Science World Journal Vol 11 (No 3) 2016
www.scienceworldjournal.org
ISSN 1597-6343

Full Length Research Article

A STUDY OF THE SLOPE OF COX PROPORTIONAL HAZARD AND WEIBULL MODELS: SIMULATED AND REAL LIFE DATA APPROACH

Adejumo, A. O. and Ahmadu, A. O.

Department of Statistics,
University of Ilorin, Ilorin, Nigeria.

Email Address: aodejumo@unilorin.edu.ng

ABSTRACT
Parametric models require that the distribution of survival time is known and the hazard function is completely specified except for the values of the unknown parameters. These include the Weibull model, the exponential model, and the log-normal model. In this research work, Weibull Model is used for modelling survival time situations because it is flexible and also allows the inclusion of covariates. However, when the distributional assumptions for Weibull Model is not satisfied, Cox Proportional Hazard Model will be used, although semi-parametric, because it possessed a similar characteristic of covariates inclusion. The main objective of this research work is to determine if the cox proportional hazard model depend on the shape parameter of the Weibull model. And to investigate if there exist an advantage of using a parametric form of the survival distribution (Weibull distribution) instead of the semi parametric cox proportional hazard model when the parametric form of the model is known. This has two phases, the simulated and the real life data approach. We observed that the shape parameter of the Weibull model does not depend or have effect on the performance of the Cox Proportional Hazard model. And as the sample size increases the Mean Squared errors of the Maximum likelihood estimates of proportional hazard function of both the Weibull and Cox Proportional Hazard Models approximately the same.

Keywords: Cox Proportional Hazard Model, Weibull Model, Slope, Shape parameters, Scale parameter, Survival time.

INTRODUCTION
Survival analysis studies the amount of time that it takes before a particular event, such as death, occurrence of a disease, marriage, divorce, occurs. However, the same techniques can be used to study the time until any event occur. While a time-to-event study is theoretically simple to undertake, in practice there are a number of problems if the event being studied is relatively rare or takes a long time to occur. For instance, a study of death rates might be highly difficult to undertake if most subjects outlive the term of the study, or drop out of the study while it is in progress. Results from such analysis are used to help in calculating insurance premiums. Survival analysis is for analyzing the expected duration of time until one or more events happen, such as death in biological organisms, lifetime of the battery in a laptop, and the employment time of employees for a certain company and failure in mechanical systems.
Survival data could be derived from laboratory studies of animals or from clinical and epidemiologic studies. Survival data could relate to outcomes for studying acute or chronic diseases. Survival time refers to a variable which measures the time from a particular starting time (time initiated the treatment) to a particular endpoint of interest (attaining certain functional abilities). The time to event or Survival time can be measured in days, weeks, years, and so on. It is important to note that for some subjects in the study a complete survival time may not be available due to censoring.
To this effect, both uncensored and censored data were simulated and used as well as real life data on Tuberculosis (TB) Patients, collected from University College Hospital (UCH), Ibadan, with the Age, sex, date admitted, date discharged, length of stay, and censoring obtained from each of 132 TB patients. The time it takes for this event or disease to take place in a given individual is called survival time.
TB is a widespread disease and in many cases fatal infectious disease caused by various strains of mycobacteria. It is spread through the air when people who have an active TB infection, cough, sneeze, or otherwise transmit respiratory fluids through the air. Most infections do not have symptoms, known as latent tuberculosis (Mason, 2015).
Multiple factors contribute to the global increase in TB infection with the human immunodeficiency virus (HIV) which causes acquired immune deficiency syndrome (AIDS), is the single greatest risk progression of TB infection to active disease. People with HIV have a weakened immune system that increases their susceptibility to TB, and in these people, TB often progresses rapidly from the primary to the secondary stage. The increase of TB incidence is highest in Africa and Asia, areas with the highest number of people infected with HIV (WHO, 2009a; 2009b).
A second factor contributing to TB resurgence is the failure of patients to complete the full 6 months of antibiotics therapy required to cure the disease. Many people stop taking antibiotic when they begin to feel healthier, but unsuccessful treatment of TB requires therapy beyond the period of obvious symptoms. When patients fail to follow the prescribed treatment, they may become actively infectious, spreading the disease to others. Failure to complete the full round of treatment also can cause the emergence of TB bacterial strains with acquired drug resistance further complicating treatment by increasing the length and cost of therapy. Other factors are migration, international air travel, and tourism also had contributed to the global spread of TB. The extreme difficulty of screening immigrants and travelers for TB allows the disease to cross international borders easily. The substantial increase in homelessness, and the related circumstances of poverty, overcrowding, and malnutrition, also contributed to the increased incidence of TB in the United States and other industrialized countries during the early

