International Journal of Science and Research (IJSR)
ISSN: 2319-7064
Impact Factor 2024: 7.101

An Analytical Study on Survival of Oral Cancer Patients of Assam by Using Life-Table and Semi-Partial Correlation

Jayanta Sarma Kakoty[1], Dr. Parbin Sultana[2]

[1]Research Scholar, Department of Mathematics, University of Science & Technology, Meghalaya, India
[2]Director, Research Division, University of Science & Technology, Meghalaya, India

Abstract: ***Oral cancer is a major public health issue in Assam, largely due to the prevalent use of tobacco, areca nut, smoking and alcohol. The purpose of the study is to analyse the survival of the cancer patients with the effect of the factors categorized as Personal (age, gender, behavioural factors, education, family income and the locality [rural or urban] and Clinical (cancer directed treatment [CDT], follow-up, topography sites, stage groups, treatment given prior to registration [TGPR], metabolic risks). The study is based on the sample of 1000 patients from the patient database (2015 – 2020) provided by North East Cancer Hospital & Research Institute (NECHRI), Guwahati, Assam, India. The study uses Life-Table method to examine the effect of several predictor variables (personal and clinical data of the patients) on the survival of the patient. The influence of the significant factors on the survival of the patient is interpreted by using semi-partial correlation. From the study, it is observed that behavioural factor, family income, CDT, follow up, topography sites, stage group, TGPR, MRF and locality (rural / urban) more significant on the survival rate of the patient. On the other hand, gender, educational level and food habits are not significant on the survival rate of the patient.***

Keywords: behavioural risk factors, clinical treatment outcomes, Life-Table, semi-partial correlation, socioeconomic influence, rural urban differences

1.Introduction

Cancer is a disease caused by an uncontrolled division of abnormal cells in the parts of the body. Cancer has become more than just a global health problem and is also a cause of deep suffering to the individual patients, their families and to the community at large. The GLOBOCAN 2022 estimates that there were 19.4 million new cases of cancer and almost 10 million deaths from cancer in 2021. According to ICMR-NCDIR latest report, oral cancer is among the top three cancers in India and it is among the top six cancers in Assam. The factors used in this study are categorized as 'Personal' and 'Clinical'. The personal factors with their variables used in this study are Age (X_1), Gender (X_2), Behavioural factors, Education (X_4), Family Income (X_5), Food habits (X_6) and Localities (Rural/Urban) (X_{13}). The clinical factors with their variables used in this study are CDT (Cancer Directed Treatment) (X_7), Follow up (X_8), Topography sites (X_9), Stage group (X_{10}), TGPR (Tratment given prior to registration) (X_{11}) and MRF (Metabolic risk factor) (X_{12}). The study is trying to analysis the survival rates of oral cancer patients by using the personal and clinical factors in context of Assam by applying Life-Table method. To study how the significant factors influence on the survival rate of the patient, semi-partial correlation is applied.

2.Review of Literatures

Different literatures on related topics have been reviewed and some of them are present as follows.

Moghimi-Dehkordi, B., et. al., (2023) conducted life-table method to analyse the survival rates and patterns of survival in gastric cancer based on patient's database of gastric cancer registered in the Cancer Registry Centre of Research Centre of Gastroenterology and Liver disease of Shaahid Beheshti University of Medical Science, Iran from Dec 21, 2001 to Dec 21, 2006. The study interpreted that stages of tumour, histology grade, histologic type of cancer, tumour size, age at diagnosis and surgery approach were independent prognostic factors. However, variables such as sex, body mass index, ethnicity and level of education did not show significant effects on Survival.

Abdus, R., et. al., (2012) estimated the survival rate of bladder cancer based on life table method. The data was based on cancer registry center of Shiraz University of Medical Sciences from 2001-2009. The findings showed that there was significant difference in survival rate of bladder cancer among the age groups and treatment types. There was no difference in survival time based on smoking, alcohol and educational level of patients.

