Modelling of Cancer Patients: An Accelerated Model with Composite Distribution

Bandana Sharma

Associate Professor, Cotton University, Guwahati, Assam

Abstract A cancer is always referred to by the stage at the time of its diagnosis. The progression of the disease cancer can be specified on the basis of area of infection under the disease. The area of infection for any type of cancer can be detected by using different methods of impression by the Doctors. Once the body of a normal person is invaded by cancer, the biological process of that person gets interrupted and starts behaving in a very unnatural way i.e., the life of the cancer patient gets automatically accelerated and goes through a sequence of transitions viz., from a cancer free stage to a stage with accelerated cancer situation and then to a stage to be followed by death (Kay, 1986). The disease has a chronic or stable phase, which is followed by transition to an accelerated phase (Klein et. al, 1984). Modeling is an experimental approach to investigate the complex biological systems of cancer for better understanding of the disease (Renee et.al, 2016). Over the years, various stochastic models were developed for insights into biological transitions of cancer through cell mutations (Jonathan et. al, 2014). With the accelerated behavior of the life of the cancer patient, in this paper an accelerated model of cancer patients by using Composite distribution under Power Rule Model has been developed along with prediction of the various survival rates of cancer patients as compared to normal persons.

Keywords Accelerated behavior of cancer patients, Stochastic process, Composite distribution, Power rule model

1. Introduction

Cancer is the leading cause of death in the economically developed countries and the second leading cause of death in the developing countries. According to WHO's health report, the estimated death worldwide due to cancer was 8.3 million in 2010 and in 2015 it was almost 9 million. In 2015, out of the total deaths due to cancer, 5.1 million were men and 3.9 million were female.

A cancer is always referred to by the stage at the time of diagnosis. It has been observed that survival of cancer patients mainly depends on the area of infection. In order to develop an effective cure and control of cancer, basic understanding of the mechanisms leading to the cancerous cells from normal cells is essential and highly desirable. Due to some unknown reasons, the cells of a particular part of the body started multiplying in a disorderly and uncontrolled manner. The cancerous cells grow beyond their usual boundaries and then started spreading through various pathways like local infiltration in the neighborhood of the organ of origin, lymphatic system or the blood vessels, leading to secondary deposits or metastases.

Classification of clinical extent of cancer is of great importance in the planning of treatment and for better understanding of the disease. Cancers can be classified either in accordance to the affected types of cells and organs of the body or in terms of spreading of cancerous cells to different organs of the body viz.,

(a) **Classification of cancers in accordance to the affected types of cells and organs:**

(i) **Carcinoma** is the most common type of cancer arises in the epithelium i.e. in the layers of cells covering the body's surface or the lining of internal organs and various glands.

(ii) **Sarcoma** originates in the supporting or connective tissues like bones, muscles and blood vessels.

(iii) **Myeloma** starts in the plasma of bone marrow.

(iv) **Leukemia** starts in the cells of lymph system i.e. the body's circulatory network for filtering body's impurities.

(v) Lymphoma originates at the lymphatic system.

(vi) Mixed types are of combination of two or more components of different origins of cancer.

(b) **Classification of cancers in terms of spreading of cancerous cells to different organs:**

This type of classification of cancers can be well understood from figure 1 (National Cancer Control Programmes, 2002):
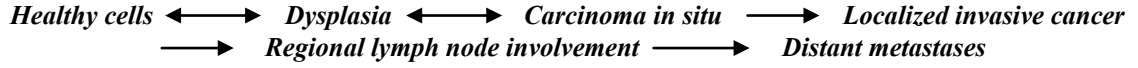
Healthy cells ⟷ *Dysplasia* ⟷ *Carcinoma in situ* ⟶ *Localized invasive cancer* ⟶ *Regional lymph node involvement* ⟶ *Distant metastases*

Figure 1.   Different phases of cancer development

Here dysplasia is the first indication of abnormality in the character of human cells. Carcinoma in situ is used to refer to the changes in cell nuclei but with no penetration of the underlying membrane that holds them in the tissue of origin. Localized invasive cancer is the abnormal cell growth involving areas underlying the tissue of origin. Regional lymph node involvement is the extension of cancerous cells to the regional lymph node that drains the area. Distant metastases are due to the spread of cancerous cells through the blood vessels or lymphatic system that affect other organs rather than the organ of origin (National Cancer Control Programmes, 2002).

