis E virus	338	177 (52.4)	161 (47.6)	5	11	-		1.00	
	Non E	744	272 (36.6)	472 (63.4)	3	6	19	< 0.001	1.37 (1.13, 1.67)	0.001
Albumin (gm%)	> 3.5	194	93 (47.9)	101 (52.1)	4	9	17		1.00	
	≤ 3.5	798	339 (42.5)	459 (57.5)	4	8	23	0.88	1.06 (0.85, 1.32)	0.60
Urea (mg%)	≤ 50	965	395 (40.9)	570 (59.1)	3	7	22		1.00	
	> 50	134	57 (42.5)	77 (57.5)	3	7	-	0.70	1.04 (0.82, 1.31)	0.77
AST (IU)	≤ 300	473	224 (47.4)	249 (52.6)	4	9	-		1.00	
	> 300	517	209 (40.4)	308 (59.6)	4	7	20	0.08	1.36 (1.13, 1.63)	0.001
ALT (IU)	≤ 470	445	195 (43.8)	250 (56.2)	4	8	26		1.00	
	> 470	553	240 (43.4)	313 (56.6)	4	8	22	0.80	1.17 (0.97, 1.42)	0.09
Pre-enc period (days)										
No PEP		18	7 (38.9)	11 (61.1)	3	3	17		1.00	
1 - 7		530	242 (45.7)	288 (54.3)	3	9	26		0.86 (0.47, 1.58)	0.63
≥ 8		527	190 (36.1)	337 (63.9)	3	6	18	0.03	1.02 (0.56 1.86)	0.95
Icterous to Encephalopathy interval (days)										
No IEI		132	50 (37.9)	82 (62.1)	3	6	26		1.00	
1 – 4		474	208 (43.9)	266 (56.1)	3	8	23		0.98 (0.76, 1.27)	0.90
≥ 5		477	183 (38.4)	294 (61.6)	3	7	20	0.26	1.07 (0.83, 1.37)	0.61

*: Based on Log-rank test $: Based on Cox's regression Num: Number

The hazard ratio of mortality for each study variable is also shown in Table1. Univariate Cox regression (Table1) identified 8 significant predictors of mortality ($p < 0.05$) and two predictors of moderate significance ($p \leq 0.10$). Though there are similarities in the significant predictors identified by the K-M method and the Cox model, differences with respect to the significance by the two methods was noticed for two variables. The data for the present study spans over a long period starting from 1986 to 2005, covering 20 years. The quality of treatment and the profile of the ALF patients attending the liver clinic might be changing with time. Keeping these in mind, the total data span of 20 years was made into three parts, namely, 1986 -1992, 1993 - 1999 and 2000 – 2005. The differences with respect to the significance by K-M method and univariate Cox regression could be due to the fact that the study period variable as defined above was adjusted in the Cox regression, which could not be done in the K-M analysis. So only those variables showing at least a moderate association ($p \leq 0.10$) in the univariate Cox regression models were considered as candidate variables for building a multivariable Cox model.

Multivariable CPH model

The candidate variables for the multivariable Cox model were age, pregnancy status, total serum bilirubin, cerebral edema, hepatic encephalopathy grade, prothrombin time, serum creatinine, etiology, AST and ALT. These variables were considered in a forward stepwise manner with an entry probability 0.05 and removal probability 0.051.

Table 2. Multivariable Cox regression model for ALF data.

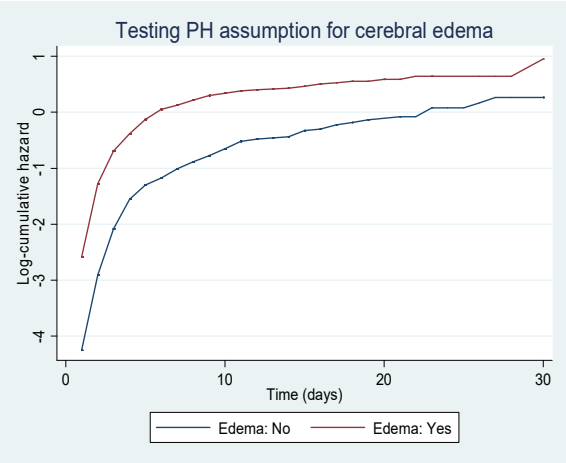
Variable	*HR* (95% C.I.)	SE	*p* value	Variable	*HR* (95% C.I.)	SE	*p* value
Cerebral Edema				Age (years)			
Absent	1.00			< 40	1.00		
Present	2.38(1.99, 2.85)	0.22	< 0.001	≥ 40	1.41(1.15, 1.72)	0.14	< 0.001
Total S.bilirubin (mg/dl)				S. creatinine (mg%)			
< 15	1.00			≤ 1	1.00		
≥ 15	1.49(1.27, 1.76)	0.12	< 0.001	>1	1.32(1.10, 1.57)	0.12	0.002
Prothrombin time (Seconds)				Etiology			
				Hep. E	1.00		
≤ 25	1.00			Non E	1.33(1.09, 1.63)	0.14	0.006
≥ 25	1.66(1.41, 1.96)	0.14	< 0.001				

