Journal of Atomic, Molecular, Condensed Matter & Nano Physics
Vol. 9, No. 1, pp. 1–9, 2022
ISSN 2582-8215
Published by RGN Publications

http://www.rgnpublications.com

Special Issue:

Nuclear Track Detectors, Luminescence and Their Applications
Guest Editors: P.D. Sahare, Rajesh Kumar, Ashok Kumar

DOI: 10.26713/jamcnp.v9i1.1951

Research Article

Relative Risk for a Class of Patients Based on Progressive Censored Data Exposed to Radiations Using Cox's Proportional Hazard with a Set of Covariates

Shishir Kumar Jha

Department of Statistics, Ramjas College (University of Delhi), Delhi 110007, India
skjha_statistics@ramjas.du.ac.in

Received: December 16, 2021 **Accepted:** April 26, 2022

Communicated by: Tarun Kumar

Abstract. An attempt has been made to estimate Relative Risks (*RR*) of a category of patients assumed to be progressive censored and exposed to multiple hazards including nuclear radiations and suffering from chronic non-communicable disease by fitting Cox's proportional hazard regression model. Covariates are different age groups of patients, nature of stages of patients and treatment given to patients. The time dependent Weibull hazard rates have been estimated by using Maximum likelihood method. Any one of the three covariates considered here is taken as poorer immunity of a section of population because of exposure to high energy radiation of different kinds. The Relative Risk and Longevity estimates can further be used to construct life tables for such class of population, considering the censoring aspect of the data.

Keywords. Covariates, Hazard rate, Progressive censoring, Relative risk

PACS. 02.50.0-r

1. Introduction

The present investigation is an attempt to estimate Relative Risk for a randomly censored data of a group of patients (population) suffering from chronic non-communicable disease exposed to different kinds of hazards. One of such hazards is nuclear radiation, particularly in the vicinity of nuclear plants. Studies included atomic bomb survivors also. Ron [16] studies on epidemiology proved that exposure to radiations may lead to cancer causing effects on human beings. These studies proved that radiation exposure increases the risk of cancer in different organs including colon, stomach, breast, lung and skin. Further, Ron [15] found that about 15% of the ionizing radiation to the human beings is due to medical radiotherapy. Similarly, microscopic air pollutants particularly in congested large cities and industrial area and untreated water supply for drinking are hazards causing the communicable and non-communicable diseases. These findings may be utilized for constructing the life table for such type of population. Besides life expectancy, life tables are used to compute the indicators: death probabilities, probabilities of survival at some age, years of life lived and the number of survivors at different ages. Martel *et al*. [11], Chiang [3–5], Greville [7], Reed and Merrell [14], Keyfitz [9] and King [10] developed different methods of constructing life tables without considering the progressive censoring nature of the data. By using randomly censored data of a group of patients from different covariate groups, Relative Risks are estimated using Cox's hazard model. The three covariates are the condition of patients, the treatment given to patients and the category of patients with respect to age. On the line of Biswas and Jha [2] and Morgan and Elasoff [13], the hazard functions and hence the Relative Risks have been estimated using Weibull survival model. Martel *et al*. [11] described different methods for constructing life tables for different set of people. This leads to estimate the longevity of patients at different age groups and construction of life table for such populations.

2. Assumptions and Notations

Suppose that covariates under consideration X, Y and Z denotes nature of treatment, condition of patient with respect to morbidities and category of patient with respect to age respectively. Further, suppose that X, Y and Z are dichotomous random variables and can take only two values 0 or 1 such that

$$X = \begin{cases} 0 & \text{implies the standard treatment,} \\ 1 & \text{implies the new treatment;} \end{cases}$$

$$Y = \begin{cases} 0 & \text{implies that the patient has no premorbid condition prior to the onset} \\ & \text{of the disease,} \\ 1 & \text{implies that the patient has premorbid condition (poor immunity);} \end{cases}$$

$$Z = \begin{cases} 0 & \text{implies that the patient's age is} \le 60 \text{ years,} \\ 1 & \text{implies that the patient's age is} \ge 60 \text{ years.} \end{cases}$$

Let $h_{ijk}(t)$ denotes the Weibull hazard rate with respect to ith $(i = 0,1)$ treatment, jth $(j = 0,1)$ condition and kth $(k = 0,1)$ age group of the patient such that

$$h_{ijk}(t) = \theta_{ijk} t^{m_{ijk}-1}, \quad i,j,k = 1,2, \tag{2.1}$$

where

$$\theta_{ijk} = e^{\mu+\alpha X+\beta Y+\gamma Z} \tag{2.2}$$

is the function of covariates X, Y and Z and α, β and γ are corresponding parameters.

