RESEARCH ARTICLE

A Model Approach to Calculate Cancer Prevalence from 5 Years Survival Data for Selected Cancer Sites in India – Part II

Ramnath Takiar[1]*, Sathish Kumar Krishnan[2], Varsha Premchandbhai Shah[3]

Abstract

Objective: Prevalence is a statistic of primary interest in public health. In the absence of good follow-up facilities, it is often difficult to assess the complete prevalence of cancer for a given registry area. An attempt is made to arrive at the complete prevalence including limited duration prevalence with respect of selected sites of cancer for India by fitting appropriate models to 1, 3 and 5 year cancer survival data available for selected registries of India. Methodology: Cancer survival data, available for the registries of Bhopal, Chennai, Karunagappally, and Mumbai was pooled to generate survival for the selected cancer sites. With the available data on survival for 1, 3 and 5 years, a model was fitted and the survival curve was extended beyond 5 years (up to 30 years) for each of the selected sites. This helped in generation of survival proportions by single year and thereby survival of cancer cases. With the help of estimated survived cases available year wise and the incidence, the prevalence figures were arrived for selected cancer sites and for selected periods. In our previous paper, we have dealt with the cancer sites of breast, cervix, ovary, lung, stomach and mouth (Takiar and Jayant, 2013). Results: The prevalence to incidence ratio (PI ratio) was calculated for 30 years duration for all the selected cancer sites using the model approach showing that from the knowledge of incidence and P/I ratio, the prevalence can be calculated. The validity of the approach was shown in our previous paper (Takiar and Jayant, 2013). The P/I ratios for the cancer sites of lip, tongue, oral cavity, hypopharynx, oesophagus, larynx, nhl, colon, prostate, lymphoid leukemia, myeloid leukemia were observed to be 10.26, 4.15, 5.89, 2.81, 1.87, 5.43, 5.48, 5.24, 4.61, 3.42 and 2.65, respectively. Conclusion: Cancer prevalence can be readily estimated with use of survival and incidence data.

Keywords: Survival - model - prevalence - oral cavity - larynx - NHL - colon - prostate - leukemia

Asian Pac J Cancer Prev, **15 (14)**, 5681-5684

Introduction

Prevalence is a statistic of primary interest in public health because it identifies the level of burden of disease or health-related events on the population and health care system. Prevalence represents new and pre-existing cases alive on a certain date. Prevalence is a function of both the incidence of the disease and survival.In the absence of good follow-up facilities for cancer in Indian registries, it is not possible to estimate the cancer prevalence either at registry level or at country level. In our previous paper (Takiar and Jayant, 2013), it was shown that by fitting an appropriate curve to limited cancer survival data of 1,3 and 5 years, it is possible to generate the survivals for 1-30 years. It was also shown there that for the selected cancer sites of breast, cervix, ovary, lung, stomach and mouth, the prevalence to incidence ratios tend to stabilize around 30 years. Further, for any selected cancer site, the stabilized P/I ratio can be used to find out the prevalence by multiplying it with the known incidence cases. In the current paper, the same approach was extended to 11 more cancer sites.

Objectives: 1) Using the available Cancer Survival data of 1 year, 3 years and 5 years for selected cancer sites, a best fit model to be identified for each site; 2) The best fit model equation is used to extend the cancer survival curve beyond 5 years say up to 30 years; 3) From the extended survival curve, generate the proportion of survivals by single years for 1-30 years duration; 4) The survival proportions as obtained above by single years in combination with incidence cases estimated for single years can be used in determining the prevalence by single years; 5) To arrive at the stabilized prevalence to incidence ratios of selected cancer sites using the model approach so that they can be used in determining the prevalence of the respective sites.

Materials and Methods

In Tumor Based Cancer, an attempt is made to find out the number of primary cancers diagnosed among individuals living on a specified date while in Person Based Cancer, an attempt is made to find out the number of individuals living with cancer on a specified date.