Log-likelihood = -3714.503, N = 1026

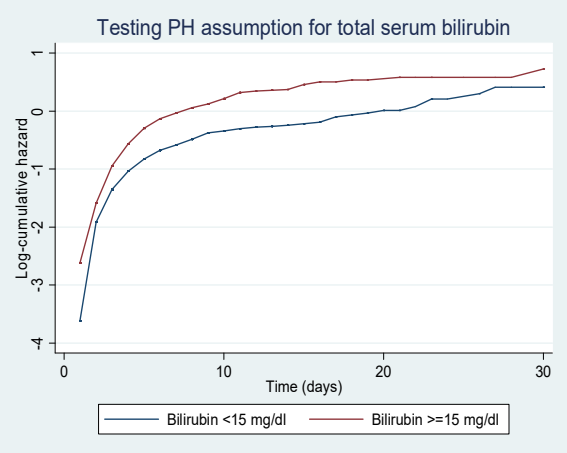
The stepwise procedure picked up six variables viz. age, prothrombin time, cerebral edema, serum bilirubin, serum creatinine and etiology of ALF which is considered as the final CPH model with these six variables (Table2). The hazards of mortality [HR: 2.38; 95% C.I.: (1.99, 2.85), $p < 0.001$] is observed to be 2.4 times higher in patients presenting cerebral edema as compared with the patients in absence of cerebral edema, which is the highest risk among all these prognostic factors.

Checking the assumption of proportional hazards: Log-cumulative hazards plot

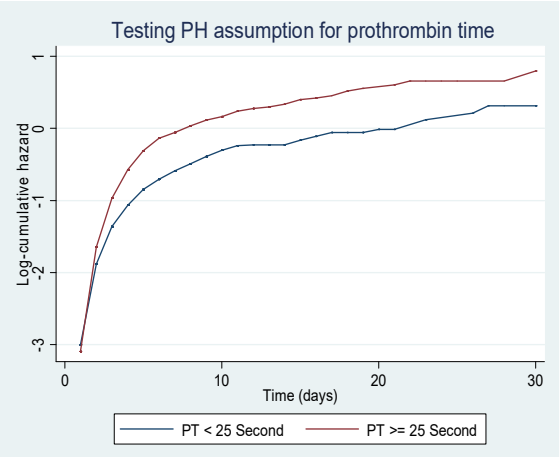
The Cox model assumes that the hazards are proportional. In order to test this proportionality of hazards (PH) assumption, the graph of logarithm of cumulative hazard versus survival time for each of the strata of each variable of the final multivariable Cox regression model was plotted, keeping the values of the other variables constant as shown in Figure 1(a) - 1(f).



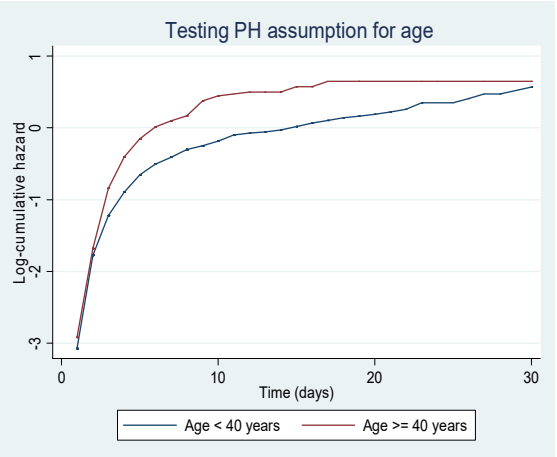
1(a)



