EUROPEAN ACADEMIC RESEARCH
Vol. V, Issue 10/ January 2018

Impact Factor: 3.4546 (UIF)
DRJI Value: 5.9 (B+)

ISSN 2286-4822
www.euacademic.org

Proportional Hazard Models vs. Accelerated Models Comparative Study in Large Censoring Data

KHALID SHOAIB[1]
United Nations Development Program, Program Unit
Khartoum, Sudan
Sudan University of Science and Technology
College of Science, Statistics Department, Khartoum, Sudan
Dr. AHMED HAMDI
Sudan University of Science and Technology
College of Science, Statistics Department, Khartoum, Sudan
Dr. AL TAIYB AHMED

Abstract:

The objective of this paper was to support an argument for the consideration of the accelerated failure time (AFT) models as an alternative to the proportional hazards (PH) models in the analysis of survival sample of Children<5 years real dataset. Critiqued PH model and assessed the lack of fit to overcome the violation of PHs, Cox model used with time-dependent covariates.

The methodological developments of survival analysis that had the most profound impact are the Cox PH model for examining the covariate effects on the hazard function and the proposed AFT model but seldom used. The basic concepts presented was semiparametric methods (Cox PH and Cox models with time-dependent covariates) and parametric methods (Parametric PH and the AFT models) for analyzing survival data.

The objective of the analysis was to determine whether the samples of five diseases preventive therapies affected the rate of Children<5 years mortality and survivor. The conclusion was considered the AFT model as an alternative to the PH model in the analysis and evaluation of the effects of large censoring survival data.

[1] Corresponding author: khalid_stat@hotmail.com

After a comparison of all models and the assessment of goodness-of-fit found that the log-logistic AFT model fits better, more valuable and realistic alternative to PH model in some situations. Also provided the predicted hazard functions and survival functions, median survival times and time ratios. AFT model could easily interpret the effected results upon the expected median duration of effected Children in a clinical setting. Thus, PH model suggested may not appropriate in some situations.

Key words: Censor Data; Survival Time; Parametric and Semiparametric Model; Cox PH Models; AFT Models; Children<5 years Mortality.

1. INTRODUCTION

This study is based on the large number of Sudanese Children<5 years participants, where the prevalence of 5 samples of diseases infection were very high. The study shown that the benefit of the preventive therapies to delay the death or infected Children's was not confirmed, however, there was effective in reducing the Children incidence. The Cox PH model, Cox model with time dependent variables, piecewise exponential model and the AFT model to this dataset have been applied as well the corresponding of the results and compared the main methods of Cox and AFT. However, the interaction between these diseases increased fatality [1], caused significant Children morbidity and mortality.

It has been observed that during the period of study, a huge number of Sudanese Children<5 years affected by the most 10 types of diseases. This paper applied PH models vs. AFT models in Children survivor for the most common 5[2] of 10 diseases in period 2012-2016. Data collected from the one of the biggest Children hospital in Sudan *"Jaffar Ibn Oaf"* at Khartoum, due to up normal increasing of the Children's number with chronic diseases AFT method has been

[2] Renal Failure in Acute "http://dx.doi.org/10.1016/S0749-0704(05)70329-8"; Congenital DeformityHeart"http://www.who.int/mediacentre/factsheets/fs370/en/";Leukemia"https://www.webmd.com/cancer/lymphoma/childhood-leukemia-symptoms-treatments"; Septicemia "http://kidshealth.org/en/parents/sepsis.html/ sickle-cell-anemia.html"; Sickle cell disease "http://kidshealth.org/en/parents/sickle-cell-anemia.html".

preferred in survival analysis to reduce the Children healthy challenges and difficulties have been faced.

The complexities provided by the presence of censored observations led to the development of a new field of statistical methodology. The methodological developments in survival analysis were largely achieved in the latter half of the 20th century. Although Bayesian methods in survival analysis [26] are well developed and becoming quite common for survival data, but this application was focused on frequents methods.

The study touches only the partial likelihood ratio inference for Cox's type of models and AFT models. It was demonstrated that the parametric and semiparametric models provided various flexibility in modeling survival data. For analysis of asymptotic properties of the non-or semi-parametric components in Cox's type of models, counting processes and their associated martingales play an important role. For details, interested readers can consult with Fan, Gijbels, and King (2007) and Cai, Fan, Jiang, and Zhou (2007). However, there were many other approaches to model survival data. Parametric methods for censored data are covered in detail by Kalbfleisch and Prentice (1980, Chapters 2 and 3) and by Lawless (1982, Chapter 6). Semiparametric models with unspecified baseline hazard function are studied in Cox and Oakes (1984). Martingale methods were also used to study the parametric models (Borgan 1984) and the semiparametric models (Fleming and Harrington 2005; Andersen et al, 1993). While parametric methods work well for homogeneous samples, they don't determine whether certain variables are related to the survival times. The standard multiple linear regression model was not well suited to this survival data for several reasons. Firstly, survival times are rarely normally distributed. Secondly, censored data result in missing values for the dependent variable (survival time) [35]. Although, Cox PH model became the most widely used for the analysis of survival data in the presence of covariates or prognostic factors because of its simplicity, and not being based on any assumptions about the survival distribution and the model assumes that the underlying hazard rate was a function of the independent covariates, but no assumptions are made about the nature or shape of the hazard function. In the last several years, the theoretical basis for the model has been solicited by connecting to the study of counting processes and martingale theory, which was

discussed in the books of *Fleming and Harrington* [18] and of *Andersen et al* [2]. These developments have led to the introduction of several new extensions to the original model. However, the Cox PH model may not be appropriate in many situations and other medications such as striated Cox model or Cox model with time-dependent variables can be used for the analysis of survival data. The AFT [10] model become another alternative method for the analysis of survival data.

1.1 CENSORING DATA TYPES

The reasons behind why right censoring occurred, due to no event before the study ends, loss to follow-up during study period or withdrawal from the study because of some reasons. The last reason may be caused by competing risks. The right censored survival time was then less than the actual survival time.

Censoring could also have occurred if the presence of a condition has been observed but don't know where it began. In this case called left censoring and the actual survival time was less than the observed censoring time. While, if Children<5 years was known to have experienced an event within an interval of time but the actual survival time was unknown, this called interval censoring. The actual occurrence time of event was known within an interval of time.

Right censoring was very common in survival time data, but the left censoring was rare. The term "censoring" will be used in this paper to mean in all instances "right censoring". An important assumption for methods presented for the analysis of censored survival data was the individuals who were censored at the same risk of subsequent failure as those who were still alive and uncensored i.e., a subject whose survival time is censored at time C must be representative of all other individuals who have survived to that time. If this was the case, the censoring process called non-informative. Statistically, if the censoring process was independent of the survival time, *i.e.*

$$P(X \geq x, C \geq x) = P(X \geq x)\, P(C \geq x).$$

1.2 SURVIVAL TIME DISTRIBUTION

Let T be a random variable denoting the survival time. The distribution of $S(t)$ is characterized by any of three functions: *the*

survival function, *the probability density function or the hazard function*. The following definitions are based on textbook [32]. Note the survival function was defined for both discrete and continuous$T$. The probability density and hazard functions are easily specified for discrete and continuous$T$. The definition of the survival function is defined as the probability that the survival time was greater or equal to t.

$$S(t) = P(T \geq t), t \geq 0,$$

for a discrete random variable T taking well-ordered values $0 \leq t_1 \leq t_2 \leq \ldots \leq \tau_j$ let the probability mass function be given by $P(T = t_i) = f(t_i),\ i = 1,2, \ldots$ then the survival function is

$$S(t) = \sum_{j|t_{j\geq t}} f(t_j) = \sum f(t_j) I_{t_{j\geq t}},$$

Where the indicator function $I_{t_{j\geq t}} := \begin{cases} 0\ if\ t_j < t \\ 1\ if\ t_j \geq t \end{cases}$

In this case, the hazard function $h(t)$ is defined as the conditional probability of failure at time t_j given that the individual has survived up to timet_j,

$$h_j = h_{(j)} = P\big(T = t_j \big| T \geq t_j\big) = \frac{f(t_j)}{S(t_j)} = \frac{S\big(t_j\big) - S(t_{j+1})}{S(t_j)} = 1 - \frac{S\big(t_{j+1}\big)}{S\big(t_j\big)}.$$

Thus, $1 - h\big(t_j\big) = \frac{S(t_{j+1})}{S(t_j)}$, and

$$\prod_{j|t_{j<t}} (1 - h\big(t_j\big)) = \frac{S(t_2)}{S(t_1)} \times \frac{S(t_3)}{S(t_2)} \times \ldots \times \frac{S(t_{j+1})}{S(t_j)} = S(t), \qquad \textbf{(1.1)}$$

Because $S(t_1) = 1$ and $S(t) = S\big(t_j + 1\big)$. Moreover, $f(t_j) = h(t_j)S(t_j)$

$$= h\big(t_j\big) \prod_{i=1}^{j-1} (1 - h(t_i)). \qquad \textbf{(1.2)}$$

For a continuous variable $\boldsymbol{T}$, the probability density function of $\boldsymbol{T}$ is $f(t) = F'(t) = -S'(t),\ t \geq 0$. The hazard function defined as gives the instantaneous failure rate at t given that the individual has survived up to time$t$, i.e. $h(t) = \lim_{\Delta t \downarrow 0} P \frac{(T \leq t + \Delta t\ |T \geq t)}{\Delta t}, t \geq 0$. There was a clearly defined relationship between $S(t)$ and $h(t)$ given by the formula $h(t) = \frac{f(t)}{S(t)} = -\frac{s'(t)}{S(t)} = \frac{-\log S(t)}{dt}$ **(1.3)**

$H(t) = -log\ S(t)$ Equivalently to $S(t) = exp(-H(t))$

$S(t) = exp\left[-\int_0^t h(u)du\right] = exp\ (-H(t)), t \geq 0.$ **(1.4)**

Where $H(t) = -\int_0^t h(u)du$ is called the cumulative hazard function, which can be obtained from the survival function since $H(t) = -\log S(t)$ the probability density function of T can be written

$$f(t) = h(t)\, exp\left[-\int_0^t h(u)du\right], t \geq 0$$

These three functions gave mathematically equivalent specification of the distributions of the survival time T. If one of them was known, the others two are determined. The survival function is most useful for comparing the survival progress of two or more groups. The hazard function gaves a more useful description of the risk of failure at any time point.

2. PARAMETRIC AND SEMIPARAMETRIC

In survival analysis, it was always a good idea to present numerical or graphical summaries of the survival times for the individuals. In general, survival data are conveniently summarized through estimates of the survival function and hazard function. The estimation of the survival distribution provides' estimates of descriptive statistics such as mean of the survival time. These methods are parametric or semiparametric since assumptions of the distribution of survival time are required.

