

A STOCHASTIC MODEL FOR LUNG CANCER

P. Vantha¹, P. Pandiyan, and V.S. Venkatesh Kumar

ABSTRACT. Cancer is one of the most common health problems in the world today. A rapidly growing concern for millions of people around the world is the overall risk of cancer. Cancer is emerging as a major problem globally, both in more developed and less developed countries. The study showed that a person predisposed to lung cancer increases tumor growth, with exposure shortening the expected lifespan. The present study estimates that tumor development for carcinoma rapidly increases the threshold level of the infected individual. Once a person is infected, the tumor cells proliferate, and once infected the lungs will probably be affected. The length of the time interval for infection depends on the amount of smoke exposure the infected person has had. We have provided stochastic models for adenocarcinoma disease onset and progression, including the partial observational capability of tumourient's prostate cancer health status.

1. INTRODUCTION

Cancer is one of the most common health problems in this world. A rapidly growing concern for millions of people throughout the world is the widespread threat of cancer. Cancer is up-and-coming as a most important problem globally; in collaboration in further developed and in less residential countries. Annually,

¹corresponding author

2020 *Mathematics Subject Classification.* 00A71, 92-10.

Key words and phrases. Cigarette Smoking, Lung cancer, Tumor growth, Expected life time.

Submitted: 13.08.2022; *Accepted:* 30.08.2022; *Published:* 26.10.2022.

there are over 10 million new cases of cancer and more than 6 million deaths occur due to cancer worldwide. The contribution of the developing world to this figure is more than half. Cancer is considered the most progressive and devastating disease posing a threat of mortality to the entire world despite the significant advances in medical technology for its diagnosis and treatment. By 2020, the number of new cancer cases is expected to reach at least 15 million a year and cancer deaths to 10 million a year. In a wide range of cancer, some portion of the body's cells, start to separate without halting and spread into encompassing tissues. The Human body is made up of cells that are trillions in number. Tumors can begin at any place within these cells. Routinely, the human cells create and separate to shape new cells as expected by the body. At the point when cells are old, new cells typically substitute them. Be that as it may, when malignant growth happens, this interaction doesn't occur as it expected to be. The old cells do not die and the new cells are formed without necessity. The cells keep on dividing without any restrictions and form outgrowths in the body called tumors. These growths are typically strong and are solid masses of tissue. Malignancies of the blood, for example, leukemia, by and large, do not shape strong tumors. The cancerous tumors can spread into, or attack the tissues close to them, so these are called malignant. Moreover, as these cancers grow, some development cells can cut off and move to far-off spots in the body either through the blood or the lymph hub system and shape new cancer-causing growths quite far from the first growth area. However, harmless growths dislike threatening cancers. They do not spread or attack the tissues surrounding them, or the tissues close to them. After removal either by surgery or by other treatment procedures, benign tumors do not grow back. This is not normal for threatening cancers, which once in a while recover after evacuation. Generally, benign tumors are not life-threatening, except for the benign tumors that occur in the brain. The Cerebrum's harmless growths can be hazardous and could be the reason for the death of an individual.

Lung cancer is the most common cancer worldwide contributing nearly 13% of the total number of new cases diagnosed in 2008. Breast cancer is the second most common cancer in women only with nearly 1.4 million new cases in 2008. Colorectal cancer is the third most common cancer with over 1.2 million new cases in 2008. Of the 10 million new cancer cases each year worldwide, 4.7 million are

in the more developed countries and nearly 5.5 million are in the less developed countries. Although the disease has often been regarded as a problem in the developed world, more than half of all cancers occur in developing countries. In developed countries, cancer is the second most common cause of death and epidemiological evidence points to the emergence of a similar trend in developing countries. Cancer is the leading cause of death worldwide, it accounted for 7.4 million deaths in 2004. About 30% of cancer deaths can be prevented if detected during the initial stages of its occurrence. In 2005, cancer killed approximately 826000, people in India of which 519000 were under the age of 70. Lungs cancer is among the five main types of cancer leading to overall cancer mortality contributing to about 1.3 million deaths per globally. Tobacco use is the single most important risk factor for cancer. The economic cost of treating four major tobacco related diseases such as cancer, cardiovascular diseases, respiratory diseases, and tuberculosis in India is as high as US 1.7 billion (WHO, 2002).