The time period for the development of a clinically detectable cancer from a tumour cell is very short relative to the time for transformation process (Moolgavakar et. al, 1981). It has been observed that complete remission is possible only when the patient is diagnosed in the early stage of cancer i.e. in any of the stages viz., dysplasia or carcinoma in situ.

2. Model

By keeping the cancer patient under constant observation in the interval $(0,t]$, $t \geq 0$ and on basis of the transition graph in Fig 2, a stochastic process $\{X(t), t \geq 0\}$ with state space S = { 0, 1, 2, 3, 4, 5 } can be defined (Biswas, 1993; Medhi, 1998) with

X (t) = 0, if the patient is in the disease stage known as healthy cells

= 1, if the patient is in the disease stage known as dysplasia

= 2, if the patient is in the disease stage known as carcinoma in situ.

= 3, if the patient is in the disease stage known as localized invasive cancer.

= 4, if the patient is in the disease stage known as regional lymph node involvement.

= 5, if the patient is in the disease stage known as distant metastases.

Here, state 5 is the absorbing stage as there is no return for a cancer patient from this stage.

The corresponding transition matrix is given by

$$
R = \begin{array}{c} \\ H \\ D \\ C \\ L \\ R \\ DM \end{array}
\begin{pmatrix}
1-\lambda_h & \lambda_h & 0 & 0 & 0 & 0 \\
\mu_d & 1-(\lambda_d + \mu_d) & \lambda_d & 0 & 0 & 0 \\
0 & \mu_c & 1-(\lambda_c + \mu_c) & \lambda_c & 0 & 0 \\
0 & 0 & 0 & 1-\lambda_l & \lambda_l & 0 \\
0 & 0 & 0 & 0 & 1-\lambda_r & \lambda_r \\
0 & 0 & 0 & 0 & 0 & 1
\end{pmatrix}
\begin{array}{c} H\ \ D\ \ C\ \ L\ \ R\ \ DM \end{array}
$$

where λ_h, λ_d, λ_c, λ_l and λ_r are the failure rates of transitions from states H, D, C, L and R respectively and μ_d, μ_c are the cure rates at states D and C respectively.

For the sake of simplicity, the study has been restricted to the analysis of cancer patients in the stages viz., carcinoma in situ, localized invasive cancer, regional lymph node involvement and distant metastases only.

3. Assumptions

Following are the assumptions that are made for the proposed model

(1) Patients with unconfirmed cases of cancer are to be discarded.

(2) Patients belonging to the adult group (Barlow and Proschan, 1975) are considered to have lifetime with constant

hazard rate λ.

(3) The life of a normal person gets accelerated with the detection of cancer under Power rule model (Mann et al., 1973).

(4) The ailments score other than cancer at the time of invasive cancer detection is r_i ($i = 2, 3, 4, 5$). This r_i varies from patient to patient and $r_i \geq 0$.

(5) Severity of the disease is to be determined by the area of infection v_i in state i

$$(i = 2, 3, 4, 5) \text{ and } v_i > 0$$

Now, in accordance with the assumption (3), the hazard rate for accelerated life of a cancer patient becomes

$$\lambda_i = k \, v_i^{r_i} \quad r_i > 0, \text{ k being a constant and } (i = 2, 3, 4, 5) \text{ and } k > 0.$$

(6) In state 2, hazard rate λ_2 remains proportional to the area of infection i.e.,

$$\lambda_2 \propto v_2^{r_2} \Rightarrow \lambda_2 = k \, v_2^{r_2}$$

(7) In state 3, the area of infection becomes proportional to time i.e., $v_3 \propto t$

$$\Rightarrow v_3 = \alpha t \Rightarrow \lambda_3 = k \, (\alpha t)^{r_3}$$

(8) In state 4, the area of infection further becomes proportional to time i.e.,

$$v_4 \propto t^2 \Rightarrow v_4 = \beta \, t^2 \Rightarrow \lambda_4 = k \, (\beta t^2)^{r_4}$$

(9) The proportionality of infection to time is assumed to be k, which is a constant for all the states.