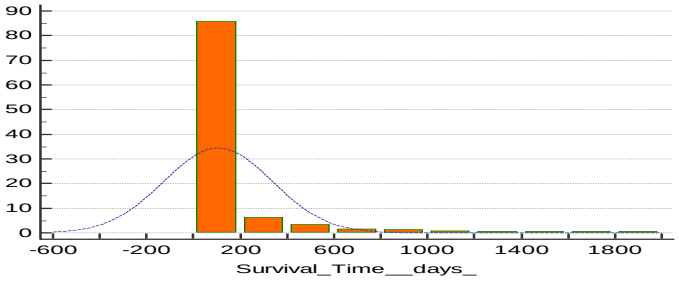


Figure 2.1: Relation between Survival Time and Events Frequencies

2.1 GREENWOODS FORMULA

Confidence intervals for the survival probability can also be calculated by the well-known Greenwoods formula [23]. First, the variances of the $\widehat{h_j}s$ is needed. Let the number of individual at risk at $t_{(j)}$ be r_j and the number of deaths at $t_{(j)}$ be d_j. Given r_j, the number of individuals

surviving through the interval $[t_{(j)}; t_{(j+1)})$, $r_j - d_j$, can be assumed to have binomial distribution with parameters r_j and 1-h_j.

The conditional variance of $r_j - d_j$ is given by $V(r_j - d_j|r_j) = r_j h_j (1 - h_j)$.

The variance of $\widehat{h_j}$ is $V(\hat{h}_j|r_j) = V(1 - \hat{h}_j) = V\left(1 - \frac{d_j}{r_j}\right) = \frac{h_j(1-h_j)}{r_j}$.

Since $\hat{h}_j$ is conditional independent of $\hat{h}_1, \dots, \hat{h}_j$ given $r_1, \dots, r_{j-1}$; the delta method [11] can be used to obtain. $V\left(ln\widehat{S}(t)\middle|r_j: t_{(j)} < t\right) = V\left[\sum_{j:t_{(j)}<t}(ln\left(1 - \hat{h}_j\right)|r_j\right] = \sum_{j:t_{(j)}<t} V\left[ln(1 - \hat{h}_j)|r_j\right]$

$\approx \sum_{j:t_{(j)}<t}(\frac{d}{dx} ln(1-x))^2{}_{x=\hat{h}_j} V(\hat{h}_j|r_j) = \sum_{j:t_{(j)}<t} \left\{-\frac{d}{1-\hat{h}_j}\right\}^2 \frac{h_j(1-h_j)}{r_j}, j = 1, \dots, r.$

Could estimate this by simply replacing h_j withb $\hat{h}_j = d_j/r_j$, which gives $\hat{V}\left(\ln \hat{S}(t)\right) = \sum_{j:t_{(j)}<t} \frac{d_j}{r_j(r_j - d_j)}$, $r_j = 1, \dots, r.$

Let Y= $\ln \hat{S}(t)$ again using the delta method, then we got $\hat{V}\left(\hat{S}(t)\right) \approx \left[\hat{S}(t)\right]^2 \sum_{j:t_{(j)}<t} \frac{d_j}{r_j(r_j - d_j)}$. This known as Greenwoods formula. The K-M estimator and functions of it have been proved to be asymptotically normal distributed [2], [18]. Thus, the confidence intervals can be constructed by the normal approximation based on $S(t)$.

2.1.1 ESTIMATING THE MEDIAN AND PERCENTILE OF SURVIVAL TIME

Since the distribution of survival time tends to be positively skewed, the median is preferred for a summary measure. The median survival time is the time beyond which 50% of the individuals under study are expected to survive the 95% confidence interval for the pth percentile $\hat{t}(p)$ has limits of t ̂(p)±1.96 SE {t ̂(p)}.

Table 2.1: Variable Summary Report, Break per Gender=1 "male"

Variables	Count	Mean	Median	SE	Min	Max	Interquartile Range	25th Percentile	75th Percentile
Age	628	493.50	90	26.162	1	1820	719	11	730
Stage	628	2	2	0	2	2	0	2	2
Symptoms	628	1.201	1	0.0160	1	2	0	1	1
Disease Type	628	3.753	4	0.0471	1	5	2	3	5
Disease History	628	2	2	0	2	2	0	2	2
Height (cm)	452	56.513	48.5	1.355	0	155	23	39	62
Weight (kg)	617	7.150	3.1	0.424	0	161	7.7	2.3	10
Freq. Visits	627	1.258	1	0.023	1	4	0	1	1
Status	628	0.207	0	0.016	0	1	0	0	0
Time	628	100.631	17	8.849	0	1719	66.75	7	73.75

Table 2.2: Variable Summary Report, Break per Gender=2 "Female"

Variables	Count	Mean	Median	SE	Min	Max	Interquartile Range	25th Percentile	75th Percentile
Age	472	604.896	210	32.762	1	1820	1082	13	1095
Stage	471	2	2	0	2	2	0	2	2
Symptoms	472	1.210	1	0.019	1	2	0	1	1
Disease Type	472	3.699	4	0.059	1	5	2	3	5
Disease History	472	2	2	0	2	2	0	2	2
Height (cm)	305	58.285	47	1.865	0	165	43	37	80
Weight (kg)	461	8.641	3.6	0.499	0	110	10.6	2.4	13
Freq. Visits	468	1.259	1	0.027	1	4	0	1	1
Status	472	0.227	0	0.019	0	1	0	0	0
Time	472	111.248	14.5	11.265	0	1816	67.5	6	73.5

2.2. COMPARISON OF PARAMETRIC, SEMIPARAMETRIC AND NONPARAMETRIC OF SURVIVAL DISTRIBUTIONS

Although one commonly used in non-parametric tests for comparison of two or more survival distributions was the log-rank test [40] for the survival curves could give an insight about the difference of survival functions in two or more groups, but whether these observed differences were statistically significant requires a formal statistical test. So, there were number of methods that used to test equality of the survival functions in different groups.

Let's take two groups as an example. Let $t_{(1)} < t_{(2)} < \ldots < t_{(k)}$ be the ordered death times across two groups. Suppose that d_j failures occur at $t_{(j)}$ and that r_j subjects are at risk just prior to $t_{(j)}$ $(j = 1,2,\ldots,k)$. Let d_{ij} and r_{ij} be the corresponding numbers in group $i\,(i = 1,2)$.

The log-rank test compares the observed number of deaths with the expected number of deaths for group i. Consider the null hypothesis: $S_1(t) = S_2(t)$; i.e. there is no difference between survival curves in two groups. Given r_j andd_j, the random variable d_{1j} has the hypergeometric distribution

$$\frac{\binom{d_j}{d_{1j}}\binom{r_j - d_j}{r_{1j} - d_{1j}}}{\binom{r_j}{r_{1j}}}.$$

Under the null hypothesis, the probability of death at $t_{(j)}$ does not depend on the group, i.e., the probability of death at $t_{(j)}$ is $\frac{d_j}{r_j}$. So, that the expected number of deaths in group one is $E(d_{1j}) = e_{1j} = r_{1j} d_j r^{-1}{}_j$. The test statistic is given by the difference between the total observed and expected number of deaths in group one $U_L = \sum_{j=1}^{r}(d_{1j} - e_{1j})$. **(2.2)**

Since d_{1j} has the hypergeometric distribution, the variance of d_{1j} is given by

$$v_{1j} = V(d_{1j}) = \frac{r_{1j}r_{2j}d_J}{r^2{}_j(r_j-1)} \quad \textbf{(2.3)}$$

So, that the variance of U_L is $V(U_L) = \sum_{j=1}^{r} v_{1j} = V_{L.}$

Under the null hypothesis, statistic (2.2) has an approximate normal distribution with zero mean and variance $V_{L.}$ this then follows $\frac{U^2{}_L}{V_L} \sim x_1^2$. There are several alternatives to the log-rank test to test the equality of survival curves, for example, the Wilcoxon test [20]. These tests may be defined generally as $\frac{\sum_{j=1}^{r} w_j(d_{1j}-e_{1j})}{\sum_{j=1}^{r} w_j^2 v_{1j}}$. Where, w_j are weights whose values depend on the specific test. The Wilcoxon test uses weights equal to risk size at $t_{(j)}$, $w_j = r_j$, This gives less weight to longest survival times. Early failures receive more weight than later failures. The Wilcoxon test places more emphasis on the information at the beginning of the survival curve where the number at risk was large. This type of weighting may be used to assess whether the effect of treatment on survival was strongest in the earlier phases of administration and tends to be less effective over time. Whereas the log-rank test uses weights equal to one at $t_{(j)}$, $w_j = 1$. This gives the same weight to each survival time. Therefore, Wilcoxon statistic was less sensitive than the log-rank statistic to difference of d_{1j} from e_{1j} in the tail of the distribution of survival times.

The log-rank test is appropriate when hazard functions for two groups are proportional over time, i.e., $h_1(t) = \emptyset h_2(t)$: So, it is the most likely to detect a difference between groups when the risk of a failure was consistently greater for one group than another.

Table 2.3: Compare of survival distribution for diseases groups "Means and Medians for Survival Time".

Disease Type	Meana				Median			
	Estimate	Std. Error	95% Confidence Interval		Estimate	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound			Lower Bound	Upper Bound
Acute Renal Failure	624.119	77.280	472.650	775.588	.	.	.	.
Congenital Deformity Heart	322.634	61.279	202.527	442.740	202.000	86.146	33.154	370.846
Leukemia	304.162	19.961	265.039	343.286	.	.	.	.
Septicemia	194.599	31.240	133.368	255.829	32.000	6.977	18.326	45.674
Sickle cell disease	1693.109	31.152	1632.051	1754.167	.	.	.	.
Overall	1154.517	40.068	1075.983	1233.051	.	.	.	.

a. Estimation is limited to the largest survival time if it is censored.

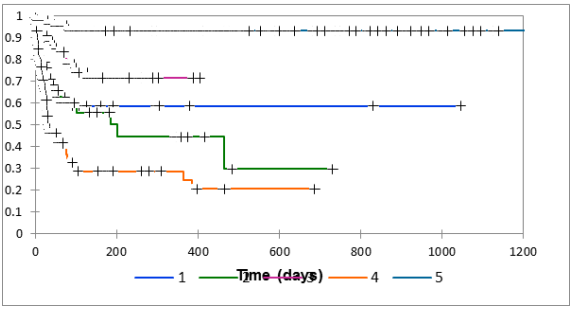


Figure 2.2: Survival distribution function / Event

3. COX REGRESSION MODELS

When have several prognostic variables, multivariate approaches must be used. But multiple linear regression or logistic regression could not use because they cannot deal with censored observations. One very popular model in survival data was the Cox proportional hazards model, which is proposed by Cox [12]. The Cox Proportional Hazards model is given by

$$h(t|x) = h_0\,(\mathrm{t})\,\exp(\beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_3 x_3) = h_0(\mathrm{t})\,exp(\beta' x),$$

where h_0(t) is called the baseline hazard function, which is the hazard function for an individual for whom all the variables included in the model are zero., $x = (x_1, x_2, \ldots, x_p)'$ is the values of the vector of explanatory variables for a particular individual, and $\beta' = (\beta_1, \beta_2, \ldots, \beta_p)$ is a vector of regression coefficients. The corresponding survival functions are related as $S(t|x) = S_0(\mathrm{t})^{\exp \sum_{i=1}^{p} \beta_i x_i}$.