3. Methodology

The time dependent hazard rate is

$$h_{ijk}(t) = e^{\mu+\alpha X+\beta Y+\gamma Z} t^{m_{ijk}-1}. \tag{3.1}$$

Let

$$t^{m_{ijk}-1} = t^l \phi(\alpha,\beta,\gamma) \tag{3.2}$$

where

$$\phi(\alpha,\beta,\gamma) = ve^{\xi\alpha X+\psi\beta Y+\delta\gamma Z}, \tag{3.3}$$

l, v, ξ, ψ and δ are arbitrary constants.
Therefore,

$$t^{m_{ijk}-1} = t^l ve^{\xi\alpha X+\psi\beta Y+\delta\gamma Z}. \tag{3.4}$$

Substituting (3.4) in (3.1), we get

$$\begin{aligned} h_{ijk}(t) &= e^{\mu+\alpha X+\beta Y+\gamma Z} t^l ve^{\xi\alpha X+\psi\beta Y+\delta\gamma Z} \\ &= ct^l ve^{\alpha' X+\beta' Y+\gamma' Z} \end{aligned} \tag{3.5}$$

where

$$c = ve^v,$$

$$\alpha' = \alpha + \xi\alpha,$$

$$\beta' = \beta + \psi\beta,$$

$$\gamma = \gamma + \delta\gamma,$$

$h_{ijk}(t) = ct^l ve^{\alpha' X+\beta' Y+\gamma' Z}$ is the Cox's regression model (cf., Cox [6]),

where $ct^l = h_0(t)$ is the baseline hazard rate and $e^{\alpha' X+\beta' Y+\gamma' Z}$ is the hazard rate corresponding to covariates X, Y and Z. $h_0(t)$ and $e^{\alpha' X+\beta' Y+\gamma' Z}$ are independent.
The estimate of Relative risk (RR) between $X = 1$ and $X = 0$ is

$$\widehat{RR(X)} = \frac{h_{1jk}(t)}{h_{0jk}(t)} = e^{\alpha'}. \qquad \text{(using eq. (3.5))} \tag{3.6}$$

Similarly, the estimate of Relative risk (RR) between $Y = 1$ and $Y = 0$ is

$$\widehat{RR(Y)} = \frac{h_{i1k}(t)}{h_{i0k}(t)} = e^{\beta'} \tag{3.7}$$

and the estimate of Relative risk (RR) between $Z = 1$ and $Z = 0$ is

$$\widehat{RR(Z)} = \frac{h_{ij1}(t)}{h_{ij0}(t)} = e^{\gamma'}. \tag{3.8}$$

Estimation of α', β' and γ' by Partial Likelihood Method Given by Cox [6]. Let us assume that m patients quit the trial due to death or censoring out of a sample of n patients. (i.e. m patients leave the trial by time t out of a set of n patients).

Therefore,

$$\frac{h_0(t)e^{\alpha' T_{1i}+\beta' T_{2i}+\gamma' T_{3i}}}{h_0(t)\sum_1^n e^{\alpha' T_{1i}+\beta' T_{2i}+\gamma' T_{3i}}} \tag{3.9}$$

is the probability of the patient i $(i = 1,2,3,\dots,n)$ leaving the trial in $(0,t)$, where T_{ji} is the value of jth covariate of ith individual $(j = 1,2,3;\ i = 1,2,3,\dots,n)$.
Therefore,

$$P_L = \prod_{i=1}^{m} \frac{h_0(t)e^{\alpha' T_{1i}+\beta'' T_{2i}+\gamma' T_{3i}}}{h_0(t)\sum_1^n e^{\alpha' T_{1i}+\beta' T_{2i}+\gamma' T_{3i}}} \text{ is the Cox's partial likelihood.}$$