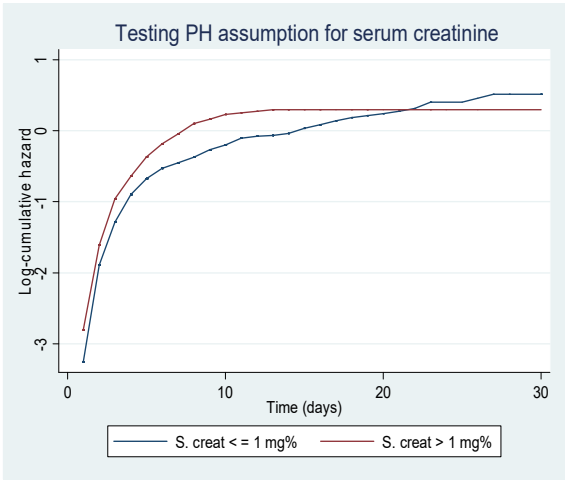
1(b)



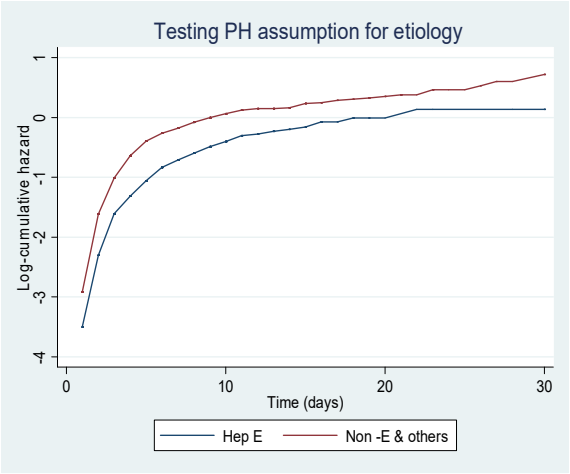
1(c)



1(d) (Contd…)



1(e)



1(f)

Fig. 1. Log-Cumulative hazard plots for (a) cerebral edema, (b) total serum bilirubin, (c) prothrombin time, (d) age, (e) serum creatinine and (f) etiology.

Figure 1 indicates that the proportionality of hazards is valid for all the variables, except for serum creatinine. The curves for the two strata of serum creatinine are intersecting each other, but almost at the end of follow up time. It might be because of few numbers of events in the end part of the study. On the whole, the assumption of constant vertical differences between the curves is not grossly violated and hence the proportionality hazard assumption can be reasonable.

PH test based on Schoenfeld residuals

The correlation (ρ) between Schoenfeld residuals and survival time for each covariate(Table3) indicate that all variables satisfied the proportional hazards assumption as the correlation between Schoenfeld residuals and survival time is not significant except for two variables cerebral edema and serum creatinine. However, the correlation coefficient itself is small (0.15 and 0.08 respectively) and the statistical significance might be due to the large number of cases. Further, the cumulative hazards plot for the variable cerebral edema showed almost perfect parallel curves.

Adding interaction term with time

The assumption of proportionality hazards was also checked for each covariate in the Cox regression model by adding an interaction term with time (Table4). The β coefficient for interaction effect of each covariate with time is found non - significant indicating the hazard is independent of time, except for the variable cerebral edema. However, as seen in Figure 4.2a, the log-cumulative

hazards curves are almost perfectly parallel for cerebral edema. So it can be taken as that there is no violation of proportionality of hazards assumption for all variables.

Table 3. Summary of test based on Schoenfeld residuals.

Variable	ρ	χ^2	d.f.	p value
Cerebral edema	-0.14890	13.21	1	0.0003
Total S. bilirubin	-0.00454	0.01	1	0.9099
Age	-0.02163	0.29	1	0.5896
Prothrombin time	0.05386	1.76	1	0.1845
Serum creatinine	-0.08432	4.29	1	0.0383
Etiology	-0.03116	0.58	1	0.4465
Global test		21.27	6	0.0016

Table 4. Results for Cox model after adding interaction term with time.