Hence under the above mentioned assumptions, $\{X(t), t \geq 0\}$ will be a stochastic process with state space $S = \{2, 3, 4, 5\}$ where the transition probabilities of the process are of the types $i \rightarrow i + 1$ ($i = 2, 3, 4$). The corresponding transition matrix $P = (p_{ij})_{i \times j}$ for $i = 2,3,4,5$ and $j = 2,3,4,5$ is given by

$$P = \begin{array}{c} \\ 2 \\ 3 \\ 4 \\ 5 \end{array} \begin{pmatrix} 1 - k \, v_2^{r_2} & k \, v_2^{r_2} & 0 & 0 \\ 0 & 1 - k \, (\alpha t)^{r_3} & k \, (\alpha t)^{r_3} & 0 \\ 0 & 0 & 1 - k(\beta t^2)^{r_4} & k(\beta t^2)^{r_4} \\ 0 & 0 & 0 & 1 \end{pmatrix} \begin{array}{c} 2 \quad 3 \quad 4 \quad 5 \end{array}$$

4. Probabilistic Analysis

The lifetime distribution of a cancer patient is being denoted by $g(t)$ as

$$g(t) = g_2(t), \quad 0 \leq t \leq d_1$$
$$= g_3(t), \quad d_1 \leq t \leq d_2$$
$$= g_4(t), \quad t \geq d_2$$

provided $\int\limits_0^\infty g(t)dt = \int\limits_0^{d_1} g_2(t)dt + \int\limits_{d_1}^{d_2} g_3(t)dt + \int\limits_{d_2}^\infty g_4(t)dt = 1$

where

$g_i(t)$ = p.d.f of a cancer patient to transits from state i to $i + 1$ ($i = 2,3,4$), $t \geq 0$ so that

$$g_2(t) = k \, v_2^{r_2} \, \exp(-k \, v_2^{r_2} \, t), \qquad 0 \leq t \leq d_1$$

$$g_3(t) = k \, (\alpha \, t)^{r_3} \, \exp\{-k \, (\alpha \, t)^{r_3} \, t\}, \qquad d_1 \leq t \leq d_2, \; \alpha > 0$$

$$g_4(t) = k\left(\beta\ t^2\right)^{r_4} \exp\{-k\left(\beta\ t^2\right)^{r_4}\ t\},\quad t \geq d_2,\ \beta > 0$$

Here the time interval $(o, t]$ is divided into three parts so that we can identify two-partition parameters d_1 and d_2 such that the patient remains in stage 2 for $0 \leq t \leq d_1$, in stage 3 for $d_1 \leq t \leq d_2$ and in stage 4 for $t \geq d_2$. The role of partition parameters seems to be of great importance to cancer patients because a patient transits from one state to another state depending upon the area of infection and the patient's physical capabilities to fight the disease. Here, partition parameters are function of parameters of the p.d.f of the lifetime distribution of the cancer patients. Now $g(t)$ may be defined as a composite distribution (Mann et al., 1973) viz.,

Here

$\overline{G}_i(t) =$ Survival function at state $i(i = 2,3,4)$ corresponding to the density $g_i(t)$, $t \geq 0$.