The prevalence of adult smoking in the united states was born from 42.4% in 1965 to 25.5% in 1990, and progress has been slow since the 1990s (26.5% in 1992, 24.7% in 1995, and 23.3% in 2000). This is often partly due to high rates of relapse following quit tries among smokers. This is often revealed by the actual fact that the prevalence of bringing to an end within the United States accumulated from 24.3% in 1965 to 49.6% in 1993 and then flattened to 48.8% in 2000. Surveys show that top smoking prevalence is a minimum of partly as a result of high rates of relapse among smokers World Health Organization attempt quitting. A major problem once learning addiction behavior is that participants generally build many quit attempts before they with success quit. Thus, for economical development, targeting, and analysis of interventions, it's necessary to differentiate transient cessation (temporarily smoking-free however relapse later) from permanent cessation (lifelong smoking-free) and establish the chance factors related to permanent cessation. Smoking may be an advanced behavior influenced by social, economic, environmental, behavioral, and physiological factors. Our objectives are to identify and quantify baseline factors related to the success of permanent smoking cessation and describe the complete stochastic nature of the smoking addiction pattern. Within the remainder of this section, we have a tendency to describe the dataset, covariates, and modeling strategy to attain these objectives.

Any component or device when exposed to shocks that cause damage to the device or system is likely to fail when the total accumulated damage exceeds a level called the threshold [7]. The pace of aggregation of harm decides the lifetime of the part or gadget.. One can see more detail the threshold to attain the expected time in 8, [9] and [10].

The mathematical model is developed in the areas of biology, medicine, and engineering. These models help the concerned people in the application of the same in real-life situations. The use of mathematical models in the study of tumor cell growth and also in many other studies related to other diseases, are very useful to arrive at a conclusion. In any research study, mathematical as well as stochastic models are developed taking into consideration the problems which occur in real-life problems.

2. ASSUMPTIONS OF THE MODEL

These assumptions are somewhat artificial but have been made to account for the lack of detailed real-world information on the one hand and to explain the procedure on the other.

- (i) Cigarette smoking is the only source of tumor development for lung cancer.
- (ii) The threshold of any individual is a random variable. If the total damage crosses a threshold level Y which itself is a random variable, the causes occur and a person is recognized as infected.
- (iii) The arrival times of successive habits, the order of damage, and the threshold are all mutually independent.

NOTATIONS

- X_i : a continuous random variable denoting the amount of contribution to the threshold due to the smoking and salty foods in the i^{th} habit, in other words the damage caused to the lung cancer growth in the i^{th} habit, (Shock), $1, 2, \dots, k$ and X_i 's are i.i.d.
- Y_1, Y_2 : Continuous random variable denoting the threshold levels for the two grades which follows Generalized Rayleigh distribution.
- U_i : a random variable denoting the inter-arrival times between habit with c.d.f. $F_i(\cdot)$, $i = 1, 2, \dots, k$

- $g(.)$: The probability density function of X_i .
- $g * (.)$: Laplace transform of $g(.)$
- $g_k(.)$: The k - fold convolution of $g(.)$ i.e., p.d.f. of $\sum_{j=1}^k X_i$
- $f(.)$: p.d.f. of random variables showing between consistent guidelines announcement with the corresponding c.d.f. $F(.)$.
- $F_k(.)$: k-fold convolution of $F(.)$.
- FEV1: Forced expiratory volume
- FVC: forced vital capacity
- $S(.)$: Survival function.
- $V_k(t)$: Probability of exactly k habits.
- $L(t)$: $1 - S(t)$.

3. MODEL DESCRIPTION

In this paper having the threshold which follows Generalized Rayleigh distribution is discussed with the shape parameter ($\alpha = n$) is been considered. The expected time and variance are obtained. The two-parameters generalized Rayleigh distribution is a particular member of the generalized Weibull distribution, originally proposed by Mudholkar and Srivastava [11].