Variable	β	p value	Variable	β	p value
Cerbral Edema	1.45	< 0.001	Prothrombin time (PT)	0.40	0.002
(Cerb. Edema)*(Time)	-0.09	< 0.001	(PT)*(Time)	0.01	0.464
Total S. bilirubin	0.50	< 0.001	Serum creatinine	0.52	0.001
(Bilirubin)*(Time)	-0.0006	0.977	(Serum creatinine)*(Time)	-0.05	0.083
Age	0.50	0.001	Etiology	0.63	< 0.001
(Age)*(Time)	-0.01	0.609	(Etiology)*(Time)	-0.03	0.167

Goodness of the fitted CPH model: Cox-Snell residuals plot

The plot of the Cox-Snell residuals against the cumulative hazard of Cox-Snell residuals (Figure2) indicates that the overall fit of the Cox regression model is good. However, some deviation about the 45^0 line is seen in the right side which can be expected even if we have a well-fitting Cox model (Cleves, Gould, & Gutierrez, 2004) because of the reduced effective sample size caused by prior failures and censoring.

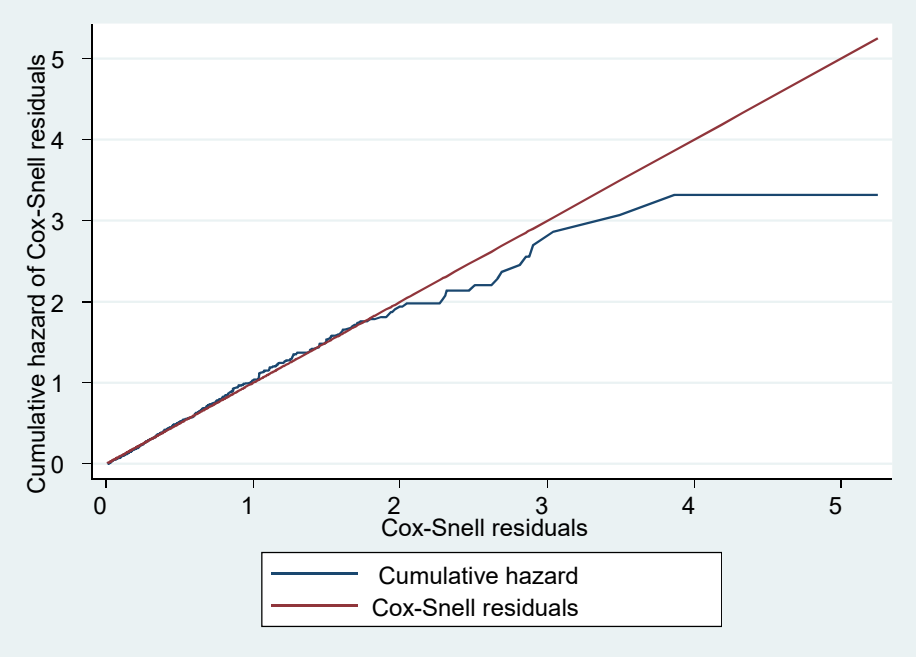


Fig. 2. Cox-Snell residuals plot for Cox regression model.

Goodness of the fitted CPH model: Visual assessment of observed vs predicted survival pattern

In this method, a prognostic index (*PI*), the weighted (considering respective regression coefficients of the covariates of the model as weights for each covariate value) sum of the values of variables (found in the final CPH model) was calculated for each subject under study. The values of *PI* ranged from 0.5742 to 11.0812. *PI* values were ranked and divided into four strata (Table5) in such a way that there was approximately equal number of events in each stratum. The observed and predicted survival probabilities for each stratum are compared graphically (Figure3). The observed and predicted curves are observed to be closer with each other in each stratum except at the end of follow up time indicating good fitting of the model.

Table 5. Distribution of *PI* in 4 risk strata.

Risk stratum	Prognostic Index (*PI*)	Total	Events observed
I	< 2.3665	430	143
II	2.3665 – 3.5129	234	147
III	3.5129 – 5.2784	185	145
IV	≥ 5.2784	177	165
Total		1026	600

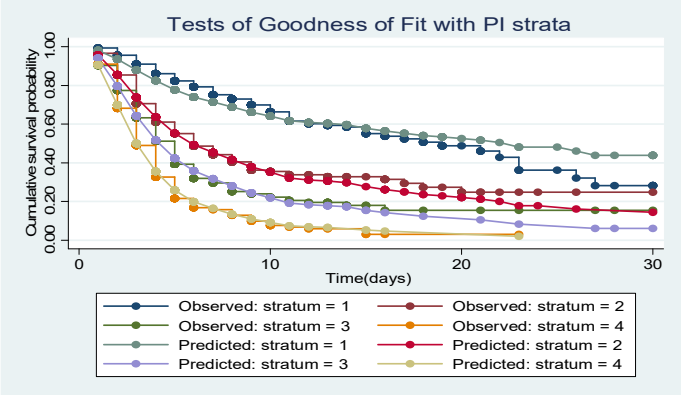


Fig. 3. Observed and model predicted survival curves in 4 risk strata.