The model is referred to as a semi-parametric model. The Cox approach for this vagueness creates no problems for estimation. Even though the baseline hazard is not specified, we could still get a good estimate for regression coefficients, hazard ratio, and adjusted hazard curves. The measure of effect is called hazard ratio. The hazard ratio of two individuals with different covariates x and x^* was

$$\widehat{HR} = \frac{h_0 \exp(\hat{\beta}' x)}{h_0 \exp(\hat{\beta}' x^*)} = \exp(\sum \hat{\beta}'(x - x^*).$$

This hazard ratio was time-independent, this why was called the proportional hazards model.

Table 3.1: Cox Regression Report for the Numerical Variables

Independent Variable	Regression Coefficient(B)	Standard Error of B	Risk Ratio Exp(B)	Mean	Wald Z-Value	Prob Level	Pseudo R^2
Gender	0.216603	0.138812	1.2419	1.4137	1.5604	0.1187	0.0118
Age	-0.001329	0.000286	0.9987	489.7507	-4.6475	0	0.0956
Stage	0	0	1	2.0189			
Symptoms	2.577964	0.183191	13.1703	1.2695	14.0725	0	0.4922
Disease Type	-0.190225	0.066023	0.8268	3.7278	-2.8812	0.004	0.039
Disease History	0	0	1	2.0189			
Height (cm)	-0.002792	0.004017	0.9972	57.7763	-0.6952	0.4869	0.0024
Weight (kg)	0.004757	0.012746	1.0048	7.385	0.3732	0.709	0.0007
Freq. Visits	-1.009026	0.212097	0.3646	1.2466	-4.7574	0	0.0997

Table 3.2: Confidence Limits the Numerical Variables

Independent Variable	Regression Coefficient(B)	Lower 95.0% Confidence Limit of B	Upper 95.0% Confidence Limit of B	Risk Ratio Exp(B)	Lower 95.0% C.L. of Exp(B)	Upper 95.0% C.L. of Exp(B)
Gender	0.216603	-0.05546	0.488669	1.2419	0.946	1.6301
Age	-0.001329	-0.00189	-0.00077	0.9987	0.9981	0.9992
Stage	0	0	0	1	1	1
Symptoms	2.577964	2.218916	2.937013	13.1703	9.1974	18.8594
Disease Type	-0.190225	-0.31963	-0.06082	0.8268	0.7264	0.941
Disease History	0	0	0	1	1	1
Height (cm)	-0.002792	-0.01067	0.00508	0.9972	0.9894	1.0051
Weight (kg)	0.004757	-0.02022	0.029738	1.0048	0.98	1.0302
Freq. Visits	-1.009026	-1.42473	-0.59332	0.3646	0.2406	0.5525

Table 3.3: Log Likelihood & R^2

Term(s) Omitted All Terms	*DF*	*Log Likelihood*	*R^2 of Remaining Term(s)*	*Reduction from Model R^2*
All Terms	9	-1288.87	0	0.4734
Gender	1	-1052.14	0.4717	0.0017
Age	1	-1062.64	0.4565	0.0169
Stage	1	-1050.93	0.4734	0
Symptoms	1	-1185.84	0.2425	0.2309
Disease Type	1	-1054.81	0.4679	0.0055
Disease History	1	-1050.93	0.4734	0
Height (cm)	1	-1051.17	0.4731	0.0003
Weight (kg)	1	-1050.99	0.4733	0.0001
Freq. Visits	1	-1067.31	0.4496	0.0238
None(Model)	9	-1050.93	0.4734	0

3.2 PARTIAL LIKELIHOOD ESTIMATE FOR COX PROPORTIONAL HAZARDS MODEL

Due to fit the Cox proportional hazards model, we wish to estimate h_0(t) and ß. One approach is to attempt to maximize the likelihood function for the observed data simultaneously with respect to h_0(t) and ß. A more popular approach is proposed by Cox [13] in which a partial likelihood function that does not depend on h_0(t) is obtained for ß. Partial likelihood is a technique developed to make inference about the regression parameters in the presence of nuisance parameters

(h_0(t) in the Cox PH model). we constructed in the partial likelihood function based on the proportional hazards model.

Let t_1, t_2, …, t_n be the observed survival time for n individuals. Let the ordered death time of r individuals be $t_{(1)}< t_{(2)}<$ … $< t_{(n)}$and let R ($t_{(j)}$) be the risk set just before $t_{(j)}$ and r_j for its size. So that R ($t_{(j)}$)is the group of individuals who are alive and uncensored at a time just prior to $t_{(j)}$. The conditional probability that the ith individual dies at $t_{(j)}$ given that one individual from the risk set on R ($t_{(j)}$) dies at $t_{(j)}$ is P(individual i dies at $t_{(j)}$ | one death from the risk set R($t_{(j)}$) at $t_{(j)}$)

$= \frac{P(individual\ i\ dies\ at\ t_{(j)})}{P(one\ death\ at\ t_{(j)})} = \frac{P(individual\ i\ dies\ at\ t_{(j)})}{\sum_{k\in R(t_{(j)})} P(one\ death\ at\ t_{(j)})} \simeq \frac{P(individual\ i\ dies\ at\ (t_{(j)},t_{(j)}+\Delta t)\}/\Delta t}{\sum_{k\in R(t_{(j)})} P(individual\ k\ dies\ at\ (t_{(j)},t_{(j)}+\Delta t)\}/\Delta t}$

$= \frac{\lim_{\Delta t\downarrow 0} P(individual\ i\ dies\ at\ (t_{(j)},t_{(j)}+\Delta t)\}/\Delta t}{\lim_{\Delta t\downarrow 0} \sum_{k\in R(t_{(j)})} P(individual\ k\ dies\ at\ (t_{(j)},t_{(j)}+\Delta t)\}/\Delta t} = \frac{h_i\, t_{(j)}}{\sum_{k\in R(t_{(j)})} h_k\, t_{(j)}} = \frac{h_0\, t_{(j)})\ exp(\beta'\ x_i\, t_{(j)})}{\sum_{k\in R(t_{(j)})} h_k\, t_{(j)})\ exp(\beta'\ x_k\, t_{(j)})}$

$= \frac{exp(\beta'\ x_i\, t_{(j)})}{\sum_{k\in R(t_{(j)})}\ exp(\beta'\ x_k\, t_{(j)})}$. Then the partial likelihood function for the Cox PH model is given by

$$L(\beta) = \prod_{j=1}^{r} \frac{exp(\beta'\ x_i\, t_{(j)})}{\sum_{k\in R(t_{(j)})}\ exp(\beta'\ x_k\, t_{(j)})}, \qquad \textbf{(3.1)}$$

in which $x_i\, t_{(j)})$ was the vector of covariate values for individual i who dies at $t_{(j)}$.The general method of partial likelihood was discussed by Cox [13].

Note that this likelihood function was only for the uncensored individuals. Let $, t_2, \dots, t_n$ be the observed survival time for n individuals and $\boldsymbol{\delta_i}$ be the event indicator, which is zero if the ith survival time is censored, and unity otherwise. The likelihood function in equation (3.1) can be expressed by

$$L(\beta) = \prod_{i=1}^{n} \left[\frac{exp(\beta'\ x_i\, (t_i)}{\sum_{k\in R(t_{(j)})}\ exp(\beta'\ x_k\, (t_i)}\right]^{\delta_i} \qquad \textbf{(3.2)}$$

Where R($t_{(j)}$) was the risk set at time t_i. The partial likelihood was valid when there were no ties in the dataset. That means there was no two subjects who had the same event time.

3.3 PROPORTIONAL HAZARD ASSUMPTION CHECKING

The main assumption of the Cox proportional hazards model was proportional hazards. PH mean that the hazard function of one individual was proportional to the hazard function of the other individual, i.e., the hazard ratio was constant over time.

3.3.1 GRAPHICAL METHOD

Cox PH survival function by the relationship between hazard function and survival function can obtained as $S(\mathrm{t}, x)= S_0(\mathrm{t})^{\exp(\sum_{i=1}^{p}\beta_i x_i)}$, where X= $(x_1, x_2, \dots, x_p)'$, Which is the values of the vector of explanatory variables for a patient. When taking the logarithm twice, we could easily have got $\ln[-\ln S(\mathrm{t,x})] = \sum_{i=1}^{p}\beta_i x_i + \ln[-\ln S_0(t)]$. Then the difference in log-log curves corresponding to two different individuals with variables $x_1 = (x_{11}, x_{12}, \dots, x_{1p})$ and $x_2 = (x_{21}, x_{22}, \dots, x_{2p})$ is given by $\ln[-\ln S(\mathrm{t}, x_1)] - \ln[-\ln S(\mathrm{t}, x_2)] = \sum_{i=1}^{p}\beta_i(x_{1i} - x_{2i})$, which does not depend on *t*.. By plotting estimated log(-log(survival)) versus survival time for two or more groups we would see parallel curves if the hazards were proportional. This method doesn't work well for continuous predictors or categorical predictors that have many levels because the graph became "cluttered". Furthermore, the curves are sparse when there are few time points and it may be difficult to tell how close to parallel was close enough. However, looking at the K-M curves and log(-log(survival)) was not enough to be certain of proportionality since they are univariate analysis and do not show whether hazards will still be proportional when a model includes many other predictors. But they supported our argument for proportionality. Some other statistical methods for checking the proportionality were shown below.

3.3.2 ADDING TIME-DEPENDENT COVARIATES IN THE COX MODEL

If the predictor of interest was X_j, then a time-dependent covariate creates X_j(t), X_j(t) = $X_j \times$ g (t), where g(t) was a function of time *t, log t* or Heaviside function of *t*. The model assessing PH assumption for X_j adjusted for other covariates is $h(t, x\,(t)) = h_0\,(t)\,exp[\beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_j x_j + \dots + \beta_p x_p + \delta x_j \times g(t)]$,

Where, $x\,(t) = \left(x_1, x_2, \dots, x_p, x_j(t)\right)'$ was the values of the vector of explanatory variables for a Children<5 years. The null hypothesis to check proportionality is that $\delta = 0$. The test statistic carried out using either a Wald test or a likelihood ratio test. In the Wald test, the test statistic was $w = (\hat{\delta}\,/se\,(\hat{\delta}\,))^{2}$.

The likelihood ratio test calculates the likelihood under null hypothesis, L_0 and the likelihood under the alternative hypothesis, L_α

The LR statistic was then $LR = -2\,ln\,(L_0 = L_\alpha) = -2\,(l_0 - l_\alpha)$, where l_0, l_α are log likelihood under two hypotheses respectively. Both statistics have a chi-square distribution with one degree of freedom under the null hypothesis. In the same way, the PH assumption for several predictors simultaneously could assessed.