*[1]National Cancer Registry Programme, Bangalore, Ex-Scientist G, Bangalore, India, [2]National Cancer Registry programme, [3]National Cancer Registry programme (NCDIR), Statistician, Bangalore, India *For correspondence: ramnath_takiar@yahoo.co.in*

The registries with long duration of working (more than 30 years) and good follow-up facilities could only give the complete prevalence. The Complete Prevalence represents the proportion of people alive on a certain day that previously had a diagnosis of the disease, regardless of how long ago the diagnosis was, or if the patient is still under treatment or is considered cured. The registries with short duration of working say 15-20 years can at-most provide the information on Limited Duration Prevalence. It represents the proportion of people alive on a certain day that had a diagnosis of the disease, within the past 'X' years. Prevalence can be examined over various durations of time say 5 years, 10 years or 15 years.

In a Scientific Report of IARC (IARC, 2011), the registries of Bhopal, Chennai, Karunagappally and Mumbai provided their cancer survival rates for three selected periods (1 year, 3 years & 5 years) and for 16, 20, 22 and 28 sites, respectively. The following eleven sites: Lip (C00), Tongue (C01-02), Oral cavity (C03-06), Hypopharynx (C12-13), Oesophagus (C15), Larynx (C32), NHL (C82-85 & C96), Colon (C18), Prostate (C61), Lymphoid Leukemia (C91), Myeloid Leukemia (C92-94) provided with respective ICD10 codes are considered for the present communication. The survival data was pooled separately for each selected site by periods for all the registries.

Using the pooled-survival data of 1 year, 3 years and 5 years of selected cancer site, an attempt is made to fit an appropriate survival curve with the help of SPSS (SPSS, 2005). Then, with the help of the equation of survival curve, the survival proportions will be obtained for each single year up to 30 years duration or till such duration that the survival proportion tend to stabilize.

For calculation of prevalence, it is assumed that the cancer incidence rate for the registry is constant over the years and the incidence cases of the registry are 100,000 for the base year 2001. Further, the incidence cases are expected to grow proportionately to the population growth. Then, applying the survival proportions and the incidence estimates for the year 2001-2030, the prevalence figures are estimated. The details of such calculations for cancer of all sites can be seen in our previous paper (Takiar and Jayant, 2013). The prevalence to incidence ratio is also calculated for each single year, for each selected 11 cancer sites, for the period of 1-30 years. Then, such a ratio can serve as a multiplier to the incidence cases for the selected period to arrive at the point prevalence.

Results

The percentage survival of cases of selected cancer sites by different years-pooled for four registries is provided in Table 1. The 1 year survival was below 40% for the sites of Lymphoid Leukemia, Oesophagus and Myeloid Leukemia. Hypopharynx, NHL, Colon and Tongue cancer sites had the 1 year survival around 50%. Oral Cavity, Larynx, Prostate and Lip cancer sites had 1 year survival above 60%. The 5 years survival was below 20% for the cancer sites of Lymphoid Leukemia, Hypopharynx, Myeloid Leukemia and Oesophagus. The 5 years survival was between 20-30% for the cancer sites of Larynx, NHL, Colon, Prostate and Tongue. Oral Cavity and Lip cancer had the 5 years survival above 30%.

Table 1. Percentage Survival of Cases by Selected Cancer Sites by Years (Pooled for four Registries) and the Constants of Best Fitted Curve

Cancer Site	No. of cases	1 year	3 years	5 years	a	b	R^2 value
Lip	247	77.3	58.7	50.6	0.78	-0.26	0.999
Tongue	3,307	55.1	29.2	23.6	0.54	-0.53	0.995
Oral Cavity	4,726	61.5	38.3	31.7	0.61	-0.41	0.999
Hypopharynx	3,070	46.5	22.4	16.0	0.46	-0.66	1.000
Oesophagus	5,413	34.8	16.0	10.5	0.35	-0.74	0.999
Larynx	2,708	61.6	37.7	29.2	0.62	-0.46	0.999
NHL	3,305	50.2	35.7	27.8	0.51	-0.36	0.986
Colon	1,626	52.5	34.8	27.4	0.53	-0.40	0.997
Prostate	1,495	64.9	39.5	23.9	0.67	-0.59	0.956
Lymphoid Leukaemia	1,602	37.5	24.1	18.0	0.38	-0.45	0.992
Myeloid Leukaemia	1,831	33.7	21.2	13.8	0.35	-0.53	0.966

With the help of SPSS, a suitable curve is fitted to survival data of each site. The type of curve fitted and respective constants derived are shown in Table 1. In general, the power curve ($y=at^b$) was found to be fitting well to the survival data. The R^2 value for each curve fitted is also shown which in general was above 0.95 suggesting that the fits are good.