Using Maximum Likelihood Method to estimate α', β' and γ', the three estimating equations are

$$\sum_{i=1}^{m} T_{1i} - \frac{m\sum_{i=1}^n T_{1i}e^{\alpha' T_{1i}+\beta' T_{2i}+\gamma' T_{3i}}}{\sum_{i=1}^n e^{\alpha' T_{1i}+\beta' T_{2i}+\gamma' T_{3i}}} = 0, \tag{3.10}$$

$$\sum_{i=1}^{m} T_{2i} - \frac{m\sum_{i=1}^n T_{2i}e^{\alpha' T_{1i}+\beta' T_{2i}+\gamma' T_{3i}}}{\sum_{i=1}^n e^{\alpha' T_{1i}+\beta' T_{2i}+\gamma' T_{3i}}} = 0, \tag{3.11}$$

$$\sum_{i=1}^{m} T_{3i} - \frac{m\sum_{i=1}^n T_{3i}e^{\alpha' T_{1i}+\beta' T_{2i}+\gamma' T_{3i}}}{\sum_{i=1}^n e^{\alpha' T_{1i}+\beta' T_{2i}+\gamma' T_{3i}}} = 0. \tag{3.12}$$

Using the estimates of α', β' and γ' and equations (3.6), (3.7) and (3.8); $\widehat{RR(X)}$, $\widehat{RR(Y)}$, and $\widehat{RR(Z)}$ are estimated. We get different sets of equations (3.10), (3.11) and (3.12) for different values of t and hence different sets of estimates of α', β' and γ'.

Estimation of c and l of $h_0(t)$. Put $X = 0$, $Y = 0$ and $Z = 0$ in (3.5), we get

$$h_0(t) = ct^l. \tag{3.13}$$

Putting

$$l = l' - 1 \text{ and } c = c'l', \tag{3.14}$$

we get

$$h_0(t) = c'l't^{l'-1}. \tag{3.15}$$

Therefore,

$$f(t) = c'l't^{l'-1}e^{-c't^{l'}}. \qquad \text{(Since, } h(t) \text{ is Weibull hazard rate)} \tag{3.16}$$

Suppose that n is the total number of observations and n' is the total number of uncensored observations.
Therefore, Maximum likelihood equations are given (cf. Miller [12])

$$\frac{n'}{c'} - \sum_{t=1}^{n} t^{l'} = 0, \tag{3.17}$$

$$\frac{n'}{l'} - \sum_{t'=1}^{n'} \log t' - c'\sum_{t=1}^{n} t^{l'} \log t = 0, \tag{3.18}$$

where t' is the uncensored observation's failure time.
On solving these equations, we get the estimates of l' and c'.
Therefore, $\hat{l} = l' - 1$ and $\hat{c} = c'l'$

We take the initial values of l' as $(l')_0$ and c'as $(c')_0$ and use Newton-Raphson method to solve for l' and c'.
$(l')_0$ and $(c')_0$ are solutions for fitting the baseline hazard rate (uncensored part of data).
i.e. $h_0(t') = ct'^l$; t' denotes uncensored observation's failure time. Taking log, we get

$$H_0(t') = c_1 + l\tau\,,$$

where $H_0(t') = \log h_0(t')$, $c_1 = \log(c)$ and $\tau = \log(t')$.
Therefore,

$$\sum_{t'} H_0(t') = \sum_{t'} c_1 + l \sum_{t'} \tau \tag{3.19}$$

$$\Rightarrow \quad \sum_{t'} \tau H_0(t') = \sum_{t'} (\tau c_1) + l \sum_{t'} (\tau^2) \tag{3.20}$$

Solving (3.19) and (3.20), we get estimates of c and l as $(c)_0$ and $(l)_0$.
Therefore,

$$(l')_0 = (l)_0 + 1 \text{ and } (c')_0 = \frac{(c)_0}{(l')_0}. \qquad \text{(using (3.14))} \tag{3.21}$$