1990s.Industrialised nations with good public health systems have been able to control the recent TB resurgence, curbing the spread of TB on a global scale will require ongoing international efforts. In the future, combating TB throughout the world will require advances in molecular biology, researches into the genetics of TB in order to understand drug resistance, and the continuous development of new drugs, as well as the prospect of synthesizing additional vaccines (Atun et al. 2005; Rios et al. 2000, Bacaër et al. 2008).
Mean Squared Error (MSE) and Standard Error (SE) are being used in comparing the two models, Weibull and Cox proportional hazard model, because we seek estimators that are unbiased and have minimal standard error.
The main objective of this research work is to determine if the Cox proportional hazard model depend on the shape parameter of the Weibull model. And to investigate if there exist an advantage of using a parametric form of the survival distribution (Weibull distribution) instead of the semi parametric Cox proportional hazard model when the parametric form of the model is known.

METHODOLOGY

Parametric models require that the distribution of survival time is known and the hazard function is completely specified except for the values of the unknown parameters. Examples include the Weibull model, the exponential model, and the log-normal model (Richards, 2012).

The analysis of these survival times is best done when all the survival times are known. However, there are many instances when this is not the case. Observations in this category are said to be censored data. A terminally ill patient may live to end of the study, or a mechanical component may not malfunction during the times it is being observed. In these cases, the survival times of the observations are not known, but it is known to be at least as long as the time of the study. This is called Type I censoring when all censored data have the same length (Singh and Mukhopadhyay, 2011, Brostrom, 2012; Angella, 2008).
Type II censoring is a type of censoring in which all individuals begin at the same time and the study is terminated once a specified number of failures is reached. The remaining observations are then censored to the point at which the longest uncensored observation failed (Angella, 2008; Richards, 2012; Brostrom, 2012).
Two important functions for describing survival data are the survival function and the hazard function

The survival function, $S(t)$, of an individual is the probability that they survive until at least time *t*

$$S(t) = \Pr(T > t) \quad \text{..... (1)}$$

where *t* is a time of interest and *T* is the time of event

The survival curve is non-increasing (the event may not reoccur for an individual) and is limited within [0,1]. Note that the event might not happen within our period of study and we call this right-censoring.
In terms of the cumulative distribution function F(t), the survival function can be written as:

$$S(t) = 1 - P(an \quad individual \quad fails \quad before \quad time \quad t) = 1 - F(t) \quad \text{.. (2)}$$

From this, it is easy to see that S(t) is non-increasing and has the following properties

$$S(t) = 1 \; for \quad t = 0$$

$$S(t) \to 0 \; as \quad t \to \infty \quad \text{..... (3)}$$

The rate of survival can be depicted using a survival curve, in which a steep curve would indicate a low rate and a gradual curve would represent a high rate of survival (Kalbfleisch,and Prentice, 2002; Brostrom, 2012).
The hazard function $\lambda(t)$ is a related measure, telling us the probability that the event *T* occurs in the next instant $(t + \delta t)$ given that the individual has reached time step *t*

$$\lambda(t) = \lim_{\delta t \to 0} \frac{\Pr(t \le T < t + \delta t / T > t)}{\delta t} \quad \text{..... (4)}$$

The hazard function $\lambda(t)$ is non-parametric, so we can fit a pattern of events that is not necessarily monotonic.
The hazard function is the rate of death/failure at an instant t, given that the individual survives up to time t. It measures how likely an observation is to fail as a function of the age of the observation. This function is also called the instantaneous failure rate or the force of mortality (Angella, 2008; Singh and Mukhopadhyay, 2011). It is also defined as

$$\lambda(t) = \frac{f(t)}{1 - F(t)} = \frac{f(t)}{S(t)} \quad \text{..... (5)}$$

where $f(t)$ is the probability density function of T.
Hence, in terms of the survival function,

$$\lambda(x) = -\frac{d}{dx} \log S(x) \quad \text{..... (6)}$$

Thus,

$$\log \; S(t) = -\int_0^t \lambda(x)\, dx \quad \text{..... (7)}$$

and since

$$S(0) = 1$$

$$S(t) = \exp\left\{-\int_0^t \lambda(x) dx\right\} \quad \text{..... (8)}$$

Therefore, the pdf of the distribution can be found from the hazard and survival functions

$$f(t) = \lambda(t) \exp\left\{-\int_0^t \lambda(x) dx\right\} \quad \text{..... (9)}$$