With the hazard rates of the transitions from i to $i+1$ $(i = 2,3,4)$ as

$$h_2(t) = k\ v_2^{r_2},\quad 0 \leq t \leq d_1$$

$$h_3(t) = k(\alpha\ t)^{r_3},\quad d_1 \leq t \leq d_2$$

$$h_4(t) = k\left(\beta\ t^2\right)^{r_4},\quad t \geq d_2$$

Case 1: When t = d_1 then $\quad d_1 = \dfrac{v_2^{r_1/r_3}}{\alpha},\ \alpha > 0$

Case 2: When t = d_2 then $\quad d_2 = \left(\dfrac{\alpha^{r_3}}{\beta^{r_4}}\right)^{1/(2r_4 - r_3)},\ \beta > 0$

Let *m* be the mean lifetime of a cancer patient such that

$m = m_2 + m_3 + m_4$, where $m_i(i = 2,3,4)$ is the mean stay time in state i $(i = 2,3,4)$. Now m_2, m_3 and m_4 are calculated as follows:

When $0 \leq t \leq d_1$, we have $m_2 = \{1 - \exp(-k\ v_2^{r_2} d_1)\}\left(d_1 + \dfrac{1}{k\ v_2^{r_2}}\right)$

When $d_1 \leq t \leq d_2$, we have

$$m_3 = \frac{\dfrac{1}{(r_3 + 1)}\Gamma\left(\dfrac{r_3 + 2}{r_3 + 1}\right)}{k^{1/r_3+1}\alpha^{r_3/r_3+1}}\left[I\left(k\alpha^{r_3} d_2^{r_3+1}, \frac{1}{(r_3+1)}\right) - I\left(k\alpha^{r_3} d_1^{r_3+1}, \frac{1}{(r_3+1)}\right)\right]$$

where $I_\lambda(k+1) = \dfrac{1}{\Gamma(k+1)}\displaystyle\int_0^\lambda e^{-\xi}\xi^k\,d\xi$, is the Incomplete Gamma function.

When $t \geq d_2$, We have $m_4 = \exp(-k\ \beta^{r_4} d_2^{2r_4+1})\ d_2(1 + d_2^{4r_4+1})$

So that mean life time in one state becomes

$$m = \{1 - \exp(-k\ v_2^{r_2} d_1)\}\left(d_1 + \frac{1}{k\ v_2^{r_2}}\right) + \exp(-k\ \beta^{r_4} d_2^{2r_4+1}) + d_2(1 + d_2^{4r_4+1}) +$$

$$\frac{\dfrac{1}{(r_3 + 1)}\Gamma\left(\dfrac{r_3 + 2}{r_3 + 1}\right)}{k^{1/r_3+1}\alpha^{r_3/r_3+1}}\left[I\left(k\alpha^{r_3} d_2^{r_3+1}, \frac{1}{(r_3+1)}\right) - I(k\alpha^{r_3} d_1^{r_3+1}, \frac{1}{(r_3+1)})\right]$$

5. Sampling Scheme

Under the sampling scheme to estimate the characteristics viz., average lifetime, survival probabilities etc. the following assumptions are made:

(i) N numbers of patients under incurable phase are taken for study. The patients are medically examined and their stages of cancer are determined.

(ii)   Once a patient transits from state i to $i+1$ $(i=2,3,4)$, observation on the patient is terminated and his future development regarding the disease cancer is ignored. The time epoch at which the patient reaches the state $i+1$ from state i is recorded.

(iii)   The experiment is continued till the transitions of all the N patients to the next states are recorded.

(iv) N_i, the number of patients under state i $(i=2,3,4)$ and t_{ij}, the time point at which the jth patient transits to state $i+1$ from state i $(i=2,3,4)$. The corresponding three random samples of sizes N_2, N_3, N_4 are $\left(t_{21},t_{22},----t_{2N_2}\right)$, $\left(t_{21},t_{22},----t_{2N_2}\right)$ and $\left(t_{21},t_{22},----t_{2N_2}\right)$ respectively.