Grønnesby and Borgan test modified by May & Hosmer

The observed and expected numbers of events were compared in each decile by using Z score (Table6).

Table 6. Observed number of events, estimated number of events, Z-score and two tailed *p*-values within each decile of risk based on CPH model.

Decile of risk	Risk group	Observed events	Expected events	Z	*p* value
1	< 1.4251	66	71.3603	-0.635	0.525
2	1.4251 - 2.1739	60	60.7420	-0.095	0.924
3	2.1739 - 2.4640	60	49.7255	1.457	0.145
4	2.4640 - 3.0593	54	55.3993	-0.188	0.851
5	3.0593 - 3.5129	51	56.5891	-0.743	0.457
6	3.5129 - 3.9887	61	66.8280	-0.713	0.476
7	3.9887 - 4.7327	70	54.0200	2.174	0.030
8	4.7327 - 5.8759	60	57.2102	0.369	0.712
9	5.8759 - 7.7866	54	48.9265	0.725	0.468
10	≥ 7.7866	64	79.1968	-1.708	0.090
	Total	600	600.00		

Only in the 7th decile the observed vs expected number of events seems to be significantly different ($p < 0.05$). If we use Bonferroni's correction for multiple testing, then none of the deciles has a significant difference between the observed and expected counts. Therefore, it can be concluded that the difference between observed and expected number of events within each of the 10 deciles of risk are due to chance and hence the fit of the Cox regression model is good.

R^2 type statistic

The statistic computed for six covariates (R_p^2) based on the partial log-likelihood of the final Cox PH model has come out to be 59.89% indicating that the 60% of variability in partial log-likelihood is explained by the CPH model with the final set of six covariates, indicating a good fit of the Cox regression model.

Prediction of survival probability based on the fitted model

The prognostic index (*PI*) for a subject defines his/her place within the prognostic spectrum defined by the model and it can be used further in the estimation of survival curve, the probability of surviving a given time etc for that subject. These predicted probabilities would be of help to the clinicians as to a quick assessment of the likely prognosis of a new patient at hospital admission. The predicted survival probabilities for different values of *PI* (combination of risk factor variables) are presented in Table7.

Table 7. Estimated survival probability (%) as a function of Prognostic Index (*PI*).

Survival time (days)	*PI* = 1	*PI* = 3	*PI* = 5	*PI* = 7	*PI* = 9	*PI* = 10
1	98.6	95.9	93.3	90.7	88.2	87.0
2	95.0	85.7	77.4	69.8	63.0	59.8
3	90.5	74.0	60.6	49.6	40.6	36.7
4	86.0	63.7	47.2	34.9	25.8	22.2
5	82.1	55.4	37.4	25.2	17.0	14.0
6	79.1	49.4	30.9	19.3	12.1	9.5
7	76.9	45.5	26.9	15.9	9.4	7.3
8	74.7	41.7	23.3	13.0	7.3	5.4
9	72.5	38.2	20.1	10.6	5.6	4.0
10	70.8	35.5	17.8	8.9	4.5	3.2
11	68.7	32.5	15.3	7.2	3.4	2.4
12	68.0	31.4	14.5	6.7	3.1	2.1
13	67.6	30.9	14.1	6.4	2.9	2.0
14	67.0	30.1	13.5	6.1	2.7	1.8
15	65.4	28.0	12.0	5.1	2.2	1.4
16	64.3	26.6	11.0	4.5	1.9	1.2

The presented survival probabilities are considered up to 16 days as the median survival probability of ALF patients is 7 days. It is seen from the table that higher the value of *PI*, lower the survival

probability for a given time. Nearly 9%, 26%, 39%, 50%, 59% and 63% of ALF patients with a *PI* of 1, 3, 5, 7, 9 and 10 respectively die by 3 days of hospital stay.

Prediction of survival probability on the number of risk factors

The predicted survival probability according to the prognostic factors present (out of six predictors identified by the Cox model) has been presented in Table 8.