With the help of the best fit equation, the survival proportions are estimated for 1 year to 30 years and shown with 5 years interval in Table 2. The 10 years survival was between 10-20% for the cancer sites of Hypopharynx, Myeloid Leukemia, Lymphoid Leukemia, Tongue and Prostate. The 10 years survival was between 20-30% for the cancer sites of Colon, Larynx, NHL and Oral Cavity. Oesophagus (6.4%) and Lip (42.5%) cancer sites had the extreme 10 years survival values. The 30 years survival was below 10% for the cancer sites of Tongue, Hypopharynx, Oesophagus, Prostate, Lymphoid Leukemia and Myeloid Leukemia. It was between 10-20% for the cancer sites of Oral Cavity, Larynx, NHL and Colon. Lip had the highest 30 years survival of 31.9%. The absolute fall in survival was higher between 10 to 20 years. There after the absolute fall was not much and remained mostly between 1-3 units.

Following the procedure described earlier: Takiar and Jayant, (2013), the incidence and prevalence figures are generated for selected cancer sites with the help of survival curves.Then, prevalence to Incidence ratio is calculated for each single year but shown for years which are multiple of five (5,10,15,20,25,30,35) in Table 2. The 5 years P/I ratio was below 2 for the sites of Oesophagus (0.93), Myeloid Leukemia (1.07), Lymphoid Leukemia (1.25), Hypopharynx (1.30), Tongue (1.67), NHL (1.79), Colon (1.80) and Prostate (1.98). Larynx (2.01), Oral Cavity (2.06) and Lip (2.97) had P/I ratio above 2.

The 30 years P/I ratio was below 3 for the sites of Oesophagus (1.87), Myeloid Leukemia (2.65) and Hypopharynx (2.81). The P/I ratio was between 3-5 for the sites of Lymphoid Leukemia (3.42), Tongue (4.15) and Prostate (4.61). The remaining sites like Colon (5.24), Larynx (5.43), NHL(5.48), Oral Cavity (5.89) and Lip (10.26) had the P/I ratio above 5.

Table 2. Estimated Percentage Survival by 5 Years Interval (5-30 years) and Prevalence to Incidence Ratio Arrived Using the Model Approach

Cancer Site	Percentage survival by 5 years interval						Prevalence to incidence ratio by 5 years interval					
	5	10	15	20	25	30	5	10	15	20	25	30
Lip	50.9	42.5	38.2	35.4	33.4	31.9	2.97	4.99	6.63	8.01	9.21	10.26
Tongue	23.0	15.9	12.8	11.0	9.7	8.8	1.67	2.48	3.06	3.50	3.85	4.15
Oral Cavity	31.4	23.6	19.9	17.7	16.1	14.9	2.06	3.23	4.10	4.81	5.39	5.89
Hypopharynx	16.0	10.1	7.7	6.4	5.5	4.9	1.30	1.83	2.18	2.44	2.65	2.81
Oesophagus	10.7	6.4	4.7	3.8	3.3	2.8	0.93	1.27	1.49	1.65	1.77	1.87
Larynx	29.4	21.4	17.7	15.5	14.0	12.9	2.01	3.08	3.86	4.48	5.00	5.43
NHL	28.6	22.3	19.3	17.4	16.1	15.0	1.79	2.88	3.72	4.41	4.99	5.48
Colon	27.8	21.1	17.9	16.0	14.6	13.6	1.80	2.84	3.63	4.26	4.79	5.24
Prostate	25.9	17.2	13.5	11.4	10.0	8.9	1.98	2.87	3.48	3.94	4.31	4.61
Lymphoid Leukaemia	18.5	13.5	11.3	9.9	9.0	8.3	1.25	1.92	2.42	2.82	3.14	3.42
Myeloid Leukaemia	14.7	10.2	8.2	7.0	6.2	5.7	1.07	1.58	1.95	2.23	2.46	2.65

Table 3. Prevalence Cases Estimated for Selected Cancer Sites Using P/I Ratio - India - (2015-2025)