3.3.3 TESTS BASED ON THE SCHOENFELD RESIDUALS

The other statistical test of the proportional hazards assumption is based on the Schoenfeld residual [48]. The Schoenfeld residuals are defined for each subject who is observed to fail. If the PH assumption holds for a particularly covariate then the Schoenfeld residual for that covariate will not be related to survival time. So, this test is accomplished by finding the correlation between the Schoenfeld residuals for a particularly covariate and the ranking of individual survival times. The null hypothesis was that the correlation between the Schoenfeld residuals and the ranked survival time was zero. Rejection of null hypothesis concludes that PH assumption is violated.

3.4 COX PROPORTIONAL HAZARDS MODEL DIAGNOSTICS

After a model has been fitted, the adequacy of the fitted model needs to be assessed. The model checking procedures below are based on residuals. when censored observations are presented and partial likelihood function is used in the Cox PH model, the usual concept of residual is not applicable. Many residuals have been proposed for use about the Cox PH model, three major residuals in the Cox model are described: The Cox-Snell residual, the deviance residual and the Schoenfeld residual. Then we will talk about influence assessment.

3.4.1 COX-SNELL RESIDUALS AND DEVIANCE RESIDUALS

The Cox-Snell residual is given by Cox and Snell [15]. The Cox-Snell residual for the ith individual with observed survival time t_i is defined as r_{ci}= exp $(\hat{\beta}'\mathcal{X}_i)$ $\widehat{H}_0(t_i)$= $\widehat{H}_i(t_i)$= - log $\hat{S}_i(t_i)$, where $\widehat{H}_0(t_i)$ is an estimate of the baseline cumulative hazard function at time t_i; which was derived by Kalbfliesch and Prentice [31].

The martingale residuals take values between negative infinity and unity. They have a skewed distribution with mean zero [3]. The deviance residuals are a normalized transform of the martingale residuals [53]. They also have a mean of zero but are approximately symmetrically distributed about zero when the fitted

model is appropriated. Deviance residual can also be used like residuals from linear regression. The plot of the deviance residuals against the covariates can be obtained. Any unusual patterns may suggest features of the data that have not been adequately fitted for the model. In a fitted Cox PH model, the hazard of death for the *i*th individual at any time depends on the value of $\exp(\beta' x_i)$ which is called the risk score.

3.4.2 SCHOENFELD RESIDUALS

All the above three residuals were residuals for each individual. Covariate-wise residuals will have described. Schoenfeld residuals [48]. The Schoenfeld residuals were originally called partial residuals because the Schoenfeld residuals for *i*th individual on the *j*th explanatory variable X_j is an estimate of the *i*th component of the first derivative of the logarithm of the partial likelihood function with respect to β_j. From equation (3.2), this logarithm of the partial likelihood function is given by $\frac{\partial \log L(\beta)}{\partial \beta_j} = \sum_{i=1}^{p} \delta_i \{x_{ij} - a_{ij}\}$, where x_{ij} the value of the *j*th explanatory variable j = 1, 2,...,p for the *i*th individual and $a_{ij} = \frac{\sum_{l \in R(t_i)} x_{jl} \exp(\beta' x_l)}{\sum_{l \in R(t_i)} \exp(\beta' x_l)}$. The Schoenfeld residual for *i*th individual on x_j is given by $r_{pji} = \delta_i \{x_{ji} - a_{ji}\}$. The Schoenfeld residuals sum to zero.

3.4.3 DIAGNOSTICS FOR INFLUENTIAL OBSERVATIONS

Observations that have an undue effect on model-based inference are said to be influential. In the assessment of model adequacy, it was important to determine whether are any influential observations. The most direct measure of influence is $\hat{\beta}_j - \hat{\beta}_{j(i)}$, where $\hat{\beta}_j$is the *j*th parameter, j = 1, 2, …, p, in a fitted Cox PH model and $\hat{\beta}_{j(i)}$ is obtained by fitting the model after omitting observation *i*. In this way, we must fit the n + 1 Cox models, one with the complete data and n with each observation eliminated. This procedure involves a significant amount of computation if the sample size was large. We would like to use an alternative approximate value that does not involve an iterative refitting of the model. To check the influence of observations on a parameter estimate, *Cain and Lange* [9] showed that an approximation to $\hat{\beta}_j - \hat{\beta}_{j(i)}$ is the *j*th component of the vector $r'_{S_i} V\ (\hat{\beta})$,

where $r'_{S_i}$is the $p \times 1$ vector of score residuals for the ith observation [10], which are modifications of Schoenfeld residuals and are defined for all the observations, and $V(\hat{\beta})$ is the variance-covariance matrix of the vector of parameter estimates in the fitted Cox PH model. The jth element of this vector is called delta-beta statistic for the jth explanatory variable, i.e., $\Delta_i\hat{\beta}_j \approx \hat{\beta}_j - \hat{\beta}_{j(i)}$. Therefore, we could check whether there are influential observations for any particularly explanatory variable.

3.5 STRATEGIES FOR ANALYSIS OF NON-PROPORTIONAL DATA

Supposed that statistic tests or other diagnostic techniques gave strong evidence of non-proportionality for one or more covariates. To deal with this, we would describe two popular methods: stratified Cox model and Cox regression model with time-dependent variables which were particularly simple and could be done using available software. Another way to consider was to use a different model. A parametric model such as an AFT model might be more appropriate for this dataset.

3.5.1 STRATIFIED COX MODEL

Which stratifies on the predictors not satisfying the PH assumption. The data are stratified into subgroups and the model is applied for each stratum. The model is given by $h_{ij}(t) = h_{0g}(t)\exp(\beta' x_{ig})$, where g represents the stratum. Note that the hazards were non-proportional because the baseline hazards may be different between strata. The coefficients β are assumed to be the same for each stratum g. A drawback of this approach was that we cannot identify the effect of this stratified predictor.

3.5.2 COX REGRESSION MODEL WITH TIME-DEPENDENT VARIABLES

The second method to consider is to model non-proportionality by time-dependent covariates. The violation of PH assumptions was equivalent to interactions between covariates and time. The PH model assumes that the effect of each covariate was the same at all points in time. If the effect of a variable varies with time, the PH assumption is violated for that variable. To model a time-dependent effect, one could

create a time-dependent covariate $X(t)$, then $\beta X(t) = \beta X \times g(t)$. $g(t)$ is a function of t such as t, $log\ t$ or Heaviside functions, etc. The choice of time-dependent covariates may be based on theoretical considerations and strong clinical evidence. The Cox regression with both time independent predictors X_i and time-dependent Covariates $X_j(t)$ can be written $h(t|x(t))=h_0(t)$ exp $[\sum_{i=1}^{p1} \beta_i X_i + \sum_{j=1}^{p2} \alpha_j x_j(t)]$. The hazard ratio at time t for the two individuals with different covariates x and x^* is given by $\widehat{HR}(t)=\exp\ [\sum_{i=1}^{p1} \hat{B}_i(x^*_i - X_i + \sum_{j=1}^{p2} \widehat{\alpha_j}\ (x^*_j(t) - X_j(t))]$. $\widehat{\alpha_j}$ represents overall effect of $X_j(t)$ considering all times at which this variable has been measured in this study. This means that the hazards of event at time t was no longer proportional and the model was no longer a PH model. One of the earliest applications of the use of time-dependent covariates was in the report by *Crowley and Hu* [16] on the Stanford Heart Transplant study. Time-dependent variables are usually classified to be internal or external.

3.6 COX PH MODEL

Univariate analysis used to check all the risk factors before proceeding to more complicated models. Then used a univariate Cox PHs regression for every potential risk factor. The Wald test is considered in each univariate Cox PH model. Variables are identified as significant using a 0.1 significance level in the univariate model. Then after fitted the full multivariate Cox PH model including all the potential risk factors and treatment arms.

3.7 COX MODEL WITH TIME-DEPENDENT VARIABLES

The Cox model has shown that displayed nonproportionality for variables, although the aim of study was not focus in the relation of variables and Children disease progression but, it's a believable there was a significant interaction between variables and their survivor time.

Figure 3.2: The status of Survivor time and Children<5 at risk per disease type

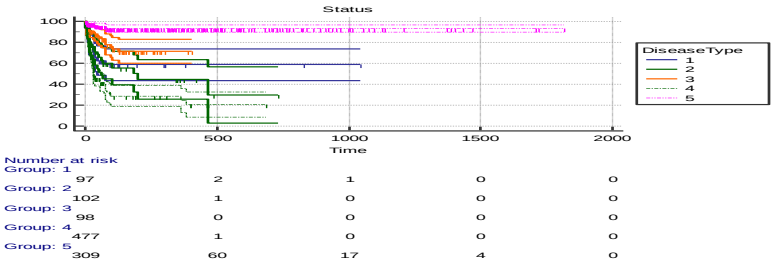
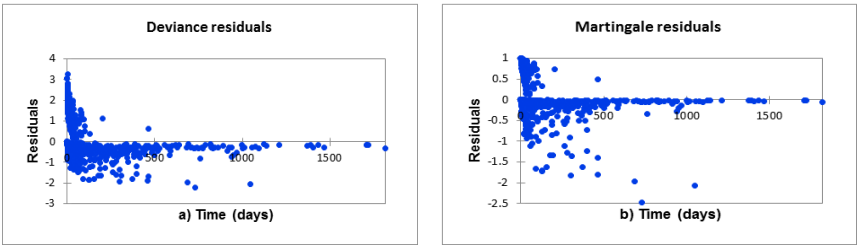


Figure 3.3: Deviance and Martingale residuals plotted against the risk score for Cox PH model



4. PARAMETRIC PROPORTIONAL HAZARDS MODEL

The parametric proportional hazards model was the parametric versions of the Cox proportional hazards model. It is given with the similar form to the Cox PH models. The hazard function at time t for the Children<5 years with a set of p covariates $(x_1, x_1, \ldots., x_p)$ is given as follows: $h(t \mid x) = h_0(t) \exp(B_1x_1 + B_2x_2 + \ldots + B_px_p) = h_0(t) \exp(B'x)$ Hazard ratios have the same interpretation and proportionality of hazards is still assumed. Many different parametric PH models may be derived by choosing different hazard functions. The commonly applied models were exponential, Weibull or Gompertz models.

4.1 WEIBULL PH MODEL

Suppose that survival times are assumed to have a Weibull distribution with scale parameter λ and shape parameter, so the survival and hazard function of a $W(\lambda,\gamma)$ distribution are given by $S(t) = exp(-\lambda t^{\gamma})$, $h(t) = \lambda\gamma(t)^{\gamma-1}$, with λ, $\gamma > 0$. The hazard rate increased when $\gamma > 1$ and decreased when $\gamma < 1$ as time goes on. When $\gamma = 1$, the hazard rate remains constant, which is the special exponential case.