Table 8. Estimated survival probability (%) on the basis of number of risk factors present.

Day	Number of risk factors present				
	None	Any 1	Any 2	Any 3	≥ 4
1	99.2	98.8	97.1	94.8	92.0
2	-	95.7	89.9	82.3	73.8
3	94.5	91.8	81.4	68.6	55.4
4	91.9	88.0	73.5	56.9	41.3
5	89.6	84.6	66.9	47.9	31.5
6	87.7	81.9	61.9	41.6	25.3
7	86.4	80.1	58.6	37.6	21.6
8	85.0	78.2	55.3	33.7	18.2
9	83.6	76.2	52.0	30.2	15.3
10	82.5	74.6	49.5	27.6	13.3
11	81.2	72.9	46.7	24.8	11.2
12	80.7	72.2	45.7	23.8	10.6
13	80.4	71.9	45.1	23.3	10.2
14	-	71.4	44.4	22.6	9.8
15	-	70.0	42.4	20.8	8.5
16	78.4	69.0	41.0	19.5	7.8

It can be clearly seen from Table 8 that higher the number of risk factors presents lowers the survival probability for a given time. For example; the probabilities of a person surviving 3 days for presence of any one, any two, any three and four or more risk factors were 91.8%, 81.4%, 68.6% and 55.4% respectively.

DISCUSSION

Acute Liver Failure (ALF), one of the catastrophic liver diseases, is a highly complex syndrome arising when acute liver cell damage causes breakdown of vital functions of the normal liver within a few days or weeks. The causes of ALF are not common in each region of the world. There is high prevalence of acetaminophen associated ALF in the United Kingdom and in the United States. The most common etiologies in Taiwan (Ho et al., 2014) were viral (45.4%, mainly Hepatitis B virus) and followed by alcohol/toxin (33%). Prognostic factors identified in Taiwan was alcohol consumption (*HR*: 1.67), malignancy (*HR*: 2.90), frequency of checkups per week for total bilirubin (*HR*: 1.57), sepsis (*HR*: 1.85), the use of hemodialysis / hemofiltration (*HR*: 2.12) and proton pump inhibitor (*HR*: 0.94). These factors are different from this study. Because of variations in the etiology of ALF in different geographic regions, the prognostic factors may also be different. Review of literature indicates whenever studies applied CPH model in this set up but has not been found attempted to assess the proportionality assumptions and goodness of fit of the model. Nonetheless Khanal, et al. (2014) applied Accelerated Failure Time (AFT) models to identify the prognostic factors of ALF patients and identified the same set of independent variables as this study has identified. In view of these, the present study is undertaken to identify the prognostic factors in the ALF patients through the use of CPH model in the Indian scenario examining all the tests for proportionality of hazards, goodness of fit tests and predictions of survival probability based on the fitted model.

In the beginning, K - M method followed by log-rank test and univariate Cox regression were applied to select candidate variables for multivariable Cox regression. Log-rank test showed promise of association ($p \leq 0.1$) with 10 variables. Univariable Cox regression also picked up ten variables showing promise, but only nine only were common to both Log-rank test and univariable Cox regression. One variable (PEP) showing significant association ($p = 0.03$) by the log-rank test turned up as insignificant ($p > 0.10$) in the Cox regression. Similarly, the variable ALT which was insignificant ($p = 0.8$) by the log-rank test turned up as moderately significant ($p = 0.09$) in the Cox regression. Ideally, both the methods should identify the same as significant/moderately significant variables. The present study covered data for 20 years and this time period was made into 4 groups and adjusted in the Cox regression but not in K-M analysis. This might be a reason for the differences in identifying variables with some promise of association.

Taking the variables with at least moderate significant association with mortality in the univariate Cox regression, a step-wise Cox PH model was built to identify the independent predictors. Six variables namely age, cerebral edema, total serum bilirubin, prothrombine time, serum creatinine and etiology of ALF emerged as significant predictors of mortality. The assumption of proportional

hazards for the Cox model was assessed by three methods for each of these six predictor variables. All six variables satisfied this requirement by at least one method of the three. When a log-cumulative hazards plot showed a deviation from the PH assumption for a variable, the deviation was either minimal or was towards the end of the follow-up. Similarly, when the Schoenfeld residuals showed significant correlation with survival time, the magnitude of correlation itself was observed to be small, but the statistical significance might be due to the large sample of the study. Further, a significant non-proportionality may make no difference to the interpretation of a data set, particularly for large sample size (Therneau & Grambsch, 2000). Thus, it can be considered that the proportionality of hazards assumption is met by each variable in the final Cox regression model developed.