The Weibull model is characterized mainly by the shape and scale parameter of its distribution (Angella, 2008). The cumulative distribution function of the Weibull distribution is given as

$$F(t) = 1 - \exp\left\{-\theta t^{\gamma}\right\} \quad t > 0 \quad \text{..... (10)}$$

where θ is the scale parameter and γ is the shape parameter, and the probability density function of the Weibull distribution is

$$f(t) = \theta\gamma \exp\left\{-\theta t^{\gamma}\right\} t^{\gamma - 1} \quad t > 0 \quad \text{..... (11)}$$

The survival function and hazard function of the Weibull distribution are

$$S(t) = \exp\left\{-\theta t^{\gamma}\right\} \quad \text{..... (12)}$$

and

$$\lambda(t) = \theta\gamma t^{\gamma-1} \quad \text{..... (13)}$$

respectively, where θ is the scale parameter which is explained by covariate or explanatory variables given below

$$\theta = \exp(\beta X) = \exp(\beta_1 X_1 + \beta_2 X_2) \quad \text{..... (14)}$$

where β is a 1 x 2 vector of coefficients and X is a 2 x 1 vector of explanatory variables (age and sex as we have it in the data on Tuberculosis).

The parameter is estimated using Maximum Likelihood Estimation (MLE) method. The likelihood function is

$$L(\theta,\gamma) = f(t_1,t_2,...,t_n/\theta,\gamma) = \prod_{i=1}^{n} \theta\gamma \exp\left\{-\theta t^{\gamma}\right\} t^{\gamma-1} \quad \text{..... (15)}$$

The log-Likelihood is

$$\log L(\theta,\gamma) = \log(f(t_1,t_2,...,t_n/\theta,\gamma)) = \log\left(\prod_{i=1}^{n} \theta\gamma \exp\left\{-\theta t^{\gamma}\right\} t^{\gamma-1}\right) \quad \text{.... (16)}$$

$$\log L(\theta,\gamma) = n\log(\theta) + n\log(\gamma) - \sum_{i=1}^{n} \theta t_i^{\gamma} + (\gamma-1)\sum_{i=1}^{n} \log t_i \quad \text{..... (17)}$$

$\frac{\delta \log L(\theta,\gamma)}{\delta\theta} = 0$ only depends on survival time through t_i^{γ}.

The random variable $U = T^{\gamma}$ which has an exponential function

$$f(u) = \theta \exp(-\theta u), \quad u > 0 \quad \text{..... (18)}$$

Therefore, the distribution of the MLE of θ only depend on u and not on shape parameter **γ**.

The Cox Proportional Hazard model gives a semi-parametric method of estimating the hazard function at time *t,* given a baseline hazard that's modified by a set of covariates:

$$\lambda(t/X) = \lambda_0(t)\exp(\beta_1 X_1 + \ldots + \beta_p X_p) = \lambda_0(t)\exp(\beta X) \quad t > 0 \quad \text{.... (19)}$$

where $\lambda_0(t)$ is the non-parametric baseline hazard function, t is survival time and X is explanatory variables and βX is a linear parametric model using features of the individuals, transformed by an exponential function (Cox, 1972; Zhou, 2000; Singh and Mukhopadhya, 2011).

The partial likelihood (Breslow, 1974) for Cox proportional hazard model is given as

$$L_p(\beta) = \prod_{i=1}^{n} \left(\frac{\exp(X_i^T \beta)}{\sum_{k \in \Re} \exp(X_k^T \beta)} \right)^{\delta_i} \quad \text{..... (20)}$$

Assuming there is no ties, the log-partial likelihood is

$$\log L_p(\beta) = \sum_{i=1}^{n} \delta_i \left\{ X_i^T \beta - \log\left(\sum_{k \in \Re} \exp(X_k^T \beta)\right)\right\} \quad \text{..... (21)}$$

The partial likelihood depends only on the ordering of the survival times, not on the actual values, so it is invariant to monotonic transformation of time.