Under the Weibull PH model, the hazard function of a particularly Children<5 years with covariates $(x_1, x_1, \ldots., x_p)$ is given by

$$h(t|x) = \lambda\gamma(t)^{\gamma-1} exp(B_1x_1 + B_2x_2 + \ldots + B_px_p) = h_0(t) = \lambda\gamma(t)^{\gamma-1}exp(B'x).$$

Survival time for this patient observed that has the Weibull distribution with scale parameter $\lambda\, exp(B'x)$, and shape parameter γ. Therefore the Weibull family with fixed γ possesses PH property. This shown the effects of the explanatory variables in the model alter the scale parameter of the distribution, while the shape parameter remained constant. From equation (1.4), the corresponding survival function is given by

$$S(t\,|\,x) = \exp\{-\exp(B'x)\lambda t^{\gamma}\}. \quad \textbf{(4.1)}$$

After a transformation of the survival function for a Weibull distribution, we obtained

$$Log\{-log\, S(t)\} = log\lambda + \gamma\, log\, t.$$

The $\{-log\, S(t)\}$ versus $log(t)$ should give approximately a straight line if the Weibull distribution assumption was reasonable. The intercept and slope of the line will be rough estimated of $\log\lambda$ and γ respectively. If the two lines for two groups in this plot were essentially parallel, this mean that the proportional hazards model was valid. Furthermore, if the straight line has a slope nearly one, the simpler exponential distribution was reasonable. In the other way, for exponential distribution, there was $\log S(t) = -\lambda t$. Thus, it could have considered the graph of $\log S(t)$ versus t. This should be a line that goes through the origin if exponential distribution was appropriate.

If the hazard function were reasonably constant over time, this would indicate that the exponential distribution might be appropriate. If the hazard function increased or decreased monotonically with increasing survival time, a Weibull distribution or Gompertz distribution might be considered.

4.2 EXPONENTIAL PH MODEL

The exponential PH model is a special case of the Weibull model when γ = 1. The hazard function under this model is to assume that it is constant over time. The survival and hazard function are written as $S(t) = exp(-\lambda t)$; $h(t) = \lambda$.Under the exponential PH model, the hazard

function of a particularly Children<5 years is given by $h(t\,|x) = \lambda\, exp(\beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_p x_p) = \lambda\, exp(\beta' x)$.

The piecewise exponential model [7] is an extension of the exponential PH model. For the piecewise exponential model, the period of follow-up is divided into k intervals $(t_j, t_{j+1}]$, $j=1, 2,\ldots,k$; $t_1 = 0$. Assume that the baseline hazard is constant within each interval but can vary across intervals, so
$h_0\,(t) = exp(\alpha_j) = \lambda_j$ *for* $t_j < t < t_{j+1}$, i.e., the baseline hazard function is approximated by a step function. The piecewise exponential model is given by $\lambda_{ij} = \lambda_j exp(\beta' x_i)$, where λ_{ij} is the hazard corresponding to individual i in interval j and exp $(\beta' x_i)$ is the relative risk for an individual with covariate value x_i compare to the baseline at any given time. In the piecewise exponential approach, a log-linear model is used to model both the effects of the covariates and the underlying hazard function. Estimates of the underlying hazard function and the regression parameters can be obtained using maximum likelihood, which estimates of the baseline hazard function in interval i for given regression coefficients β is given by $\widehat{\lambda_j} = \frac{d_j}{\sum_{i \in R_j} exp(\beta' x_i) t_{ij}}$, where d_j is the number of events in interval j, R_j is the risk set entering interval j and t_{ij} is the observed survival time for individual i in interval j. This approach was first studied by Holford [24], also the subject of work by Holford [25] and Laird and Olivier [36]. One of the greatest challenge related to the used of the piecewise exponential model was to find an adequate grid of time-points needed in its construction and one of the advantage of this method was the ability to incorporate time-dependent covariates.

4.3 GOMPERTZ PH MODEL

The survival and hazard function of the Gompertz distribution are given by

$$S(t) = exp(\frac{\lambda}{\theta}(1 - e^{\theta t}), h(t) = \lambda\, ex\, p(\theta t),$$

for $0 \leq t < \infty$ and $\lambda > 0$. The parameter θ determines the shape of the hazard function. When $\theta = 0$, the survival time then have an exponential distribution, which also a special case of the Gompertz distribution. Like the Weibull hazard function, the Gompertz hazard increases or decreases monotonically. For the Gompertz distribution, $log(h(t))$ is linear with t. Under the Gompertz PH model, the hazard

function of a Children<5 years was given by $h(t|x) = \lambda\, x\, p(\theta t)\, exp(\, B_1 x_1 + B_2 x_2 + \ldots + B_p x_p) = \lambda ex\, p(B'x)\, ext(\theta t)$. It is straightforward to see that the Gompertz distribution has the PH property. But the Gompertz PH model is rarely used in practice.

5. ACCELERATED FAILURE TIME MODELS

The module fitted the regression relationship between a positive-valued dependent variable (often time to failure) and one or more independent variables. The distribution of the residuals (errors) is assumed to follow the exponential, extreme value, logistic, log-logistic, lognormal, lognormal10, normal, or Weibull distribution. The data may include failed, left censored, right censored, and interval observations. This type of data often arises around accelerated life testing.

The models that predict failure rates at normal stress levels from test data on items that fail at high stress levels are called acceleration models.

Basic assumption of acceleration models was that failures happen faster at higher stress levels. That was the failure mechanism is the same, but the time scale has been changed (shortened).

Although parametric PH models are very applicable to analyze survival data, there were relatively few probability distributions for the survival time that can be used with these models. In these situations, the accelerated failure time model (AFT) is an alternative to the PH model for the analysis of survival time data. Under AFT models the study measured the direct effect of the explanatory variables on the survival time instead of hazard as done in the PH model. This characteristic allowed for an easier interpretation of the results because the parameters measure the effect of the correspondent covariate on the mean survival time. Currently, the AFT model is not commonly used for the analysis of clinical trial data, although it is fairly common in the field of manufacturing. Like the PH model, the AFT model describes the relationship between survival probabilities and a set of covariates.

For a group of Children<5 years with covariate $(x_1, x_1, \ldots., x_p)$, the model is written mathematically as $S(t|x) = S_0\big(t\big|\eta(x)\big)$, where $S_0(t)$ was the baseline survival function and η was an "acceleration

factor" that is a ratio of survival times corresponding to any fixed value of S(t). The acceleration factor is given according to the formula $\eta(x) = \exp(\alpha_1 x_1 + \alpha_2 x_2 + \ldots + \alpha_p x_p)$.

Under an accelerated failure time model, the covariate effects are assumed to be constant and multiplicative on the time scale that was the covariate impacts on survival by a constant factor (acceleration factor). According to the relationship of survival function and hazard function, the hazard function for an individual with covariate $X_1, X_2, \ldots, X_P$ is given by

$$h(t|x) = [(1|\eta(x)]h_0[(t|\eta(x)] \qquad \textbf{(5.1)}$$

The corresponding log-linear form of the AFT model with respect to time is given by

$$Log\ T_i = \mu + \alpha_1 X_{1i} + \alpha_2 X_{2i} + \ldots + \alpha_p X_{pi} + \sigma\varepsilon_i$$

where μ is intercept, σ is scale parameter and ε_i is a random variable, assumed to have a distribution. This form of the model is adopted by most software package for the AFT model. For each distribution of ε_i, there was a corresponding distribution for T. The members of the AFT model class include the exponential AFT model, Weibull AFT model, log logistic AFT model, log-normal AFT model, and gamma AFT model. The AFT models are discussed in detail in textbooks [10], [14], [37]. The AFT models are named for the distribution of T rather than the distribution of ε_i or $logT$.

Table 4.1: Summary of parametric AFT models

Distribution of ε	Distribution of T
Extreme value (1 parameters)	Exponential
Extreme value (2 parameters)	Weibull
Logistic	Log-logistic
Normal	Log-normal
Log-Gamma	Gamma

The survival function of T_i can be expressed by the survival function of ε_i

$S_i(t) = P(T_i \geq t) = P(logT_i \geq logt) = P(\ \mu + \alpha_1 X_{1i} + \alpha_2 X_{2i} + \ldots + \alpha_p X_{pi} + \sigma\varepsilon_i \geq logt) =$

$$P(\varepsilon_i \geq \frac{logt-\mu-\alpha x}{\sigma}) = S_{\varepsilon_i}(\frac{logt-\mu-\alpha x}{\sigma}). \qquad \textbf{(5.2)}$$

The distributions of ε_i and the corresponding distributions of T_i are summarized in Table (4.1). And the summary of the commonly used parametric models are described in Figure (4.1).

The effect size for the AFT model was the time ratio. The time ratio comparing two levels of covariate x_i (x_i = 1 vs. x_i = 0), after controlling all the other covariates was $\exp(\alpha_i)$, which is interpreted as the estimated ratio of the expected survival times for five groups. A time ratio above 1 for the covariate implies that this covariate prolongs the time to event, while a time ratio below 1 indicates that an earlier event is more likely. Therefore, the AFT models can be interpreted in terms of the speed of progression of a disease. The effect of the covariates in an accelerated failure time model was to change the scale, and not the location of a baseline distribution of survival times.

5.1 ESTIMATION OF AFT MODEL

AFT models are fitted using the maximum likelihood method. The likelihood of the n observed survival times, $t_1, t_2, \dots, t_n$ is given by $L(\alpha, \mu, \sigma) = \prod_{i=1}^{n}\{f_i(t_i)\}^{\delta_i}\{S_i(t_i)\}^{1-\delta_i}$, where $f_i(t_i)$ and $S_i(t_i)$ are the density and survival functions for the ith individual at t_i and δ_i is the event indicator for the ith observation. Using equation (5.2), the log-likelihood function is then given by $log\ L(\alpha, \mu, \sigma) = \sum_{i=1}^{n}\{-\delta_i\ log(\ \sigma t_i + \delta_i\ log\ f_{\varepsilon_i}(z_i) + (1-\delta_i) log S_{\varepsilon_i}(z_i))\}$, where z_i = (log $t_i - \mu - \alpha_i x_{1i} - \alpha_i x_{2i} - \cdots - \alpha_p x_{ni})/\sigma$ The maximum likelihood estimates of the $p+2$ unknown parameters $\mu, \sigma, \alpha_1, \alpha_2, \dots, \alpha_p$; are found by maximizing this function using the *Newton-Raphson procedure* in SAS, which was the same method used to maximize the partial likelihood in the Cox regression model.

Other approaches have been proposed for the estimation. Classical semi-parametric approaches to the AFT model that emphasize estimation of the regression parameters include the method of Buckley and James [8] and linear-rank-test-based estimators [32]. Despite theoretical advances, all these approaches complicated were numerically to implement, especially when the number of covariates was large.

5.2 WEIBULL AFT MODEL

Suppose the survival time T has $W(\lambda, \gamma)$ distribution with scale parameter λ and shape parameter. From equation (5.1), under AFT model, the hazard function for the ith individual is

$h_i(t) = [1/\eta_i(\mathrm{x})]\ h_0[\mathrm{t}/\eta_i(\mathrm{x})] = [1/\eta_i(\mathrm{x})]\ \lambda\gamma(\frac{\mathrm{t}}{\eta_i(x)})^{\gamma^{-1}} = 1/[\eta_i(x)]^{\gamma}\ \lambda\gamma\ (t)^{\gamma^{-1}}$

where $\eta_i = \exp(\alpha_1 x_{1i} + \alpha_2 x_{2i} + \ldots + \alpha_p x_{pi})$ for individual i with p explanatory variables. So, the survival time for the ith Children<5 years was $W(1/[\eta_i(x)]^{\gamma}\ \lambda, \gamma)$. The Weibull distribution has the AFT property. If T_i has a Weibull distribution, then ε_i has an extreme value distribution (Gumbel distribution). The survival function of Gumbel distribution is given by $S_{\varepsilon_i}(\varepsilon) = \exp(-\exp(\varepsilon))$.