Tong, Y., et. al., (2022) applied Kaplan – Meier method and Life Table method to understand the survival status. The study was based on 3263 cases of occupational pneumoconiosis from 1963 to 2020 in Shizuishan City. The results interpreted that there were significant differences in the survival rates of occupational pneumoconiosis patients among different types, diagnosis age, exposure time, industry, initial diagnosis stage and whether upgraded. As the survival time increased, the survival rate of patients decreased gradually.

Lee, S.H. (2018) studied the complex relationship between tumor vascularization and metabolism and eventually help in guiding targeted therapy**.** The results interpreted that A decrease in SD(WIS), MAD(TTHP), and MAD(IAUCtthp) and an increase in SD(TTP) were associated with a significant increase in SUVmax.

Wang, J-H. & Yang, C-T. (2022). utilized the non-parametric Kendall's partial correlation method to obtain

pure correlation to determine the importance of G-E interactions concerning clinical survival data under a marginal modeling framework. The results preferred Kendall's partial correlation method to identify G-E interactions related to the clinical survival results of patients with esophageal, pancreatic, and lung carcinomas using The Cancer Genome Atlas clinical survival genetic data, and further establish survival prediction models.

3.Methodology

The selected sample is based on the patients database provided by North East Cancer Hospital and Research Institute (NECHRI) during the period 2015 – 2020. The dependent variable Y used in this study is the status of the patient (death / alive). The factors which influence the status of the patients (response variable) are considered as predictor variables. The factors and their corresponding variables with their values are mentioned below.

Personal factors		**Clinical factors**	
Factors	**Categorial values**	**Factors**	**Categorial values**
Demographic factors Age in years Gender	 a positive integer 1, if female 2, if male	CDT	1, if Yes 2, if No 3, if Treatment assigned but not detected
Behavioural factors	1, if consumes nothing 2, if consumes Areca Nut 3, if consumes Tobacco 4, if Smoking 5, if consumes Alcohol 6, if consumes Areca nut and Tobacco 7, if consumes Areca nut and Smoking 8, if consumes Areca nut and Alcohol 9, if consumes Tobacco and Smoking 10, if consumes Tobacco and Alcohol 11, if consumes Alcohol and Smoking 12, if consumes Areca nut, Tobacco and Smoking 13, if consumes Areca nut, Tobacco and Alcohol 14, if consumes Areca nut, Smoking and Alcohol 15, if consumes Tobacco, Smoking and Alcohol 16, if consumes all	Follow up	a positive integer
		Topography sites	1, Cheek 2, Esophagus / Throat / Nose 3, Gingiva 4, Lip 5, Mouth 6, Palate 7, Tongue
		Stage groups	1, if unknown or not detected 2, if Stage 1 3, if Stage 2 4, if Stage 3 5, if Stage 4
Education	1, if under Metric (UM) 2, if 10^{th} passed to undergraduate (UG) 3, if graduate or higher (GH)	TGPR	1, if Yes 2, if No
Family income (in ₹ pa)	1, if below 1 lac 2, if 1 lac to 5 lacs 3, if 6 lacs and above	MRF	1, if None 2, if Over weight 3, if Raised BG 4, if Raised BP 5, if Over weight and Raised BG 6, if Over weight and Raised BP 7, if Raised BG and Raised BP 8, if Over weight, Raised BG and Raised BP
		Rural / Urban X_{13}	1, if Rural 2, if Urban
Food habits	1, if Veg (V) 2, if Non-veg (NV)	**Dependent (response) variable** Status of the patient	0, if Death 1, if Alive

Life table describes the mortality and survival pattern of the patients who suffered oral cancer in the region. A life table provides information on parameters such as the number of survivors, the number of deaths and the life expectancy.

The survival function usually denoted by S(t), is mathematically defined as

$$S(t) = P(T > t)$$

Where,

T is the variable representing the time until the event (i.e. death) occurs.
T is a specific point in time
To test the effect of survival rate, Wilcoxon (Gehan) test statistic is applied.

Wilcoxon (Gahen Test): Wilcoxon (Gehen) test is a non-parametric test used in survival analysis to compare survival curves when one group consistently have a higher risk than the other.