Let Y be the random variable which has the c.d.f defined as

$$(3.1) \quad F(x; \alpha \lambda) = (1 - e^{-(\lambda x)})^\alpha.$$

Let Y be the random variable which has the p.d.f defined as

$$(3.2) \quad f(x; \alpha \lambda) = 2\alpha\lambda^2 x e^{-(\lambda x)^2} (1 - e^{-(\lambda x)^2})^{\alpha-1}, \quad x > 0.$$

The corresponding survival function is

$$(3.3) \quad S(x; \alpha \lambda) = 1 - (1 - e^{-(\lambda x)^2})^\alpha.$$

The Hazard function is

$$(3.4) \quad \overline{H}(x) = \sum_r^n (-1)^{r-1} \binom{n}{r} (e^{-(\lambda x)^2})^r.$$

There may be no practical way to inspect an individual item to determine its threshold y . In this case, the threshold must be a random variable. The shock

survival probability is given by

$$\begin{aligned}
 P\left(\sum_{i=1}^k X_i < Y\right) &= \int_0^\infty g_k(x) \bar{H}(x) dx \\
 (3.5) \qquad &= \int_0^\infty g_k(x) \left[\sum_{r=1}^n (-1)^{r+1} \binom{n}{r} \left(e^{-(\lambda x)^2}\right)^r \right] dx \\
 &= \sum_{r=1}^n (-1)^{r+1} \binom{n}{r} g_k^*(\lambda x)^2.
 \end{aligned}$$

Equation (3.5) denotes the K^{th} convolution. Therefore $S(t) = P[T > t]$ is the survival function which gives the probability that the cumulative will fail only after time t . Here

$S(t) = P(T > t)$ is the probability that the total damage survives beyond t ; and
 $S(t) = \sum_{k=0}^{\infty} P(\text{there are exactly } k \text{ habits in } (0, t]) * P(\text{the total cumulative } (0, t]);$

$$(3.6) \qquad S(t) = P[T > t] = \sum_{k=0}^{\infty} V_k(t) P(X_i < Y).$$

A renewal process is a counting process such that the time until the first event occurs has some distribution F , the time between the first and second event has, independently of the time of the first event, the same distribution F , and so on. When an event occurs, we say that a renewal has taken place. It may happen that successive shocks become increasingly effective in causing damage, even though they are independent. This means that $v_k(t)$ the distribution function of the k^{th} damage is decreasing in $k = 1, 2, \dots$ for each t . It is also known from the renewal process that,

$$\begin{aligned}
 &P(\text{exactly } k \text{ habit decisions in } (0, t]) \\
 &= F_k(t) - F_{k+1}(t) \text{ with } F_0(t) = 1 \\
 &= \sum_{k=0}^{\infty} \sum_{r=1}^n [F_k(t) - F_{k+1}(t)] \binom{n}{r} (-1)^{r+1} g_k^*(\lambda r)^2,
 \end{aligned}$$

$$S(t) = \sum_{r=1}^n \binom{n}{r} (-1)^{r+1} - \sum_{r=1}^n \binom{n}{r} (-1)^{r+1} (1 - g_k^*(r\lambda)^2) \sum_{k=0}^{\infty} [F_k(t)] [g^*(r\lambda)^2]^{k-1}.$$

$P(T < t) = L(t) =$ The distribution functions of life time (T),

$$(3.7) \quad L(t) = 1 - S(t) = 1 - \sum_{r=1}^n \binom{n}{r} (-1)^{r+1} \\ - \sum_{r=1}^n \binom{n}{r} (-1)^{r+1} (1 - g_k^*(\lambda r)^2) \sum_{k=1}^{\infty} [F_k(t)] [g^*(\lambda r)^2]^{k-1}.$$

Here, $[f^*(s)]^k$ is Laplace transform of $V_k(t)$ since the inter-arrival times are i.i.d. The above equation can be rewritten as,

$$L(t) = \sum_{r=1}^n \binom{n}{r} (-1)^{r+1} (1 - g^*(\lambda r)^2) \sum_{k=1}^{\infty} [F_k(t)] [g^*(\lambda r)^2]^{k-1},$$

$$E(T) = \frac{d}{ds} L^*(s)$$

given

$$S = 0$$

$$E(T^2) = \frac{d^2 L^*(s)}{ds^2},$$

given

$$S = 0.$$

From which the variance can be obtained.