Simulation Study

A simulation study was done to compare the Mean Squared Errors (MSE) of Maximum Likelihood Estimate (MLE) of PH-slope of the Weibull and the Cox proportional hazards models when data are generated from a Weibull distribution.

The data were simulated from a Weibull distribution with survival function

$$S(t) = (e^{-t^2})e^{X} \quad \text{..... (22)}$$

That is, the model is Weibull with β = 1 for the slope of the covariate X, shape parameter γ = 2, and baseline hazard function $\lambda_o(t) = e^{-t^2}$. The values of the covariate $X = \{X_1, X_2\}$, with $X_1 \sim N(55,49)$ and $X_2 \sim binomial(1,0.5)$. The total sample sizes are 15, 45, 90, 180, 450 and 1200 with one third ($\frac{1}{3}$) of each as observations for each value of X. The data were simulated using the fact that the random variable U = F(T) has a uniform distribution. Where T is a Weibull random variable with cumulative distribution function F(t).

For this study, a value of T was obtained at

$$T = (-\ln(U)\beta e^{-X\beta})^{\frac{1}{\alpha}} \quad \text{..... (23)}$$

where $U \sim U(0,1), \alpha = 2$ is the Weibull shape parameter and *β=1* is the Weibull scale parameter.

The uniform random variable was generated using the R random number generator. Data were simulated without censoring and with 10% random censoring.

The parameters in the Weibull model may be estimated in R-program with the SURVREG, an R package designed for simulation in survival analysis with Weibull Model as underlying distribution, which uses the MSE method as stated in equations (16 to 18). The parameters in the Cox Proportional Hazards model may be estimated with the COXPH, another R package also designed for simulation in survival analysis with Cox Proportional Hazard Model as underlying distribution, which uses a form of a partial likelihood function as the default option as stated in equations (20 to 21). However, it is important to understand that SURVREG and COXPH procedures use different parameterizations in estimating the parameters. The coefficients that are estimated by the two procedures are not the same, but they are related. COXPH uses the model.

The maximum likelihood estimates of PH-slope using the parametric Weibull model was obtained from SURVREG procedure, as $\hat{\beta} = -\hat{\gamma}\hat{\delta}_1$ where $\hat{\gamma}$ is the estimate of the shape parameter and $\hat{\delta}_1$ is the estimate of the slope of the Weibull model as parameterized in the package. Since the shape parameter is known to be 2, an estimate of PH-slope that takes advantage of this fact, $-2\hat{\delta}_1$ was also obtained. The estimate of PH-slope from the Cox proportional hazards model was computed using COXPH procedure.

K=1000 replications of each sample sizes were run and the Mean Square and Standard Errors of the PH slope are obtained respectively as

$$MSE = \frac{\sum_{i=1}^{k} \left(\hat{\beta}_i - \beta\right)^2}{k} \quad \text{..... (24)}$$

$$SE = \frac{SDev\left(\hat{\beta}_i - \beta\right)^2}{\sqrt{k}} = \frac{\sigma}{\sqrt{k}} \quad \text{..... (25)}$$

Where β = 1, SDev is the standard deviation (σ). The distributions of $\hat{\beta}_i$ from the maximum likelihood estimates of the Weibull parameters and from the Cox proportional hazards model do not depend on the value of the shape parameter γ. Thus, the mean square errors apply to all Weibull shape parameters.

RESULTS AND DISCUSSION

Analysis of daily average wind speed data (2004-2014) showed various sample sizes were considered and the MSE and SE for each sample sizes were replicated 1000 times. R statistical package was employed for simulation and the data analysis.
For the simulated data, we considered two cases, uncensored and censored.

Case 1a: Uncensored Data

Table 1: MSEs and Standard Errors of PH Slope for Complete Samples on Uncensored Data

Sample Size (n)		Weibull MLE with γ Known	Weibull MLE with γ Unknown	Cox-PH Slope Estimate
15	MSE	0.1680	0.830354	0.5804
	SE	(0.028224)	(0.68900)	(0.33686)
45	MSE	0.2880	0.54374	0.5309
	SE	(0.082944)	(0.2957)	(0.28623)
90	MSE	0.5160	0.60922	0.6082
	SE	(0.266256)	(0.3711)	(0.37021)
180	MSE	0.6500	0.7897	0.7816
	SE	(0.4225)	(0.6237)	(0.61920)
450	MSE	0.5680	0.96459	0.95967
	SE	(0.322624)	(0.93043)	(0.91423)
1200	MSE	0.457687	0.6755433	0.6634332
	SE	(0.33453)	(0.51233)	(0.531245)

Table 1 has the means square errors and the standard error of PH Slope for the complete sample case. Here it can be seen that when the shape parameter is unknown, the estimates of the Cox proportion hazards model and the maximum likelihood estimates of the Weibull model perform almost similarly, but when the shape parameter is known, it far out-performs the Cox proportional hazards model (lower MSE). We can then conclude that when the distributional assumptions are not known, or are not met, the Cox proportional hazards model should be considered keeping in mind that the Weibull model when the distributional assumptions are not met stand a good chance as well.