Let, $h_{ijk}(t') = \widehat{h(t')}$, then

$$\widehat{h(t)} = h_0(t') e^{\alpha' X + \beta' Y + \gamma' Z}. \tag{3.22}$$

Further,

$$h_{ijk}(t') = \widehat{h(t')} = -\frac{1}{S(t')} \frac{\Delta S(t')}{\Delta t'} \tag{3.23}$$

and $S(t') = e^{-\int_0^{t'} h(t)dt} = P$ (surviving up to time t').
Therefore, we have

$$h_0(t') = \frac{\widehat{h(t')}}{e^{\alpha' X + \beta' Y + \gamma' Z}} = \frac{-\frac{1}{S(t')} \frac{\Delta S(t')}{\Delta t'}}{e^{\alpha' X + \beta' Y + \gamma' Z}}. \tag{3.24}$$

Now, we have set of covariates as $(0,0,0)$, $(0,0,1)$, $(0,1,0)$, $(0,1,1)$, $(1,0,0)$, $(1,0,1)$, $(1,1,0)$ and $(1,1,1)$. The corresponding baseline hazard rates are $h_0^{000}(t')$, $h_0^{001}(t')$, $h_0^{010}(t')$, $h_0^{011}(t')$, $h_0^{100}(t')$, $h_0^{101}(t')$, $h_0^{110}(t')$ and $h_0^{111}(t')$. Average of these hazard rates may be taken as an estimate of $h_0(t')$.

4. Numerical Illustration

A randomly censored data of 195 cancer patients taken from Kalbfleisch and Prentice [8] is used for estimating Relative Risk with a set of three covariates. The data comes from the section of the population exposed to radiations emanating from nuclear plants, atom bomb explosions/testing and exposure to radiotherapy such as X rays. Out of 195 patients, 142 patients are uncensored (i.e. $n = 195$ and $n' = 142$). Since, experiment is stopped at the end of 1800 days, therefore, one patient who survived beyond 1800 weeks are excluded for the analysis purposes as summarized in Table 1.

Table 1. Distribution of 195 patients

Time (days) up to	Number of patients	$\sum T_{1i}$	$\sum T_{2i}$	$\sum T_{3i}$	$\sum T_{1i}T_{2i}$	$\sum T_{1i}T_{3i}$	$\sum T_{2i}T_{3i}$	$\sum T_{1i}T_{2i}T_{3i}$
300	66	36	34	31	21	18	20	12
600	126	69	45	65	27	34	28	16
900	153	78	48	78	29	39	31	18
1200	174	84	49	92	29	43	32	18
1500	188	92	51	101	30	49	33	19
1800	194	94	51	103	30	49	33	19

Estimates of α', β' and γ' and hence $\widehat{RR(X)}$, $\widehat{RR(Y)}$ and $\widehat{RR(Z)}$ along with baseline hazard rates for different time period are obtained by solving equations (3.10), (3.11) and (3.12) and using equations (3.6), (3.7) and (3.8). The results are summarized in Table 2. The Relative Risk with variations in the values of covariates with respect to time is shown in Figure 1, Figure 2 and Figure 3. Similarly, survival probabilities and the life expectancy at different time period is obtained. This proceeds towards an idea of constructing a life table for the people exposed to a set of risks, viz. nuclear radiations, pollution etc.

Table 2. Estimates of Relative Risks over time

Time (days) up to	α'	β'	γ'	$h_0(t)$	$\widehat{RR(X)}$	$\widehat{RR(Y)}$	$\widehat{RR(Z)}$
300	−0.167	0.460	0.506	0.001103	0.846	1.1584	1.658
600	−0.133	0.983	−0.236	0.001081	0.876	2.674	0.790
900	0.050	1.056	−0.158	0.001196	1.051	2.876	0.854
1200	0.228	1.196	−0.289	0.001516	1.256	3.306	0.749
1500	0.136	1.100	−0.242	0.002029	1.146	3.004	0.785
1800	0.120	1.160	−0.172	0.002254	1.128	3.189	0.844