Let the random variable U denoting inter arrival time which follows exponential with parameter c . Now $f^*(s) = \left(\frac{c}{c+s}\right)$ substituting in the below equation (3.8) we get,

$$(3.8) \quad l^*(s) = \sum_{r=1}^n \binom{n}{r} (-1)^{r+1} \frac{(1 - g^*(\lambda r)^2) f^*(s)}{(1 - g^*(\lambda r)^2) f^*(s)}$$

$$= \sum_{r=1}^n \binom{n}{r} (-1)^{r+1} \frac{c(1 - g^*(\lambda r)^2)}{(c + s - g^*(\lambda r)^2) c}.$$

The mean and variance of the time to threshold to cross the breakdown point is derived:

$$E(T) = \frac{d}{ds} L^*(s) \quad \text{given } s = 0$$

$$\begin{aligned}
E(T) &= \sum_{r=1}^n \binom{n}{r} (-1)^{r+1} \frac{1}{c(1 - g^*(\lambda r)^2)} \quad (\text{on simplification}) \\
&= \sum_{r=1}^n \binom{n}{r} (-1)^{r+1} \frac{c(1 - g^*(r\lambda)^2)}{c^2(1 - g^*(r\lambda)^2)^2} \\
E(T) &= \sum_{r=1}^n \binom{n}{r} (-1)^{r+1} \frac{1}{c(1 - g^*(r\lambda)^2)} \\
E(T^2) &= \frac{d^2 L^*(s)}{ds^2} \quad \text{given } S = 0 \\
&= \sum_{r=1}^n \binom{n}{r} (-1)^{r+1} \frac{2c(1 - g^*(r\lambda)^2)}{c^2(c + s - g^*(r\lambda)^2)^3} \quad \text{given } s = 0 \\
&= \sum_{r=1}^n \binom{n}{r} (-1)^{r+1} \frac{2c(1 - g^*(r\lambda)^2)}{c^2(c + s - g^*(r\lambda)^2)^3} \\
E(T^2) &= \sum_{r=1}^n \binom{n}{r} (-1)^{r+1} \frac{2}{c^2(1 - g^*(r\lambda)^2)^2}.
\end{aligned}$$

The inter-arrival time of the threshold follows exponential distribution. The Laplace transformation of the exponential is given by

$$\begin{aligned}
&\left(\frac{\mu}{\mu + \lambda}\right) g^*(.) \sim \exp(\mu), \quad g^*(\lambda r)^2 = \left[\frac{\mu}{\mu + (\lambda r)^2}\right] \\
(3.9) \quad E(T) &= \sum_{r=1}^n \binom{n}{r} (-1)^{r+1} \left(\frac{1}{c}\right) \left[\frac{\mu + (\lambda r)^2}{(\lambda r)^2}\right] \\
E(T^2) &= \sum_{r=1}^n \binom{n}{r} (-1)^{r+1} \left(\frac{2}{c^2}\right) \left[\frac{\mu + (\lambda r)^2}{(\lambda r)^2}\right] \\
V(T) &= E(T^2) - [E(T)]^2 \\
V(T) &= \sum_{r=1}^n \binom{n}{r} (-1)^{r+1} \left(\frac{2}{c^2}\right) \left[\frac{\mu + (\lambda r)^2}{(\lambda r)^2}\right]
\end{aligned}$$

$$- \left[\sum_{r=1}^n \binom{n}{r} (-1)^{r+1} \left(\frac{1}{c} \right) \left(\frac{\mu + (\lambda r)^2}{(\lambda r)^2} \right) \right]^2$$

Special cases: ($\alpha=1$)

The shape parameter of the generalized Rayleigh distribution α is kept fixed i.e. $\alpha = 1$. We obtained the following equation (3.11) and (3.12) as the expected time to $E(T)$ and variance $V(T)$.

$$(3.10) \quad E(T) = [(\mu + (\lambda)^2)/((\lambda)^2 c)] \quad \text{on simplification}$$

$$(3.11) \quad E(T) = \left[\frac{\mu + (\lambda)^2}{(\lambda)^2 c} \right] \quad \text{on simplification}$$

$$(3.12) \quad V(T) = \left[\frac{(\mu + (\lambda)^2)^2}{c^2 (\lambda)^4} \right] \quad \text{on simplification,}$$

where, $\mu = (\text{FEV1})$, $\lambda = (\text{FVC})$ and $C = \text{Time Interval}$.