Cancer Site	Incidence cases by year*			P/I Ratio	Prevalence cases by year**		
	2015	2020	2025		2015	2020	2025
Lip	3,952	5,347	7,394	10.26	40,545	54,856	75,864
Tongue	48,888	62,099	78,991	4.15	202,886	257,713	327,812
Oral Cavity	70,228	90,342	116,980	5.89	413,642	532,117	689,011
Hypopharynx	16,213	16,196	16,264	2.81	45,558	45,512	45,701
Oesophagus	37,344	36,058	34,981	1.87	69,833	67,429	65,415
Larynx	27,271	29,012	30,898	5.43	148,079	157,533	167,778
NHL	36,057	38,408	40,958	5.48	197,594	210,477	224,448
Colon	31,816	40,601	51,954	5.24	166,714	212,752	272,241
Prostate	39,200	51,979	68,985	4.61	180,713	239,621	318,023
Lymphoid Leukaemia	17,324	19,607	22,247	3.42	59,249	67,056	76,086
Myeloid Leukaemia	25,117	34,884	50,663	2.65	66,560	92,441	134,256

*Incidence cases as per NCRP Trend Report (2013). Using the NCRP approach, the incidence cases for 2025 are estimated; **Prevalence is derived by multiplying the incidence cases by the P/I ratio

Considering the above P/I ratio for 30 years duration for selected cancer sites and with the knowledge of estimated cancer incidence cases for India for three time periods namely 2015, 2020 and 2025, the prevalent cancer cases are calculated for each selected site and period and shown in Table 3. In India, by the year 2015, the estimated prevalent cancer cases in the decreasing order will be of Oral Cavity (413,642), Tongue (202,286), NHL(197,594), Prostate (180,713), Colon (166,714), Larynx (148,079), Oesophagus (69,833), Myeloid Leukemia (66,560), Lymphoid Leukemia (59,249), Hypopharynx (45,558) and Lip (40,545). The expected number of prevalent cancer cases for above sites is also provided for the year 2020 and 2025.

Discussion

The studies related to selected cancer sites, in particular from India and neighboring countries can be found easily in literature. The prevalence of oral cancers is high in Asian countries, especially in South and South East Asia. Asian distinct cultural practices as betel-quid chewing and varying pattern of tobacco and alcohol use are important risk factors that predispose to cancer of oral cavity (Rao et al., 2014) Oral cavity cancer is the most prevalent type amongst the males and one of the highest across the globe (Mishra and Mehrotra, 2014). The incidence of tongue cancer has been shown to be increasing in Chennai PBCR area (Krishnamurthy and Ramshankar, 2013).

In year 2010, it was estimated that nearly 0.36 million new cases and 0.19 million deaths with Non-Hodgkin Lymphoma occurred in India. In the study, cigarette smoking (OR=2.0) and bidi smoking (OR=2.8) were associated with excess risk of lymphoma. Among dietary items, only the consumption of mutton showed 7.3 fold significant excess risk of lymphomas. Consumption of milk showed a 6-fold excess risk (OR=1.5) while coffee showed a 50% reduction in risk for lymphoma (Ganesh et al., 2013). Physical inactivity and greater BMI are modifiable risk factors for colon cancers in both Western and Asian populations. Increasing age, BMI and attained adult height were associated with increased hazards of deaths from colorectal cancer and physical activity was associated with reduced hazard (Morison et al., 2013).

Prostate cancer ranks among the top ten specific causes of deaths in males. The specificity of Digital Rectal Examination (DRE) was 66.0% with a sensitivity of 90.9% and a positive predictive value of 38.5%. The sensitivity of serum PSA (≥4 ng/ml) in detecting carcinoma prostate was 100% and positive predictive value for serum PSA was 19% (Belbase et al., 2013). An Iranian study demonstrated the incidence rate of leukemia and lymphoma to be 5.9 per 100,000 persons. The highest incidence rate were obtained at age 70 and above (26.4) and lowest (2.3) at the age 0-9 (Tahmasby et al., 2013).

It was estimated that 0.51 million and 0.6 million persons are likely to die from cancer in year 2016 and 2021, respectively. The leading sites of mortality in males are lung, esophagus, prostate and stomach. In females, the leading sites of cancer deaths were breast, cervix and ovarian cancer (Dsouza et al., 2013).