CONCLUSION

The developed CPH model with these six prognostic factors has satisfied the proportionality of hazards assumptions, test of goodness of fit of the model. A risk score/prognostic index has been developed based on the prognostic factors identified and using this, the survival experience of ALF patients with specific risk score is predicted. It is clearly indicated that higher the value of *PI*, lower the survival probability for a given time. In addition, the survival probability is predicted on the basis of number of risk factors present for each day of follow up. Higher the number of risk factors present lowers the survival probability for a given time. Hence, these predicted survival probabilities can be of help to hepatologists to make a quick decision regarding the likely prognosis of a patient at admission and also be helpful in triaging the ALF patients for liver transplant.

CONFLICT OF INTEREST

The authors declared that there is no conflict of interest.

ACKNOWLEDGMENTS

Authors would like to special thanks to Prof. Dr. S. N. Dwivedi and Dr. Padam Singh for their valuable suggestions and comments during this entire study. University Grants Commission (UGC), Nepal is also acknowledged for the support of Ph.D. fellowship as this study is based on the Ph.D. thesis work. Anonymous reviewers are also highly acknowledged for their critical comments and suggestions to improve the manuscript.

REFERENCES

Acharya, S. K., Dasarathy S., Kumar, T. L., Sreenivas, V., Sushma, S, Prasanna, K. S. U., . …., & Tandon, B. N. (1996). Fulminant hepatitis in a tropical population: clinical course, cause, and early predictors of outcome. *Hepatology*, *23*, 1448-1455.

Bird, G. L, Prach, A. T., McMahon, A. D., Forrest, J. A., Mills, P. R., & Danesh, B. J. (1998). Randomised controlled double-blind trial of the calcium channel antagonist amlodipine in the treatment of acute alcoholic hepatitis. *Journal of Hepatology*, *28*, 194-198.

Boin Ide, F., Leonardi, M. I., Udo, E. Y., Sevá-Pereira, T., Stucchi, R. S., & Leonardi, L. S. (2008). The application of MELD score in patients submitted to liver transplantation: a retrospective analysis of survival and the predictive factors in the short and long term. *Arquivos de Gastroenterologia*, *45*, 275-283.

Breslow, N., & Crowley, J. (1974). A large sample study of the life table and product limit estimates under random censorship. *The Annals of Statistics*, *2*, 437-453.

Christensen, E., Neuberger. J., Crowe, J., Altman, D. G., Popper, H., Portmann, B., …., & Williams, R. (1985). Beneficial effect of azathioprine and prediction of prognosis in primary biliary cirrhosis. Final results of an international trial. *Gastroenterology*, *89*, 1084-1091.

Christensen, E., Schlichting, P., Andersen, P. K., Fauerholdt, L., Schou, G., Pedersen, B. V., …., & Tygstrup, N. (1986). Updating prognosis and therapeutic effect evaluation in cirrhosis using Cox's multiple regression model for time dependent variables. *Scandinavian Journal of Gastroenterology*, 2*1*, 163-174.

Cleves, M. A., Gould, W. W., & Gutierrez, R. G. (2004). An *introduction to survival analysis using Stata*. A Stata Press Publication, Texas: STATA Corporation.

Collett, D. (2003). *Modeling survival data in medical research*. Boca Raton Florida: Chapman & Hall/CRC.

Cox, D. R., & Snell E. J. (1968). A general definition of residuals (with discussion). *Journal of the Royal Statistical Society, Series B, 30*, 248-275.

Cox, D. R. (1975). Partial likelihood. *Biometrika*, *62*, 269 - 76.

Cox, D.R (1972). Regression models and life tables. *Journal of the Royal Statistical Society, Series B, (Methodological)*, *34*(2), 187-220.

Efron, B. (1977). The efficiency of Cox's likelihood function for censored data. *Journal of the American Statistical Association*, *72*, 557-565.

El-Serag, H. B., & Everhart, J. E. (2002). Diabetes increases the risk of acute hepatic failure. *Gastroenterology*, *122*, 1822-1828.