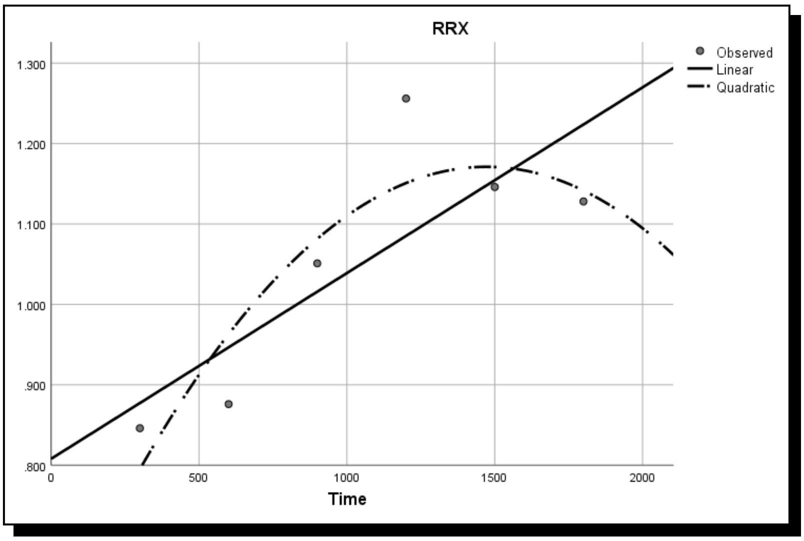

Figure 1. $RR(X)$ curve

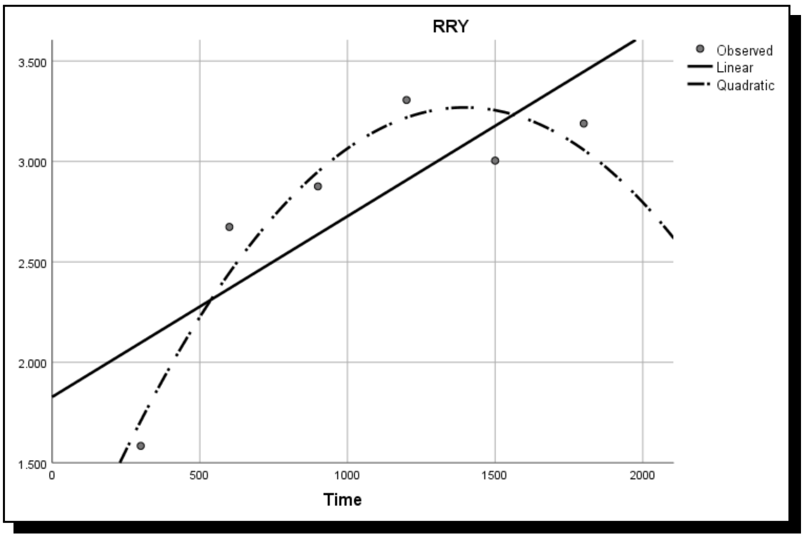

Figure 2. $RR(Y)$ curve

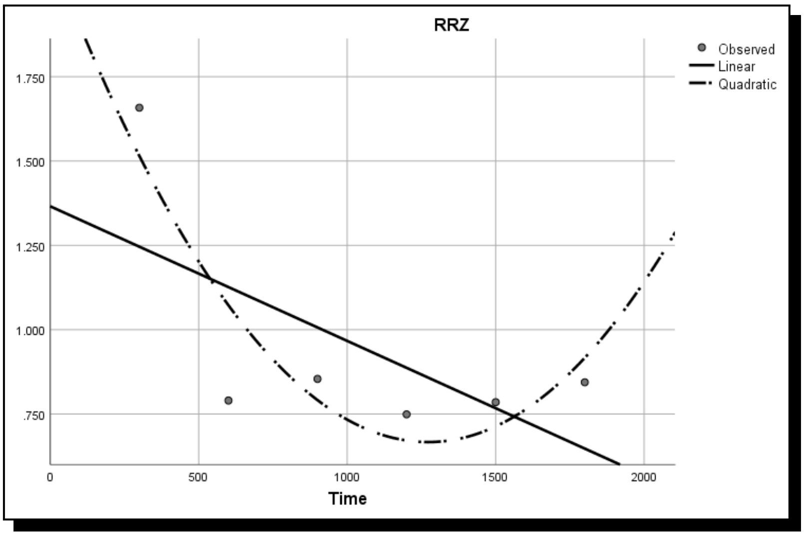

Figure 3. $RR(Z)$ curve

5. Conclusion

Figure 1 and Table 2 show that the Relative Risk of new experimental treatment in comparison to standard treatment increases with time, however the advantage of standard treatment decreases with the passage of time. Similarly, Figure 2 and Table 2 clearly show that the Relative Risk of patients with history of pre-morbidity is higher and the disadvantage of poorer health condition get reduced with time.This clearly suggests that radiation (gamma rays, x rays, α particles, β particles etc.) exposed population assumed to have low immunity has higher hazard of getting trap into cancer resulting in higher Relative Risk in comparison to unexposed population. Findings of Figure 3 and Table 2 with respect to impact of age on Relative Risk indicates that persons with higher age group has certain advantage in the beginning (this is possible because of a disciplined life style of senior citizens in comparison to younger people), however with the passage of time it gets reduced and Relative Risk increases.The concept can further be extended to construct mortality tables for the people exposed to the risk of certain radiations causing chronic diseases.

Competing Interests

The author declares that he has no competing interests.

Authors' Contributions

The author wrote, read and approved the final manuscript.

References

[1] S. Biswas and G.L. Sriwastav, *Stochastic Processes in Demography and Applications*, 1st edition, New Central Book Agency, Delhi (2011).

[2] S. Biswas and S.K. Jha, On the estimation of relative risks and longevities of a class of cancer patients based on randomly censored data, *Biometrical Journal* **36**(7) (1994), 885 – 892, DOI: 10.1002/bimj.4710360711.

[3] C.L. Chiang, A new life table technique as applied to risk assessment for a stochastic system, in: *43rd session of International Statistical Institute*, Buenos Aires, Argentina (1981).

[4] C.L. Chiang, A stochastic study of the life table and its applications: I. Probability distributions of the biometric functions, *Biometrics* **16** (1960), 618 – 635, DOI: 10.2307/2527766.