We also observed that as the sample size increases from n=180 to 1200 the MSE's for maximum likelihood estimate of the Weibull is approximately the same as that of the cox proportional hazard models.

Case 1b: Censored data

Table 2: MSEs and Standard Errors of PH Slope for Complete Samples on Censored Data

Sample Sizes (n)		Weibull MLE with γ Known	Weibull MLE with γ Unknown	Cox-PH Slope Estimate
15	MSE	0.094	0.752583	0.2597
	SE	(0.008836)	(0.566)	(0.06744)
45	MSE	0.514	0.98227	0.7118
	SE	(0.264196)	(0.9649)	(0.50666)
90	MSE	0.9118	0.9699	0.95943
	SE	(0.83137)	(0.94071)	(0.9205)
180	MSE	0.676	0.8819	0.83981
	SE	(0.456976)	(0.77775)	(0.7053)
450	MSE	0.684323	0.867746	0.853465
	SE	(0.342554)	(0.23244)	(0.13365)
1200	MSE	0.6667688	0.685438	0.674001
	SE	(0.2344455)	(0.117056)	(0.21030)

Table 2 presented the results for the censored sample case. The patterns are similar to the uncensored sample case. The MSEs are smaller for the maximum likelihood estimates and the Cox proportional hazards model estimates of PH Slope when the shape parameter is known, but much bigger for the maximum likelihood estimates of the Weibull model when the shape parameter is unknown. The MSEs for censored data are larger than uncensored data.

Smaller MSEs of Cox proportional hazard model will give it an upper hand over Weibull model when the shape parameter is unknown. This suggests that for censored data, cox proportional hazard model should be preferred over Weibull model when the distributional assumptions are not met.

Results from a Real Life Data on Tuberculosis

The real life data were used in two forms; the original and the transformed survival time of Tuberculosis patients. The data were secondary data, consist of records of 132 patients admitted over the period of six years (2009 – 2014) collected from University College Hospital (UCH), Ibadan. The patients' Age, Sex, Length of stay in the hospital (in days, which is considered as survival time), as well as their censoring code (Dead (0) or Alive (1)) were observed in this work.

Case 2a: The Original survival time

Table 3: Results Using SURVREG Code in R for Untransformed Survival Time

Parameter	Estimate	Std error	z-value	Pr(>|z|)
Intercept	3.3837	0.03838	8.82	<0.0001
Age	-0.0166	0.0053	-3.13	0.0017
Sex	0.4249	0.2007	2.12	0.0343
Log(scale)	-0.1968	0.0942	-2.09	0.0367
Scale=0.821 chisq= 13.06 df=2 p-value=0.0015				

Table 4: R result COXPH (Estimates of PH Slope from Cox Proportional Hazard Model for Untransformed Survival Time)

Covariate	PH Slope Estimate	Std. error	z-value	Pr(>|z|)
Age	0.019008	0.006524	2.914	0.00357
Sex	-0.410526	0.250585	-1.638	0.10137
Signif. Codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1				

Table 5: Estimate of PH Slope from Weibull Model for Untransformed Survival Time

Covariate	Estimate	Shape Parameter	PH-Slope
Age	-0.0166	0.821	0.0136286
Sex	0.4249	0.821	-0.3488429

Survival data from 132 patients with tuberculosis, Age is a continuous covariate, Sex is categorical covariate which takes 0 for Male and 1 female and Censor indicates censoring where Censor = 1 is a censored observation. From Tables 4 and 5, the PH slope for both Cox proportional hazard and Weibull models are almost the same with the same pattern of directions and signs.