From equation (5.2), the AFT representation of the survival function of the Weibull model is given by $S_i(t)=\exp[-\exp(\frac{\log t - \mu - \alpha_1 x_{1i} - \ldots - \alpha_p x_{pi}}{\sigma})] = \exp[-\exp(\frac{-\mu - \alpha_1 x_{1i} - \ldots - \alpha_p x_{pi}}{\sigma}) t^{1/\sigma}]$. **(5.4)**

From equation (4.1), the PH representation of the survival function of the Weibull model is given by

$S_i(t)=\exp\{-\exp(\beta_1 x_{1i} - \ldots - \beta_p x_{pi})\ \lambda t^{\gamma}\}$ **(5.5)**

Comparing the above two formulas (5.4) and (5.5), easily could see that the parameter λ, γ, β_j in the PH model can be expressed by the parameters μ, σ, β_j in the AFT model

$\lambda = \exp((-\mu/\sigma)$, $\gamma = 1/\sigma$, $\beta_j = \alpha_j/\sigma$ **(5.6)**

Using equation (1.3), the AFT representation of hazard function of the Weibull model is given by

$h_i(t) = \frac{1}{\sigma} t^{\frac{1}{\sigma}-1} \exp(\frac{-\mu - \alpha_1 x_{1i} - \ldots - \alpha_p x_{pi}}{\sigma})$ **(5.7)**

Suppose the pth percentile of the survival distribution for the ith individual is $t_i\ (p)$, which was the value such that $S_i\ (t_i\ (p)) = \frac{100-p}{100}$. From equation (5.4), we could easily get $t_i\ (p) = \exp[\sigma \log\{-\log(\frac{100-p}{100}) + \mu + \alpha' x_i]$. The median survival time is

$t_i\ (50) = \exp[\sigma \log(\log 2) + \mu + \alpha' x_i]$. **(5.8)**

To calculate the standard error of $\hat{\beta}_j$, we can use the approximate variance of a function of two parameter estimate θ_1, θ_2, which is given by $(\frac{\partial g}{\partial \theta_1})^2\ v(\hat{\theta}_1) + (\frac{\partial g}{\partial \theta_2})^2\ v(\hat{\theta}_2) + 2(\frac{\partial g}{\partial \theta_1}\ \frac{\partial g}{\partial \theta_2})\ cov\ (\theta_1, \theta_2)$.

The approximate variance of $\hat{\beta}_j$is expressed as $V(\beta_j) = (\frac{-1}{\hat{\sigma}})^2\ V(\hat{\alpha}_j) + (\frac{\hat{\alpha}_j}{\hat{\sigma}^2})^2\ V(\hat{\sigma}) + 2\left(\frac{-1}{\hat{\sigma}}\right)\left(\frac{\hat{\alpha}_j}{\hat{\sigma}^2}\right) Cov\ (\hat{\alpha}_j, \hat{\sigma})$.The square root of this is standard error of $\hat{\beta}_j$. Then 95% confidence interval can be calculated.

5.3 LOG-LOGISTIC AFT MODEL

Only limitation of the Weibull hazard function is that it was a monotonic function of time. However, the hazard function could change direction in some situations. The Weibull model would have

described in this section. The log-logistic survival and hazard function are given by

$$S(t) = \frac{1}{1+e^{\theta}\,t^{k}}\,, h\,(t) = \frac{e^{\theta}\,k\,t^{k-1}}{1+e^{\theta}\,t^{k}},$$

where θ and k are unknown parameters and $k > 0$. When $k \leq 1$, the hazard rate decreases monotonically and when $k > 1$, it increases from zero to a maximum and then decreases to zero.

Suppose that the survival times have a log-logistic distribution with parameter θ and k, then from equation (5.1), under the AFT model, the hazard function for the ith individual was

$$h_i(t) = (1/\eta_i)\,h_0(t/\eta_i) = \frac{e^{\theta}\,k\,(t/\eta_i)^{k-1}}{\eta_i(1+e^{\theta}\,(t/\eta_i)^{k})} = \frac{e^{\theta-k\log\eta_i}k\,t^{k-1}}{1+e^{\theta-k\,\log\eta_i}\,t^{k}}$$

where η_i= exp $((\beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p)$ for individual i with p explanatory variables. Therefore, the survival time for the ith individual has a log-logistic distribution with parameter $\theta - k\log\eta_i$ and k, log-logistic distribution has AFT property. If the baseline survival function is $S_0(t)$=$\{1+e^{\theta}\,\mathrm{t}^{k}\}^{-1}$, where θ and k are unknown parameters, then the baseline odds of surviving beyond time t are given by

$$\frac{S_0(t)}{1-S_0(t)} = e^{-\theta}\,\mathrm{t}^{-k}$$

The survival time for the ith individual also has a log-logistic distribution, which is

$$S_i(t) = \frac{1}{1+e^{\theta-k\,\log\eta_i\,\mathrm{t}^{k}}} \qquad \textbf{(5.9)}$$

Therefore, the odds of the ith individual surviving beyond time t is given by

$$\frac{S_i(t)}{1-S_i(t)} = e^{\log\eta_i-\theta}\mathrm{t}^{-k} \qquad \textbf{(5.10)}$$

We observed that the log-logistic distribution has the proportional odds (PO) property. So, this model also a proportional odds model, in which the odds of an individual surviving beyond time t are expressed as $\frac{S_i(t)}{1-S_i(t)}$ = exp $((\beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_p x_p\mathrm{i})\ \frac{S_0(t)}{1-S_0(t)}$. In a two-group study, using (5.10), the log (odds) of the ith individual surviving beyond time t are $\log[\frac{S_i(t)}{1-S_i(t)}] = \beta x_i - \theta - k\,\log t$,

Where, x_i is the value of a categorical variable which takes the value one in one group and zero in the other group. A plot of $log\,[(1 - S(t))/S(t)]$ versus log t should be linear if log-logistic distribution is appropriate. Therefore, we can check the suitability of log-logistic distribution using the PO property.

If T_i has a log-logistic distribution, then ε_i has a logistic distribution. The survival function of logistic distribution is given by $S_{\varepsilon_i}(\varepsilon) = \frac{1}{1+\exp(\varepsilon)}$.

Using equation (5.2), the AFT representation of survival function of the log-logistic model is given by

$$S_i(t) = [1 + t^{\frac{1}{\sigma}} \exp(\frac{-\mu-\alpha_1 x_{1i}- \ldots- \alpha_p x_{pi}}{\sigma})]^{-1} \quad \textbf{(5.11)}$$

Comparing the formula (5.9) and (5.11), we easily found a θ=- μ/σ, k=σ^{-1} According to the relationship of survival and hazard function, the hazard function for the *i*th individual is given by

$$h_i(t)=\frac{1}{\sigma t}\{1 + t^{\frac{1}{\sigma}} \exp(\frac{-\mu-\alpha_1 x_{1i}- \ldots- \alpha_p x_{pi}}{\sigma})\}^{-1} \quad \textbf{(5.12)}$$

The pth percentile of the survival distribution for the *i*th individual is $t_i(p)$, from equation (5.11), is $t_i(p) = \exp[\sigma \log\left(\frac{100-p}{100}\right) + \mu + \alpha' x_i]$. The median survival time is

$$t_i\ (50) = \exp\ (\mu + \alpha' x_i). \quad \textbf{(5.13)}$$

5.4 LOG-NORMAL AFT MODEL

When the survival times are assumed to have a log-normal distribution, the baseline survival function and hazard function are given by $S_0(t) = 1 - \Phi(\frac{logt-\mu}{\sigma})$, $h_0(t) = \frac{\phi(\frac{logt}{\sigma})}{[1-\Phi(\frac{logt}{\sigma})]\sigma t}$, where μ and σ are parameters, $\phi\ (x)$ is the probability density function and $\phi\ (x)$ is the cumulative density function of the standard normal distribution. The survival function for the *i*th individual is

$S_i(t) = S_0(t/\eta_i) = 1 - \Phi\ (\frac{logt-\alpha' x_i-\mu}{\sigma})$, where $\eta_i = \exp\ (\alpha_1 x_1 + \alpha_2 x_2 + \ \ldots + \alpha_p x_{pi})$.

Therefore the log survival time for the *i*th individual has normal $(\mu + \alpha' x_i, \sigma)$. The log-normal distribution has the AFT property. In a five-group study easily we could have got

$\Phi^{-1}[1 - S(t)] = \frac{1}{\sigma}(logt - \alpha' x_i - \mu)$, where x_i is the value of a categorical variable which takes the value one in one group and zero in another group. This implies that a plot of $\Phi^{-1}[1 - S(t)]$ versus $log\ t$ will be linear if the log-normal distribution is appropriate.

5.5 GAMMA AFT MODEL

The gamma model means the generalized gamma model in this paper. The probability density function of the generalized gamma distribution with three parameters λ, α and γ is defined by $f(t)=\frac{\alpha\lambda^{\alpha\gamma}}{\Gamma(\gamma)}t^{\alpha\gamma-1}\exp[-(\lambda t)^{\alpha}]$ $t>0,\ \gamma>0,\ \lambda>0,\ \alpha>0$, where, γ is the shape parameter of the distribution. The survival function and the hazard function do not have a closed form for the generalized gamma distribution. The exponential, Weibull and log-normal models are all special cases of the generalized gamma model. It was easily to seen that this generalized gamma distribution became the exponential distribution if $\alpha=\gamma=1$, the Weibull distribution if $\gamma=1$, and the log-normal distribution if $\gamma\longrightarrow\infty$. The generalized gamma model could take on a wide variety of shapes except for any of the special cases.

5.6 MODEL CHECKING

The graphical methods can be used to check if a parametric distribution fits the observed data specifically, if the survival time follows an exponential distribution, a plot of $log[-log\ S(t)]$ versus $log\ t$ should yield a straight line with slope of 1. If the plots are parallel but not straight then PH assumption holds but not the Weibull. If the lines for two groups are straight but not parallel, the Weibull assumption is supported but the PH and AFT assumptions are violated. The log-logistic assumption can be graphically evaluated by plotting $log[(1-S(t))/S(t)]$ versus $log\ t$.

If the distribution of survival function is log-logistic, then the resulting plot should be a straight line. For the log-normal distribution, a plot of $\Phi^{-1}[1-S(t)]$ versus log t should be linear.