To analyse the influence of the significant factors on the survival of the patients, the **semi-partial correlation** between the dependent variable and an independent variable without controlling the other independent variables, is applied.

4.Analysis and Interpretations

The effective sample size taken for this study is 1000 patients all over the Assam. The questionnaire is designed from the proforma given by Indian Council of Medical Research (ICMR). Personal data of the patients are collected by taking interview of the patients or relatives over phone and personal interviews were also conducted who are accessible and the clinical data of the patients are collected from the hospital records.

By using the Life Table (by SPSS), the results are analysed and to test the significance of survival rates among the oral cancer patients in context of Assam, Wilcoxon (Gehan) test statistic is applied. Life table is constructed by using SPSS25 by considering ages of the patients ae Time, Status of the patients as death or alive and Gender, Behavioural Factor, Education, Family Income, Food Habits, CDT, Follow Up, Topography Sites, Stage Group, TGPR, MRF and Locality (Rural or Urban) as the primary factors. (Annexure5), the estimated values of the test statistic are given in Table 1.

Table 1: Test of significance of the primary factors on survival rate of the patient

Primary factor	Wilcoxon (Gehen) Statistic	df	Sig.
Gender	2.411	1	0.120
Behavioural Factor	**278.383**	**15**	**0.000**
Education	1.469	2	0.480
Family Income	**9.476**	**2**	**0.009**
Food Habits	1.204	1	0.273
CDT	**67.458**	**2**	**0.000**
Follow Up	**356.313**	**4**	**0.000**
Topography Sites	14.632	6	0.023
Stage Group	**467.338**	**4**	**0.000**
TGPR	**200.236**	**1**	**0.000**
MRF	**306.883**	**7**	**0.000**
Locality (Rural / Urban)	**62.570**	**1**	**0.000**

From Table 1, it may be interpreted that Behavioural Factor, Family Income, CDT (Cancer Directed Treatment), Follow Up, Stage Group, TGPR (Treatment Given Prior to Registration), Metabolic Risk Factor (MRF) and Locality (Rural / Urban) are significant on the survival of the patient as their p-values < 0.01. The factors Gender, Education, Food Habits and Topography Sites are not significant on the survival of the patient as their p-values > 0.01.

To analyse the influence of the significant factors mentioned in Table 1 on the survival of the patients, the semi-partial correlation between the status of the patient (death or alive) and the given predictors (significant factors) are computed by using SPSS and are given in Table 2.

Table 2: The semi-partial correlation between status of the patient and significant factor

Significant factors	Semi-partial correlation
Behavioural Factors	-0.073
Family Income	0.097
CDT	-0.035
Follow Up	0.150
Stage Group	-0.281
TGPR	-0.149
MRF	-0.168
Rural / Urban	-0.091
a Dependent Variable: Status	

The semi-partial correlation values for a given factor are computed under the influence of the other factors. From Table 2, it is observed that Behavioural Factor, CDT (Cancer Directed Treatment), Stage Group, TGPR (Treatment Given Prior to Registration), MRF (Metabolic Risk Factor) and Locality (Rural / Urban) are negatively correlated with the status of the patients. It implies that if these factors increase, the chance of survive of the patient decreases. On the other hand, Family Income and Follow Up are positively correlated with the status of the patients. It implies that if these factors increase, the chance of survive of the patient also increases. It is also observed that the negative semi-partial correlation between Stage Group and Status is maximum followed by MRF, TGPR and Locality (Rural / Urban). The positive semi-partial correlation between Follow Up and Status is more than Family Income. Thus, it may be interpreted that;

- The survival rate of the patient decreases with the increase of behavioural factor, i.e., the patient having habituated in consuming alcohol may less survive than the patient consuming smoking, tobacco and areca nut.
- The survival rate of the patient increases when the family income increases.
- The survival rate of the patient decreases with the increase of cancer directed treatment (CDT). That is, the patient who is assigned the cancer directed treatment but not detected and the patient who is not assigned cancer directed treatment, are less survival rate than the patient who is assigned cancer directed treatment.
- The survival rate of the patient increases with the increase of follow up time.
- The survival rate of the patient decreases when the stage group increases.
- The survival rate of the patient decreases with the increase of TGPR (Treatment Given Prior to Registration).
- The survival rate of the patient decreases with the increase of MRF (Metabolic Risk Factor).
- The survival rate of the patient decreases with the increase of Locality (Rural / Urban).