Gamel, J. W. & Vogel R. L. (1997). Comparison of parametric and nonparametric survival methods using simulated clinical data. *Statistics in Medicine*, *16*, 1629-1643.

Ginés, P., Quintero, E., Arroyo, V., Terés, J., Bruguera. M., Rimola, A., …., & Rozman, C. (1987). Compensated cirrhosis: natural history and prognostic factors. *Hepatology*, *7*, 122-128.

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

Grønnesby, J. K., & Borgan, O. (1996). A method for checking regression models in survival analysis based on the risk score. *Lifetime Data Analysis*, *2*, 315-328.

Ho, C. M., Lee, C. H., Wang, J. Y., Lee, P. H., Lai, H. S., & Hu, R. H. (2014). Nationwide longitudinal analysis of acute liver failure in Taiwan. *Medicine*, *93*(4), e35. http://doi.org/10.1097/MD.0000000000000035

Hosmer, D. W., & Lemeshow, S. (1989). *Applied logistic regression*. New York: John Wiley & Sons, Inc.

Hosmer, D. W., & Lemeshow S. (1999). *Applied survival analysis: Regression modeling of time to event data.* New York: A Wiley-Inter Science Publication, John Wiley & Sons.

Kalbfleish, J. D., & Prentice, R. L. (2002). *The statistical analysis of failure time data*. Second edition, New York: Wiley.

Kaplan, E. L., & Meier, P. (1958). Nonparametric estimation from incomplete observations. *J. Amer. Statist. Assoc. 53*(282), 457-481.

Khanal, S. K, Sreenivas, V., & Acharya, S. K. (2014). Accelerated failure time models: An application in the survival of acute liver failure patients in India, *International Journal of Science of Research (IJSR)*, *3*(6), 161-166.

Lee, W. M. & Sorrell, M. F. (1996). Developing a World view toward acute liver failure. *Hepatology*, *24*, 270-271.

Ludvigsson, J. F., Elfström, P., Broomé, U., Ekbom, A., & Montgomery, S. M. (2007). Celiac disease and risk of liver disease: a general population-based study. *Clinical Gastroenterology and Hepatology*, *5*, 63-69.

May, S., & Hosmer, D. W. (1998). A simplified method for calculating a goodness-of-fit test for the proportional hazards model. *Lifetime Data Analysis*, *4*, 109-120.

Miyake, Y., Iwasaki, Y., Terada, R., Takaguchi, K., Sakaguchi, K., & Shiratori, Y. (2007). Systemic inflammatory response syndrome strongly affects the prognosis of patients with fulminant hepatitis B. *Journal of Gastroenterology*, *42*, 485-492.

Peto, R., Pike, M. C., Armitage, P., Breslow, N. E., Cox, D. R., Howard, S. V.,.,…, & Smith, P. G. (1976). Design and analysis of randomized clinical trials requiring prolonged observation of each patient. I. Introduction and design. *British Journal of Cancer*, *34*, 585-612.

Peto, R., Pike, M. C., Armitage, P., Breslow, N. E., Cox, D. R., Howard, S. V., …., & Smith, P. G. (1977). Design and analysis of randomized clinical trials requiring prolonged observation of

each patient. II. Analysis and examples. *British Journal of Cancer*, *35*, 1-39.

Schlichting, P., Christensen, E., Andersen, P. K., Fauerholdt, L., Juhl, E., Poulsen, H., & Tygstrup, N. (1983). Prognostic factors in cirrhosis identified by Cox's regression model. *Hepatology*, *3*, 889-895.

Schoenfeld, D. (1982). Partial residuals for the proportional hazards regression model. *Biometrika*, *69*, 239-241.

Therneau, T. M., & Grambsch, P. M. (2000). Modeling survival data: Extending the Cox model. New York: Springer.

Trey, C., & Davidson, C. S. (1970). The management of fulminant hepatic failure. *Progress in Liver Diseases*, 3, 282-298.

Tsiatis, A. A. (1980). A note on goodness-of –fit test for the logistic regression model. *Biometrika*, *67*, 250-251.

Reference to this paper should be made as follows:

Khanal, S. P., Sreenivas, V., & Acharya, S. K. (2018). Cox proportional hazards model for identification of the prognostic factors in the survival of acute liver failure patients in India**.** *Nep. J. Stat., 2*, 53-74.