The time epochs t_{ij}'s are random variables from $g_i(t)$ for $i=2,3,4$; $j=1,2,----,N_i$

By using the method of maximum likelihood the values of the parameters were calculated separately for states 2, 3 and 4 as

$$\hat{k}=\dfrac{N_2+N_3+N_4}{v_2^{\hat{r}_2}\sum_{i=1}^{N_2}t_{2i}+\alpha^{\hat{r}_3}\sum_{j=1}^{N_3}t_{3j}^{\hat{r}_3+1}+\beta^{\hat{r}_4}\sum_{l=1}^{N_4}t_{4l}^{2\hat{r}_4+1}} \quad \hat{\alpha}=\left[\dfrac{N_3}{\hat{k}\sum\limits_{j=1}^{N_3}t_{3j}^{\hat{r}_3+1}}\right]^{\frac{1}{\hat{r}_3}} \quad \hat{\beta}=\left[\dfrac{N_4}{\hat{k}\sum\limits_{l=1}^{N_4}t_{4l}^{2\hat{r}_4+1}}\right]^{\frac{1}{\hat{r}_4}} \quad \hat{r}_2=\dfrac{\log\dfrac{N_2}{\hat{k}\sum\limits_{i=1}^{N_2}t_{2i}}}{\log v_2}$$

No close form expressions are available for $\hat{r}_3$ and $\hat{r}_4$ but can be obtained by solving the above expressions explicitly and numerically.

Table 1.   Estimated values of r_2, r_3 and r_4 from simulated data

| Sl No.	Sample size	Estimated value of		
		r_2	r_3	r_4
1	30	2.0153609	2.8892478	3.5591778
2	31	2.0087920	2.9200161	3.5396237
3	32	1.9998897	2.9498194	3.5214944
4	33	2.0076529	2.5349118	3.5046419
5	34	2.0001454	2.5627915	3.5651138
6	35	1.9929007	2.5898771	3.5479549
7	36	1.9875862	2.6162129	3.5319087
8	37	1.9859716	2.6418386	3.5860651
9	38	2.0093835	2.6667917	3.5698028
10	39	2.0085610	2.6911071	3.5545158
11	40	2.0172997	2.7148176	3.5339474
12	41	2.0104556	2.7379483	3.5185046
13	42	2.0108141	2.7605305	3.5039284
14	43	2.0108800	2.7934827	3.5555198
15	44	2.0213873	2.8252310	3.5408320
16	45	2.0163059	2.8463108	3.5269164
17	46	2.0232252	2.8765057	3.513712
18	47	2.0192706	2.9057151	3.5011675
19	48	2.0262911	2.9254797	3.4892330
20	49	2.0195142	2.9533787	3.4778640
21	50	2.0251007	2.9804529	3.4670225
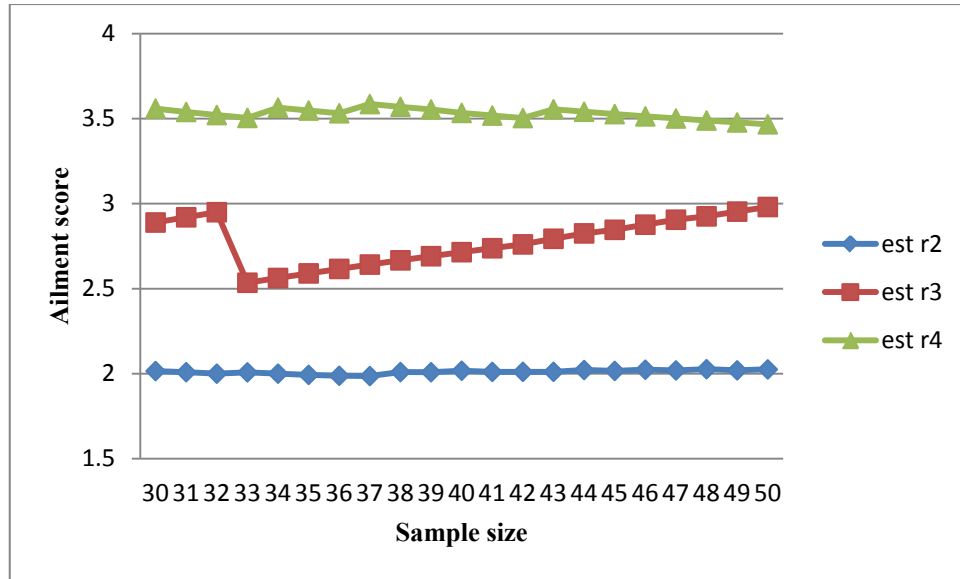
Figure 1.   Graph on ailment scores at different states of cancer