Case 2b: The Transformed survival time

Table 6: Results Using SURVREG Code in R for Transformed Survival Time

Parameter	Estimate	Std error	z-value	Pr(>|z|)
Intercept	1.69185	0.019192	8.82	<0.0001
Age	-0.0331	0.00265	-3.13	0.0017
Sex	0.81245	0.10037	2.12	0.0343
Log(scale)	-0.88994	0.09420	-9.45	<0.0001

Scale=0.411 chisq= 13.06 df=2 p-value=0.0015

Table 7: R result coxph (Estimates of PH Slope from Cox Proportional Hazard Model for Transformed Survival Time)

Covariate	PH Slope Estimate	Std. error	z-value	Pr(>|z|)
Age	0.019008	0.006524	2.914	0.00357
Sex	-0.410526	0.250585	-1.638	0.10137

Signif. Codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Table 8: Estimate of PH Slope of Weibull Model for Transformed Survival Time

Covariate	Estimate	Shape Parameter	PH-Slope
Age	-0.0331	0.411	0.0136041
Sex	0.81245	0.411	-0.33391695

From Tables 7 and 8, the PH slope for both Cox proportional hazard and Weibull models are approximately the same in patterns of direction and signs when the survival time is transformed. Lastly, comparing the two cases for real life data the PH slopes for the two are almost the same.

Conclusion

We observed that as the sample size increases the Mean Squared errors of the Maximum likelihood estimates of proportional hazard function of both the Weibull and Cox Proportional Hazard Models approximately the same.

It was noted that Cox proportional hazard model tends to be better, it exhibited smaller MSEs, than Weibull model when the shape parameter is unknown. This suggests that for censored data, Cox proportional hazard model should be preferred over Weibull model when the distributional assumptions are not met.

We also observed that the Weibull model is a better option for analyzing data on diseases like Tuberculosis used in this research work only if the distributional assumptions can be met and the shape parameter is known. But for uncensored data when the distributional assumptions are not met and shape parameter unknown, both models can be used interchangeably. The shape parameter of the Weibull model does not depend nor has effect on the performance of the proportional hazard model.

Finally, either of the two models could be adopted for modelling data on diseases like Tuberculosis based on their performances in this research.

REFERENCES

Angella, M.C. (2008). Comparison between cox proportional hazard and Weibull models. Kansas State University.

Atun R. A, Samyshkin Y. A, Drobniewski F, (2005) Seasonal variation and hospital utilization for tuberculosis in Russia: hospitals as social care institutions. *Eur J Pub Health*; 15: 350-4.

Bacaër N, Ouifki R, Pretorius C (2008). Modeling the joint epidemics of TB and HIV in a South African township. *J Math Biol*; 57: 557-93.

Breslow N. (1974). Covariance analysis of censored survival data. *Biometrics*, *30,* 89-99.

Brostrom, Göran (2012), *Event History Analysis with R (First ed.), Chapman & Hall/CRC, ISBN 978-1439831649*

Cox, D. (1972). Regression Models and Life Tables. Journal of the Royal Statistical Society, 34, 187–220.

Kalbfleisch, J. D.; Prentice, Ross L. (2002). *The statistical analysis of failure time data. New York: John Wiley & Sons. ISBN 047136357X*.

Kleinbaum, David G.; Klein, Mitchel (2012), *Survival analysis: A Self-learning text (Third ed.), Springer, ISBN 978-1441966452*

Mason, PH; Roy, A; Spillane, J; Singh, P (2015). "Social, Historical And Cultural Dimensions of Tuberculosis.". Journal of biosocial science: 1–27. *doi:10.1017/S0021932015000115. PMID 25997539*.

Richards, S. J. (2012). *"A handbook of parametric survival models for actuarial use". Scandinavian Actuarial Journal. Vol (4): 233–257. doi:10.1080/03461238.2010.506688*.

Rios M, Garcia J. M, Sanchez J. A, *et al*. (2000). A statistical analysis of the seasonality in pulmonary tuberculosis. *Eur J Epidemiol*; 16: 483-8.

Singh, R.; Mukhopadhyay, K. (2011). *"Survival analysis in clinical trials: Basics and must know areas". Perspect Clin Res. **2** (4): 145–148. doi:10.4103/2229-3485.86872*.

World Health Organization (WHO 2009a). Tuberculosis control in the South East Asia Region: *WHO annual report 2009*. Geneva.

World Health Organization (WHO 2009b). Global tuberculosis control 2009: epidemiology, strategy, financing: *WHO report 2009*. Geneva.

Zhou, Mai. (2000) Understanding the Cox regression model with time-change covariates.