Using quantile-quantile plot, an initial method for assessing the potential for an AFT model is to produce a quantile-quantile plot. For any value of p in the interval (0,100), the pth percentile was

$$t(p)=\ S^{-1}\left(\frac{100-P}{100}\right).$$

Let $t_0(p)$ and $t_1(p)$ be the pth percentiles estimated from the survival functions of the two groups of survival data. The percentiles for the two groups may be expressed as $t_0(p)=\ S_0^{-1}\left(\frac{100-P}{100}\right), t_1(p)=\ S_1^{-1}\left(\frac{100-P}{100}\right)$ where $S_0(t)$ and $S_1(t)$ are the survival functions for the two groups. So, we can get

$S_1[t_1(p)]=S_0[t_0(p)]$.

Under the AFT model, $S_1(t) = S_0(t/\eta)$, and so $S_1[t_1(p)] = S_0[t_1(p)/\eta)]$. Therefore, we have got $t_0(p) = \eta^{-1}t_1(p)$.

The percentiles of the survival distributions for two or more groups can be estimated by the K-M estimates of the respective survival functions. A plot of percentiles of the K-M estimated survival function from one group against another should give an approximate straight line through the origin if the accelerated failure time model is appropriate. The slope of this line will be an estimated of the acceleration factor η^{-1}.

Using statistical criteria, we could use statistical tests or statistical criteria to compare all these AFT models. Nested models can be compared using the likelihood ratio test. The exponential model, the Weibull model and log-normal model are nested within gamma model. For comparing models that are not nested, the Akaike information criterion (AIC) can be used instead, which is defined as

$$AIC = -2l + 2(k + c),$$

where l is the log-likelihood, k is the number of covariates in the model and c is the number of model specific ancillary parameters, the addition of $2(k + c)$ can be thought of as a penalty if non-predictive parameters are added to the model. Lower values of the AIC suggested a better model. But there was a difficulty in using the AIC in that there are no formal statistical tests to compare different AIC values. When two models have very similar AIC values, the choice of model may be hard and external model checking or previous results may be required to judge the relative plausibility of the models rather than relying on AIC values alone. Procedures based on residuals in the AFT model are particularly relevant with the Cox PH model. One of the most useful plots is based on comparing the distribution of the Cox-Snell residuals with the unit exponential distribution. The Cox-Snell residual for the ith individual with observed time t_i is defined as $r_{c\,i} = \widehat{H}(t_i/X_i) = -\log\,[\hat{S}(t_i/X_i)]$,

Where t_i is the observed survival time for individual i, X_i is the vector of covariate values for individual i, and $\widehat{S}(t_i)$ is the estimated survival function on the fitted model. From equation (5.2), the estimated survival function for the ith individual is given by $\hat{S}_i(t) = S_{\varepsilon_i}(\frac{logt-\hat{\mu}-\hat{\alpha}X_i}{\hat{\sigma}})$,

where $\hat{\mu}$, $\hat{\alpha}$ and $\hat{\sigma}$ are the maximum likelihood estimator of μ, α and σ respectively, $S_{\varepsilon_i}(\varepsilon)$ is the survival function of ε_i in the AFT model, and $\frac{logt-\hat{\mu}-\hat{\alpha}X_i}{\hat{\sigma}}$)= r_{s_i} is referred to as standardized residual.

The Cox-Snell residual can be applied to any parametric model. The corresponding form of residual based particularly AFT model can be obtained. For example, under the Weibull AFT model, since $S_{\varepsilon_i}(\varepsilon)$= exp (- e^{ε}), the Cox-Snell residual is then r_{c_i}=-log{$\widehat{S}(t_i)$} =-log $S_{\varepsilon_i}(r_{s_i})$= exp($r_{s_i}$).

Under the log-logistic AFT model, since $S_{\varepsilon_i}(\varepsilon)$= $(1 + e^{\varepsilon})^{-1}$, the Cox-Snell residual is then $r_{c_i} = \log[1 + \exp(r_{s_i})]$.

If the fitted model is appropriate, the plot of log (-log S(r_{c_i})) versus log r_{c_i} is a straight line with unit slope through the origin. These residuals led to the deviance residuals for the particularly AFT model. A plot of deviance residuals against the survival time or explanatory variables can be used to check whether there are times, or values of explanatory variables, for which the model is not a good fit.

6. RESULT

Firstly, descriptive statistics are used to give information about the distributions of the variables. We got the baseline characteristics in 1098 Children<5 years using the descriptive statistics in (Table 6.1).

Table 6.1: Study Cases summary

Number of events a	235	21.40%
Number censored b	863	78.60%
Total number of cases	1098	100.00%

a) Status = 1 b) Status = 0

Table 6.2: Overall Model Fit

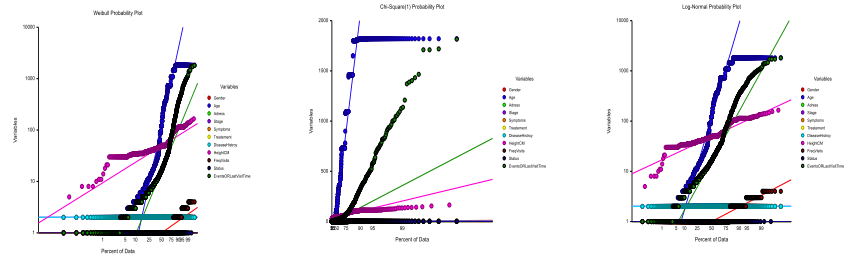
Null model -2 Log Likelihood	2996.218
Full model -2 Log Likelihood	2973.774
Chi-squared	22.444
DF	1
Significance level	P < 0.0001

Table 6.3: Coefficients and Standard Errors

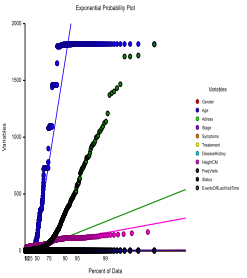
Covariate	*b*	*SE*	*Wald*	*P*	*Exp(b)*	*95% CI of Exp(b)*
Diseases Type	-0.2304	0.04688	24.1534	<0.0001	0.7942	0.7245 to 0.8707

Accelerated Failure Time Model

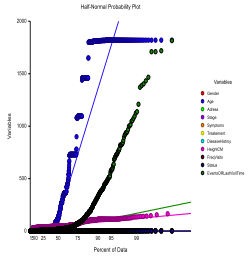
The AFT model could be used to express the magnitude of effect in a more accessible way in terms of difference between treatments in survival time. The dataset was fitted by using exponential, Weibull, log-logistic, log-normal and gamma AFT model. For each kind of model, when fitted both the univariate and multivariate AFT models, the independent variables (gender, age, date of 1st and last visits, stage when children arrived at hospital, symptoms, treatment, family history, child height (cm), child weight (kg), freq. of hospital visit, status, survivor time) were statistically significantly associated with the sample of diseases progression to Children<5 yrs. However, address and treatment didn't appear in the analysis but statistically and through observation were significant affect as the acceleration factors as well the corresponding confidence interval for every pair of groups manually. Statistically no interactions in multivariate AFT models. The results from the different AFT models applied to the time of diseases progression are presented in Figure 6. No significant difference for the estimations in the models. The Q-Q plot could use to check the AFT assumption and the Q-Q plot approximates well to a straight line from the origin indicating that the AFT model may provide an appropriate model.



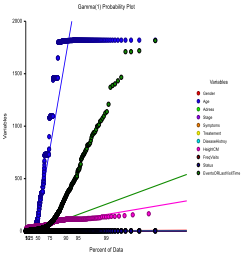
(a) (b) (c)



(d)



(e)



(f)

Figure 6.1: Probability Plot Comparison of AFT Models for time of diseases progression vs. % of children<5 years data

Table 6.4: Akaike Information Criterion (AIC) in the AFT models

Statistic	Independent	Full
Observations	235	235
DF	0	5
-2 Log(Likelihood)	2996.218	2728.487
AIC	2996.218	2738.487
SBC	2996.218	2755.785
Iterations	1	3

Table 6.5: The log-likelihoods and likelihood ratio (LR) tests, for comparing the variables selected alternative AFT models

Nbr. of variables	Variables	Variable IN/OUT	-2 Log(Likelihood)	Pr > LR
1	Freq. of visits	IN	75.133	0.000
2	Weight (kg)	IN	74.855	0.000
3	DiseaseType5	IN	66.242	0.000
4	Height (cm)	IN	41.685	0.000
5	DiseaseType-3	IN	9.815	0.081

Table 6.6: Test of the null hypothesis H0: beta=0

Statistic	DF	Chi-square	Pr > Chi²
-2 Log(Likelihood)	5	267.7306	< 0.0001
Score	5	206.5	< 0.0001
Wald	5	144.2533	< 0.0001

In addition to compare all these AFT models by used statistical criteria (likelihood ratio test and AIC) also nested AFT models can be compared by the likelihood ratio (LR) test or reliability Scale (Table 6.5). According to the LR test, the Weibull model and exponential model fits better. However, the LR test was not valid in compared

models that were not nested. In this case, AIC used to compare the models (The smaller AIC was the better). The Weibull AFT model and exponential model appears to be an appropriate AFT model according to AIC compared with other AFT models, although it was only slightly better than another AFT model. We also noted that the Lognormal10 and Loglogistic model are poorer fits according to reliability scale, LR test and AIC. This provided more evidence that the PH assumption for this data was not appropriate. Furthermore, residual plots can be checked by goodness of fit model. Noted that the exponential and Weibull AFT models are also PH models. The signs of the coefficients in the AFT model were opposite to the signs for the PH model. The estimate of shape parameter in Weibull model and exponential were approximately 0.6 and 1 respectively, this suggests that the Weibull model may be better than the exponential model.