The current communication is an extension of our

previous work: Takiar and Jayant (2013). The assessment of prevalence of cancer essentially requires good follow-up facilities including continuous knowledge of incidence and mortality of cancer at registry level. Therefore, in the absence of good follow-up facilities, the estimation of prevalence is expected to be a difficult task. The survival data presented showed that 5 years survival is quite low for the cancer sites of Oesophagus (10.5%), Myeloid Leukemia (13.8%), Hypopharynx (16.0%) and Lymphoid Leukemia (18.0%). For other sites, it was ranging between 20-30%. It was seen that the estimated % survivals falls to a great extent up to 20 years thereafter the falls are not that noticeable.

The P/I ratio calculated for selected cancer sites and for different time periods help us to calculate the limited duration prevalence. The variation in P/I ratios by different cancer sites can be explained due to their differences in survival pattern. The cancer sites with higher % survivals at 5 years period, will in general give higher life time P/I ratios as compared to those cancer sites which have lower % survivals at 5 years. For the present communication, it was assumed that the prevalence calculated with 30 years P/I ratios can be approximated to life time prevalence.

The model approach used in the paper allows us to estimate the prevalence for different cancer sites and for different time periods. The predictive reliability of the model as used in current communication has been shown to be good and beyond doubt (Takiar and Jayant, 2013). The P/I ratios calculated for selected 11 cancer sites such as Lip (10.26), Tongue (4.15), Oral Cavity (5.89), Hypopharynx (2.81), Oesophagus (1.87), Larynx (5.43), NHL (5.48), Colon (5.24), Prostate (4.61), Lymphoid Leukemia (3.42) and Myeloid Leukemia (2.65), in combination of respective incidence cases, can be used to calculate the respective prevalence for India or for any given registry area.

Realizing the fact that most of the Indian cancer registries are able to generate good incidence data but however not having good follow-up facilities, we advocate the use the P/I ratios provided by us to be used for generating the prevalence figures at their registry level. Moreover, this approach can be claimed to be very cost-effective as it only requires the knowledge of incidence number of cases and P/I ratios provided by us to calculate the prevalence figures for various cancer sites at their registry level.

References

Balasubramaniam G, Saoba S, Sarde M, et al (2013). Case control study of risk factors for non-Hodgkin lymphoma in Mumbai, India. *Asian Pac J Cancer Prev,* **14**, 775-80.

Belbase N, Agarwal C, Pokhrel P, et al (2013). Prostate cancer screening in healthy population cohort in eastern Nepal: an explanatory trial study. *Asian Pac J Cancer Prev,* **14**, 2835-8.

Dsouza N, Murthy N, Aras RY (2013). Projection of burden of cancer mortality for India 2011-2026. *Asian Pac J Cancer Prev,* **14**, 4387-92.

IARC (2011). Cancer Survival in Africa, Asia, the Caribbean and Central America. IARC Scientific Publication No. 162; Edited by Sankaranarayanan R and Swaminathan R.

Krishnamurthy A, Ramshankar V (2013). Early state oral tongue cancer among non-tobacco users – an increasing trend observed in a south indian patient population presenting at a single centre. *Asian Pac J Cancer Prev,* **14**, 5061-5.

Mishra A, Mehrotra R (2014). Head and neck cancers: global burden and regional trends in India. *Asian Pac J Cancer Prev,* **15**, 537-50.

Morison D, Parr C, Lam T, et al (2013). Behavioural, metabolic risk factors for mortality from colon and rectum cancer: analysis of data from the Asia Pacific Cohort Studies Collaboration. *Asian Pac J Cancer Prev,* **14**, 1083-7.

NCRP (2013). Time Trends in Cancer Incidence Rates 1982-2010, National Cancer Registry Programme, Indian Council of Medical Research.

Rao S, Mejia G, Thompson K, et al (2013). Epidemiology of oral cancers in Asia in past decade – an update. *Asian Pac J Cancer Prev,* **14**, 5567-77.

SPSS 14.0 (2005). SPSS for windows, Chicago, SPSS Inc.

Tahmasby B, Marnani A, Malkei M, et al (2013). Blood malignancies in Mazandaran province of Iran. *Asian Pac J Cancer Prev,* **14**, 1053-6.

Takiar R, Jayant K (2013). A model approach to calculate cancer prevalence from 5 years survival data for selected cancer sites in India. *Asian Pac J Cancer Prev*, **14**, 6899-903.