TABLE 1. FEV1, FVC infected person's stage wise

C	$\mu = (\text{FEV1})$	$\lambda = (\text{FVC})$	E(T)	V(T)
1	84	88	1.01085	1.022
2	76	83	0.50552	0.256
3	72	80	0.33708	0.114
4	72	76	0.25312	0.064
5	61	72	0.20235	0.041
6	61	65	0.16907	0.029
7	58	55	0.14560	0.021
8	55	53	0.12745	0.016
9	50	50	0.11333	0.013
10	43	49	0.10179	0.010

4. CONCLUSION

This study showed that the person predisposed to lung cancer increases tumor growth, with exposure shortening the expected lifespan. This study hypothesizes

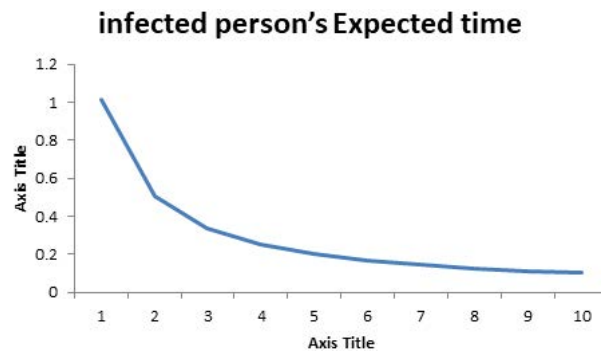


FIGURE 1. Infected person's Expected time

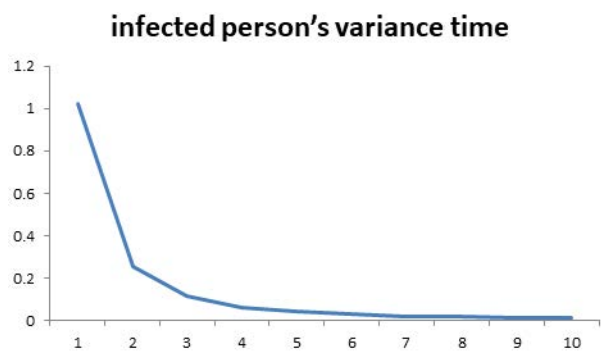


FIGURE 2. Infected person's variance time

that a person infected with tumor growth raises the threshold level very quickly for lung cancer. Once someone is infected, tumor cells multiply and once infected with the lungs, they will probably be affected. Life expectancy is the average range of years estimated for a person to measure with neoplasm cell infection. We recommend a technique for modelling participant-level random addiction behaviour. Cigarette smoking history includes self-reports of smoking (always versus other), real cigarettes per day (continuous and graded), age of smoking onset, years of recovery from smoking, and breathing while packing each year. It has been widely theorized that advanced or pathological processes increase the number of patients infected with FEV1, FVC-negative cells, thus damaging other smart cells. Oral cancer has increased in recent years, especially in developing countries. The threshold level of FEV1 - negative infected cells is calculated through the statistical model of the infected individual. Several standards of medical care support the demonstrated effects of various treatment modalities or methods. When