[5] C.L. Chiang, Constructing current life table, *Journal of the American Statistical Association* **67**(339) (1972), 538 – 541, DOI: 10.1080/01621459.1972.10481245.

[6] D.R. Cox, Regression models and life-tables, *Journal of Royal Statistical Society, Series B (Methodological)* **34** (1972), 187 – 220, DOI: 10.1111/j.2517-6161.1972.tb00899.x

[7] T.N.E. Greville, Short methods of constructing an abridged life tables, *Records of American Institute of Actuaries* **32** (1943), 29 – 43, DOI: 10.1007/978-3-642-81046-6_8.

[8] J.D. Kalbfleisch and R.L. Prentice, *The Statistical Analysis of Failure Time Data*, 2nd edition, Wiley (1980), DOI: 10.1002/9781118032985.

[9] N. Keyfitz, A life table that agrees with the data, *Journal of the American Statistical Association* **61**(314) (1966), 305 – 311, DOI: 10.1080/01621459.1966.10480866.

[10] G. King, On a short method of constructing an abridged mortality table, *Journal of Institute of Actuaries* **48**(3) (1914), 294 – 303, DOI: 10.1017/S0020268100026755.

[11] L. Martel, M. Provost, A. Lebel, S. Coulombe and A. Sherk, *Methods of Constructing Life Table for Canada, Provinces and Territories*, Report Demography Division, Catalogue number 84-538-X, Canada (2019), URL: https://www150.statcan.gc.ca/n1/pub/84-538-x/84-538-x2019001-eng.htm.

[12] R.G. Miller (Jr.), *Survival Analysis*, 2nd edition, John Wiley and Sons (2011), pages 238, URL: https://www.wiley.com/en-gb/Survival+Analysis,+2nd+Edition-p-9781118031063#series-section.

[13] T.M. Morgan and R.M. Elassof, Effects of censoring on adjusting for covariates in comparison of survival times, *Communications in Statistics – Theory and Methods* **15**(6) (1986), 1837 – 1854, DOI: 10.1080/03610928608829221.

[14] J.L. Reed and M. Merrel, A short method of constructing an abridged life table, in: *Mathematical Demography. Biomathematics*, Vol. 6. Springer, Berlin — Heidelberg (1977), DOI: 10.1007/978-3-642-81046-6_7.

[15] E. Ron, Cancer risk from medical radiation, *Health Physics* **85**(1) (2003), 47 – 59, DOI: 10.1097/00004032-200307000-00011.

[16] E. Ron, Ionizing radiation and cancer risk: Evidence from epidemiology, *Radiation Research* **150**(5) (1998), S30-41, URL: https://pubmed.ncbi.nlm.nih.gov/9806607/.