5.Conclusions

The study prevails the following conclusions;

- Gender, Education, Food habits and Topography sites are not significant on the survival of the patient.
- Behavioural factor is negatively significant on the survival of the patient. The survival rate is lower with the consumption of alcohol than that of smoking, tobacco and areca nut.
- Family income is positively significant on the survival of the patient. The survival rate of the patients are higher in higher income groups.
- CDT (cancer detected treatment) is negatively significant on the survival of the patient. the patient who is assigned the cancer directed treatment but not detected and the patient who is not assigned cancer directed treatment, are less survival rate than the patient who is assigned cancer directed treatment.
- Follow up is positively significant on the survival of the patient. The survival rates of the patients are higher in higher follow up times.
- Stage group is negatively significant on the survival of the patient. The survival rates of the patients are lower in higher stage groups.
- TGPR (Treatment given prior to registration) is negatively significant on the survival of the patient. The survival rates of the patients whose treatments are not given prior to registration are lower than that of the patients whose treatments are given prior to registration.
- MRF (metabolic risk factor) is negatively significant on the survival of the patients. The survival rates are lower in higher metabolic risk factors.
- Locality of the patient is negatively significant on his survival. The survival of the patients are higher in rural areas than that of in urban areas.

References

[1] **Datema, F. R., Moya, A., Krause, P., Bäck, T., Willmes, L., Langeveld, T., ... & Blom, H. M. (2012).** Novel head and neck cancer survival analysis approach: random survival forests versus Cox proportional hazards regression. *Head & neck*, *34*(1), 50-58. https://onlinelibrary.wiley.com/doi/abs/10.1002/hed.21698.

[2] **Farida, Y., Maulida, E.A., Desinaini, L.N., Utami, W.D., & Yuliati, D. (2021).** Breast Cancer survival Analysis Using Cox – Proportional Hazard Regression and Kaplan- Meier Method. JTAM (Vol. 5, No. 2). https://journal.ummat.ac.id/index.php/jtam/article/view/4653

[3] **Hu, Y. (2014).** Building Statistical Models for the Prediction of Oral cancer Recurrence. The University of British Colombia.

[4] https://www.stat.ubc.ca/building-statistical-models-prediction-oral-cancer-recurrence

[5] **Lee, S.H., Rimmer, A., Gelb, E., Deasy, J.O., Hunt, M.A., Humm, J.L., & Tyagi, N. (2018).** Correlation Between Tumor Metabolism and Semiquantitative Perfusion Magnetic Resonance Imaging Metrics in Non-Small Cell Lung Cancer. International Journal of Radiation Oncology*Biology*Physics. (Vol. 102, Issue 4). https://www.sciencedirect.com/science/article/abs/pii/S0360301618303304

[6] **Tong, Y., Kong, Y. Y., Bian, H., Zheng, J. Z., Wu, Y. J., & Zhang, Y. (2022).** Survival and disease burden trend analysis of occupational pneumoconiosis from 1963 to 2020 in Shizuishan City. *Zhonghua lao Dong wei Sheng zhi ye Bing za zhi= Zhonghua Laodong Weisheng Zhiyebing Zazhi= Chinese Journal of Industrial Hygiene and Occupational Diseases*, *40*(5), 341-347.

[7] **Wang, H. & Yang, C-T. (2022).** Identification of Gene-Environment Interactions by Non-Parametric Kendall's Partial Correlation with Application to TCGA Ultrahigh-Dimensional Survival Genomic Data. IMR Press. https://www.imrpress.com/journal/FBL/27/8/10.31083/j.fbl2708225