FVC negative infected cells are affected in the human body, shock appears with completely different infective variables. When the immune system does not accumulate the increase in shock, that is, the inter-arrival time, the expected life span of the human system can reach the edge. The total cumulative damage found with the shock model approach exploits the upgrade method. The expected lifetime is obtained through the distribution. The first data within the model was obtained for the expected time. In conclusion, a person infected with oral cancer FEV1, infected with FVC cells, is more likely to cross the threshold more quickly. Once the individual is infected, the growth of FEV1, FVC cells is damaged, with human tumor cells likely to affect both males and females. The time interval is for the infected person to drink. The expected lifetime is quickly reduced to the threshold level. The length of the time interval for infection depends on the duration of smoke exposure of the infected person. The model shows that once a person is infected, the immune system begins to break down, as seen in the above table and figures. We see that once a person is affected by cancer, there is a growth of good cells in the tumor and the FEV1, FVC cells are damaged by oral cancer, and his immune system's ability decreases. With proper medical advice and regular treatment, its life span can be extended. The findings serve as useful indicators for future studies. Many areas and segments in the medical field have to be updated using recent developments in statistical methods and stochastic processes. It can be seen that the principles of statistical models play an important role in the medical field. Furthermore, the important aspect of the study is to collect real-life data from different regions. The basis of which the appropriate probability distributions for random variables such as threshold and inter-arrival time can be ascertained using standard data distribution methods. The goodness of fit of such distributions can be verified using standard methods. This makes the model more application-oriented and applicable for all life situations.

REFERENCES

- [1] P. PANDIYAN, S. KOVENTHAN: *Estimating of survival time of cancer Patients*, International Journal of Recent Scientific Research, **9**(1) (2018), 23533–23534.
- [2] P. PANDIYAN: *Calculating Human Immune Virus Infected Person Survival Time Through Stochastic Model*, Bio- Science Research Bulletin, **29**(1) (3013), 1–4.

- [3] A. SARHAN, M. ZAINDIN: *Modified Weibull distribution*, Applied Sciences, bf11 (2009), 123–136.
- [4] P. PANDIYAN, S. KOVENTHAN, G. SUBASH CHANDRA BOSE: *Estimating The Survival Time Of cancer Drinker Patients Using Stochastic Model*, Indo-Asian Journal of Multidisciplinary Research, 4(4) (2018), 1464–1467.
- [5] S. JOTHIMANICKAM, P. PANDIYAN: *Calculating The Survival Time Of Cancer Patients Through Exponentiated Weibull Distribution*, Journal of International and Computational Science, 9(12) (2019), 159–165.
- [6] S. JOTHIMANICKAM, P. PANDIYAN, E. SUSIGANESHKUMAR, T. VIGNESH: (2020), *Stochastic Model for Expected Time Using Generalized Exponential Distribution*, Journal Of Xi'an University Of Architecture & Technology, 12(9) (2020), 335–341.
- [7] J.D. ESARY, A.W. MARSHAL, F. PROSCHAN: *Shock Model and Wear Processes*, Annals of Probability, 1(4) (1973), 627–649.
- [8] P. PANDIYAN, K. KANNADASAN. R. SWAMINATHAN, R. VINOTH: *Sthochastic model for mean and variance to recrut manpower in organization*, Advances in Applies Mathematical Analysis, 5(1) (2010), 77–82.
- [9] JEEVA ET.AL,: *Expected Time for Attainment Threshold Level A Shock Model Approach*, Stochastics and Quality Control, 19(1) (2004), 97–105.
- [10] R. SATHIYAMOORTHY: *Cumulative Damage model with Correlated Inter arrival Time of Shocks*, IEEE Transactions on Reliability, 29(3) (1980).
- [11] G.S. MUDHOLKAR, D.K. SRIVASTAVA: *Exponentiated Weibull Family for Analyzing Bathtub Failure-Rate Data*, IEEE Transactions on Reliability, 42 (1993), 299–302.

DEPARTMENT OF STATISTICS, ANNAMALAI UNIVERSITY, ANNAMALAI NAGAR, TAMIL NADU, INDIA.
 Email address: pvanitha2013@gmail.com

DEPARTMENT OF STATISTICS, ANNAMALAI UNIVERSITY, ANNAMALAI NAGAR, TAMIL NADU, INDIA.
 Email address: pandiyanau@gmail.com

DEPARTMENT OF COMMUNITY MEDICINE, GOVERNMENT VILLUPURAM MEDICAL COLLEGE AND HOSPITAL, VILLUPURAM, TAMIL NADU, INDIA.
 Email address: subashstat@gmail.com