It can be observed that all the M.L.Es of the unknown parameters r_i $(i = 2,3,4)$, k, α, β are consistent estimators and are also asymptotically normal (Mann et al., 1973). The estimated values are being verified by using simulated data as only truncated cancer data are available. It has been observed with simulated data as shown in table 1 and figure 1, that the estimated ailments score other than cancer increases as the patient goes to a higher state from a lower state:

Again it can be observed with simulated data as shown in table 2 and figure 2, that from the estimated values of the partition parameters d_1 and d_2 it is clear that the patient remains in stage 2 for a very short period of time whereas it may remain in stage 3 for some time with proper treatment and care .

Table 2.   Estimated values of d_1, d_2 from simulated data

| Sl No.	Sample size	Estimated value of	
		d_1	d_2
1	30	0.0661292	2.8892478
2	31	0.0683206	2.9200161
3	32	0.070612	2.9498194
4	33	0.0498702	2.5349118
5	34	0.0518191	2.5627915
6	35	0.0537396	2.5898771
7	36	0.0555456	2.6162129
8	37	0.0571387	2.6418386
9	38	0.0573930	2.6667917
10	39	0.0588687	2.6911071
11	40	0.0598089	2.7148176
12	41	0.0615504	2.7379483
13	42	0.0628823	2.7605305
14	43	0.0641454	2.7934827
15	44	2.0213873	2.8252310
16	45	2.0163059	2.8463108
17	46	2.0232252	2.8765057
18	47	2.0192706	2.9057151
19	48	2.0262911	2.9254797
20	49	2.0195142	2.9533787
21	50	2.0251007	2.9804529
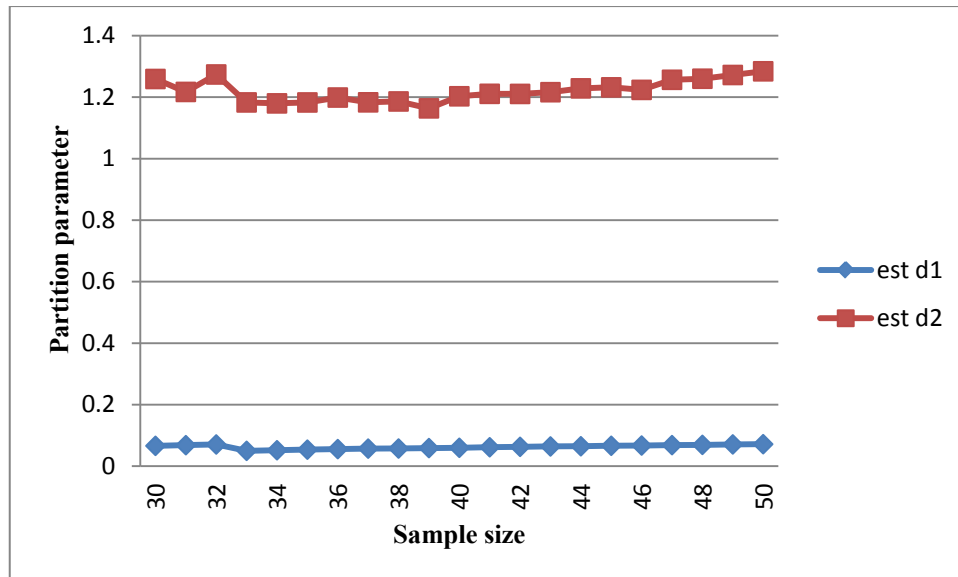
Figure 2. Graph on partition parameters

6. Discussion

While explaining the lifetime of a cancer patient, it may be observed that lifetime distribution of the patient does not remain same in the whole period of the disease. The distribution functions of life times with respect to the area of infection of a cancer patient at different stages of cancer are different and hence it becomes quite difficult to estimate the existing parameters viz., reliability, mean lifetime etc. Recently there has been a great deal of studies on the analysis of survival data on cancer patients. In this paper effort was made to have an insight to the accelerated behavior of cancer patients through transitions from one state to another state of cancer by means of partition parameters. All the parametric constants used in the model were estimated and mean life time of cancer patients was calculated. Through this model it has been observed that estimated ailments scores increases as the patient goes to higher state from the lower state.