Table 6.7: Reliability of Parametric Distributions per disease

Disease Type	Distribution	Likelihood	Shape	Scale	Threshold
1- Acute Renal Failure	Lognormal	-125.997	5.409682	2.433229	0
	Lognormal10	-125.997	2.349395	1.056738	0
	Loglogistic	-127.448	5.294861	1.397007	0
	Weibull	-129.386	0.567694	490.9924	0
	Exponential	-138.593	1	270.3333	0
	Normal	-172.699	349.3678	279.6106	0
	Logistic	-173.05	264.0041	134.1991	0
	Extreme Value	-182.169	617.0684	306.4729	0
2- Congenital Deformity Heart	Lognormal	-154.168	5.038924	2.065283	0
	Lognormal10	-154.168	2.188377	0.896941	0
	Loglogistic	-155.254	4.981978	1.178544	0
	Weibull	-157.036	0.672031	313.8727	0
	Exponential	-162.265	1	242.4	0
	Normal	-193.267	265.6699	200.9359	0
	Logistic	-194.933	233.582	111.8613	0
	Extreme Value	-202.677	422.3636	202.8603	0
3- Leukemia	Lognormal	-131.322	6.820292	2.8427	0
	Lognormal10	-131.322	2.962015	1.234569	0
	Loglogistic	-131.995	6.57346	1.514593	0
	Weibull	-132.338	0.602867	1122.779	0
	Exponential	-136.478	1	484.4737	0
	Normal	-152.436	295.1775	186.9853	0
	Logistic	-154.555	292.5666	112.1247	0
	Extreme Value	-156.314	362.3981	138.9928	0
4- Septicemia	Lognormal	-791.842	3.830626	1.678018	0
	Normal	-791.842	1.66362	0.728754	0
	Lognormal10	-791.842	1.66362	0.728754	0
	Loglogistic	-794.65	3.727124	0.930826	0
	Weibull	-812.412	0.745538	91.03753	0
	Exponential	-830.751	1	78.23225	0
	Logistic	-1020.71	60.34185	34.67565	0
	Extreme Value	-1153.51	235.2127	149.1812	0
5- Sickle cell disease	Lognormal	-129.094	13.68694	5.319224	0
	Lognormal10	-129.094	5.944163	2.31011	0
	Loglogistic	-130.251	12.33992	2.482912	0
	Weibull	-130.402	0.391783	311060.1	0
	Exponential	-144.212	1	5509.133	0
	Normal	-163.624	2199.519	1086.968	0
	Logistic	-165.746	2110.581	568.3794	0
	Extreme Value	-166.21	2272.061	609.8297	0

Table 6.8: Comparison of Cox PH model and AFT model

	Cox PH Model	AFT Model
Advantage	1. Widely used. 2. No assumption about the distribution for the survival time. 3. Survival curves can be estimated after adjusting for the explanatory variables. 4. Incorporation of time dependent covariate is convenient using SAS software	1. More informative. predicted hazard functions, predicted survival functions, median survival times and time ratios can be obtained. 2. The effect of covariate is to accelerate or delay the duration of illness by a constant amount (acceleration factor or time ratio). 3. The effect size is time ratio which is easier to interpret and more relevant to clinician
Disadvantage	1. PH assumption must hold. 2. Effect size is hazard ratio which is less relevant to clinician.	1. Relatively unfamiliar and rarely used. 2. AFT assumption must hold. 3. Need to specify the distribution of survival time, but an appropriate distribution may be difficult to identify. 4. Incorporation of time-dependent covariate is not allowed using SAS software

7. CONCLUSION

A finding of the present study was the absence of protection of Children<5 yrs. Preventive therapies on diseases progression, death and combined event of diseases progression. The study presented similar estimates of risk for the covariates with the baseline signs symptoms variables in the Cox PH model. To overcome this time-dependent covariate are incorporated into the Cox model. Also by used different AFT models to fit the data found that the Weibull and exponential AFT models fit a bit better for this dataset. The study provided the predicted hazard functions, predicted survival functions, median survival times and time ratios under the log-logistic AFT models. Thus, the independent variables were significantly associated with the diseases progression. Although the girls<5 years have longer survival time and disease progression time than boys<5 years but, their risks progression of mortality higher than boys<5 years. According to the log-logistic AFT model, variables prolongs the time to disease progression as it increases. Furthermore, AFT model makes it possible for clinicians to interpret the treatment benefit in terms of an effect on expected duration of illness. To this content the AFT model may have explanatory advantage in that covariates have a direct

effect on survival times rather on hazard functions as in the PH model.

In a review paper of survival analysis published in cancer journals [4], it was found that only five percent of all studies using the Cox PH models check PH assumption. However, PH assumption is not always satisfied in the data. If this assumption does not hold, there were various solutions to consider. One solution was to include the time-dependent variable for the predictors with non-proportional hazards. When this approach is used to account for a variable with non-proportionality, different results may be obtained from different choices of time-dependent variables. The stratified Cox model was not appropriate when the covariate with non-proportionality was continuous or of direct interests. And both ways were still based on comparison of hazards. The AFT model was an alternative method for the analysis of survival data even when hazards were not proportional. Based on asymptotic results, the AFT models should lead to more efficient parameter estimates than Cox model under certain circumferences [14], [45]. Further study of this data could attempt using a non-parametric version of the AFT model [54], which does not require the specification of the distribution can be applied in the dataset. The results from this model could then be compared with the standard AFT models and Cox PH models. In addition, further study can be carried out to evaluate the effects of practical cases such as large censoring dataset.

ACKNOWLEDGMENTS

I wish to thank *Dr. Ahmed Hamdi* and *Dr. Al Taiyb Ahmed* for carefully reviewing and who moderated this paper and in that line improved the manuscript significantly. I therefore, would also immensely grateful thanks the medical students at Sudan University of Science and Technology & the department of epidemiology at Jafar ibn Oaf Hospital for Children in Khartoum, Sudan whom have constructing the human model. As well the parents of Children for their significantly support.

REFERENCES

[1] Ackah AA, Coulibaly D, Digbeu H, et al. Response to treatment, mortality, and CD4 lymphocyte counts in HIV-infected persons with tuberculosis in abidjan, cote d.ivoire. *Lancet 345* (1995), 607.610.

[2] Andersen, P. K., Borgan, O., Gaill, R. D., and Keiding, N. *Statistical Methods Based on Counting Processes*. Springer, New York, 1993.

[3] Anderson, P. K., Borgan, O., and Gill, R. D. Cox.s regression model counting process: A large sample study. *Annals of Statistics* 10 (1982), 1100.1120.

[4] Breslow, N. E. Covariance analysis of censored survival data. *Biometrika* 30 (1974), 89.100.

[5] Buckley, I. V., and James, I. Linear regression with censored data. *Biometrika* (1979), 429.436.

[6] Collett, D. Modelling Survival Data *in Medical Research*. Chapman and Hall, London, 2003.

[7] Cox, C. Delta method. *In Encyclopedia of Biostatistics* (1998), P. Armitage and T. Colton, Eds., vol. 2, pp. 1125.1127.

[8] Cox, D. R. Regression models and life-tables. *Journal of the Royal Statistical Society. Series B 34* (1972), 187.220.

[9] Cox, D. R. Partial likelihood. *Biometrika* 62 (1975), 269.276.

[10] Cox, D. R., and Oakes, D. *Analysis of Survival Data*. Chapman and Hall, London, 1984.

[11] Cox, D. R., and Snell, E. J. A general definition of residuals with discussion. *Journal of the Royal Statistical Society. Series B 30* (1968), 248.275.

[12] Crowder, J., and Hu, M. Covariance analysis of heart transplant survival data. *Journal of American Statistical Association 78 (1977)*, 27.36.

[13] E. L. Kaplan and P. Meier, "Nonparametric Estimation from Incomplete Observations," *Journal of the American Statistical Association, Vol. 53, No. 282, 1958, pp. 457- 481*. doi:10.1080/01621459.1958.10501452.

[14] Fleming, T. R., and Harrington, D. P. *Counting Processes and Survival Analysis*. Wiley, New York, 1991.[19] Fleming A.F., Storey J., Molineaux L., Iroko E.A., Attai E.D. Abnormal hemoglobin's in the Sudan savanna of Nigeria. I. Prevalence of

haemoglobins and relationships between sickle cell trait, malaria and survival. Ann Trop Med Parasitol. 1979; 73:161–172. [PubMed].

[15] Gehan, E. A. A generalized Wilcoxon test for comparing arbitrarily singly censored samples. *Biometrika 52* (1965), 203.223.

[16] Gehan, E.A. (1969). Estimating Survival functions from the life table. Journal of Chronic Diseases, 21, 629-644.

[17] Giashuddin MS, Kabir M. Breastfeeding duration in Bangladesh and factors associated with it. Indian *J Comm Med 2003*; 28: 34-38.

[18] Greenwood, M. The natural duration of cancer. *Reports on Public Health and Medical Subjects 33 (1926), 1.26.*

[19] Holford, T. R. Life tables with concomitant information. *Biometrika 32* (1976), 587.597.

[20] Holford, T. R. The analysis of rates and of survivorship using log-linear models. *Biometrika 36* (1980), 299.305.

[21] IBRAHIM, J. G., CHEN, M.-H., and SINHA, D. *Bayesian Survival Analysis*. *Springer-Verlag*, New York, 2001.

[22] Johansen, S. The product limit estimate as a maximum likelihood estimate. *Scan-dinavian Journal of Statistics 5 (1978)*, 195.199.

[23] Jan, B., Ali Shah, W.S., Shah, S. and Qudir, F.M. (2005). Weighted Kaplan-Meier estimation of survival function in heavy censoring. Pakistan Journal of Statistics, 21(1),55-63.

[24] Jenkins, S. P. (1997). Discrete time proportional hazard regression. Stata Technical Bulletin, 39, 17-32.

[25] JP, N., MC, R., and A., K. HIV-associated tuberculosis in developing countries: epidemiology and strategies for prevention. *Tubercle and Lung Disease 73* (1992),311.321.

[26] Kalbfleisch, J. D., and Prentice, R. L. Marginal likelihoods based on Cox regression and life model. *Biometrika 60* (1973), 267.278.

[27] Kalbfleisch, J. D., and Prentice, R. L. *The statistical Analysis of Failure Time Data, 2nd ed. Wiley*, New York, 2002.

[28] Kaplan, E., and Meier, P. Nonparametric estimation from incomplete observations. *Journal of American Statistical Association 53* (1958), 457.481.

[29] Klembaum, D. G. Survival Analysis: A Self learning text. Springer, New York, 1996.

[30] Lawless, J. F. *Statistical Models and Methods for Lifetime Data Analysis. Wiley*,New York, 1982.

[31] Lee, E. T., and Wang, J. W. *Statistical Methods for Survival Data Analysis, 3rd ed*. Wiley, New York, 2003.

[32]Lim, HJ., Okwera, A., Mayanja-Kizza, H., Ellner, JJ., Mugerwan, RD., and Whalen, CC. Effect of tuberculosis preventive therapy on HIV disease progression and survival in HIV-infected adults. *HIV Clinical Trials 7* (2006), 172.183.

[33]Mantel, N. Evaluation of survival data and two new rank order statistics arising in its consideration. *Cancer Chemotherapy Report 50* (1966), 163.170.

[34]Mocroft AJ, Johnson MA, Sabin CA, Lipman M, Elford J, et al. Staging system for clinical aids patients. *Lancet 346* (1995), 12.17.

[35]Murray, J. F., and Cursed, d. HIV infection and tuberculosis. *Respiration 57* (1990), 210.220.

[36]Oakes, D. The asymtotic information in censored survival data. *Biometrika 64* (1977), 441.448.

[37]Schoenfeld, D. Partial residuals for the proportional hazards regression model *Biometrika 69* (1982), 239.241.

[38] Shafiq Mohammad, Shah Shuhrat, Alamgir (2007). Modified Weighted Kaplan-Meier Estimator.Pakistan Journal of Statistics Operation Research. 3(1), 39-44.

[39] Stapleton FB, Strother DRet al. (1988) Acute renal failure at onset of therapy for advanced stage Burkitt lymphoma and B cell acute lymphoblastic leukemia. Pediatrics 82: 863-869.

[40]Therneau, T. M., Grambsch, P. M., and Fleming, T. R. Martingale-based residuals for survival models. *Biometrika 77* (1990), 147.160.

[41] Wei, L. J. The accelerated failure time model: A useful alternative to the cox regression model in survival analysis. *Statistics in Medicine 11* (1992), 1871.1879.

[42] World Health Organization (WHO). *Global tuberculosis control: surveillance, planning, financing*. Geneva, 2007.