7. Conclusions

It has been already mentioned that life of a normal person gets quite complicated with the invasion of cancer in the body. For this complication it becomes difficult to estimate the different parameters like average lifetime, survival probabilities. Nature of Composite distribution reduces this complexity to some extent through partition parameters. The estimators obtained in the paper thus may be used to obtain the mean lifetime and survival probabilities of cancer patients at different stages. The estimates of the partition parameters d_1 and d_2 will give the transition time of a cancer patient from stage 2 to stage 3 and from stage 3 to stage 4 respectively which may help the doctors to continue treatment procedure effectively to prolong the survival rates of cancer patients.

With growing techniques the survival rate of cancer has increased and so it becomes important to contribute even the smallest help in this field favoring the survival rate. For cancer patients survival rates are considered to be the main criteria to measure the impact of treatment on cancer control.

REFERENCES

[1]   Barlow, R.E and Proschan, F. (1975). Statistical Theory of Reliability and Life Testing, Holt, Rinehart and Winston, New York.

[2]   Biswas, S. (1988). Stochastic Processes in Demography and applications, New Age International Publishers Ltd.

[3]   Biswas, S. (1993). Applied Stochastic Process: Biostatistical and Population Oriented Approach, New Age International Publishers Ltd.

[4]   Bose S.M. (1992). Cancer, National Book Trust of India, New Delhi.

[5]   Devita, Jr. Vint, Hellman, S., Roseberg, S.A. (1982). Cancer, Principles and Practice of Oncology, 2nd edition, J.B Lippincott Company, Philadephia.

[6]   Frydman, H. (1995). Semi parametric Estimation in a Three state duration dependent Markov model from interval censored observations with applications to AIDS data, Biometrics 51, 502-511.

[7]   Gray, R.J. (1994). A kernel method for incorporating information on disease progression in analysis of survival, Biometrika 81, 527-539.

[8]   Kay, R. (1986). A Markov Model For Analyzing Cancer Markers and Disease States in Survival Studies, Biometrics 42, 855-865.

[9]   Klein, J.P., Klotz, J.H. & Grever, M.R. (1984). A Biological Marker Model for Predicting Disease Transitions, Biometrics 40, 927-936.

[10] Mann, N.R., Schhafer, R., and Singpurwalla, N.D. (1974). Methods of Statistical Analysis of Reliability and life Data, John Wiley & Sons.

[11] Medhi, J. (1998). Stochastic Processes, Wiley Eastern Limited.

[12] Mukhopadhyay, P. (2000). Mathematical Statistics, Books and Allied (P) Limited.

[13] National Cancer Control Programmes, Policies and Managerial guidelines. (2002). 2nd edition, World Health Organization, Geneva.

[14] Renee M. Thomas, Terry Van Dyke, Glenn Merlino and Chi-Ping Day (2016). Concepts in Cancer Modeling: A Brief History, Cancer research, American Association for Cancer Research.

[15] Singpurwalla, N.D., Al-Khayyal (1997). Accelerated Life Tests using the Power Law Model for the Weibull Distribution in 'The Theory and Applications of Reliability', Vol II, edited by C.P. Tsokos & I.N. Shimi, Academic Press.

[16] Jonathan W. Armond, Krishanu Saha, Anas A. Rana, Chris J. Oates, Rudolf Jaenisch, Mario Nicodemi & Sach Mukherjee (2014). A stochastic model dissects cell states in biological transition processes, Scientific Reports, Nature Publishing Group, volume 4, Article number: 3692.