

RESEARCH ARTICLE

What is the Most Suitable Time Period to Assess the Time Trends in Cancer Incidence Rates to Make Valid Predictions - an Empirical Approach

Takiar Ramnath^{1*}, Varsha Premchandbhai Shah², Sathish Kumar Krishnan³

Abstract

Projections of cancer cases are particularly useful in developing countries to plan and prioritize both diagnostic and treatment facilities. In the prediction of cancer cases for the future period say after 5 years or after 10 years, it is imperative to use the knowledge of past time trends in incidence rates as well as in population at risk. In most of the recently published studies the duration for which the time trend was assessed was more than 10 years while in few studies the duration was between 5-7 years. This raises the question as to what is the optimum time period which should be used for assessment of time trends and projections. Thus, the present paper explores the suitability of different time periods to predict the future rates so that the valid projections of cancer burden can be done for India. The cancer incidence data of selected cancer sites of Bangalore, Bhopal, Chennai, Delhi and Mumbai PBCR for the period of 1991-2009 was utilized. The three time periods were selected namely 1991-2005; 1996-2005, 1999-2005 to assess the time trends and projections. For the five selected sites, each for males and females and for each registry, the time trend was assessed and the linear regression equation was obtained to give prediction for the years 2006, 2007, 2008 and 2009. These predictions were compared with actual incidence data. The time period giving the least error in prediction was adjudged as the best. The result of the current analysis suggested that for projections of cancer cases, the 10 years duration data are most appropriate as compared to 7 year or 15 year incidence data.

Keywords: Projection - cancer incidence rate - trend - time period - empirical approach - India

Asian Pac J Cancer Prev, 16 (8), 3097-3100

Introduction

Projections of cancer cases are particularly useful in developing countries to plan and prioritize both diagnostic and treatment facilities. It also helps in formulation of corresponding appropriate Government policies and budget allocation. Projection of cancer burden means a systematic way of prediction of number of cases for cancer in general or for specific sites for a specific period of time. One of the simplest ways to predict the cancer cases for the current period is to use the latest crude incidence rate and superimpose it on the latest population. However, when the question is to predict the cancer cases for the future period say after 3 years, 5 years or after 10 years, then it becomes imperative to use the knowledge of time trends in incidence rates as well as in population at risk.

Recently, NCRP has published a report on Time Trends in Cancer Incidence Rates (NCRP 2013). This report depicts the changes in incidence rates of cancer from 13 Population Based Cancer Registries. In the fifth chapter of the report, an attempt is made to project the number of cancer cases at India level for selected leadings sites. Three approaches were used in the report to assess the time trend

namely by *i*) Incidence rates by single years *ii*) Incidence rates by three years range and *iii*) Incidence rates by five years range. However, the important question remains as to which approach is suitable. Further, in view of around 25 years cancer incidence data being now available with NCRP for at least five cancer registries, the question arises as to which time period should be appropriate to assess the time trend so that meaningful projections are made. Further review of literature suggests that the cancer sites that are published related to assessment of time trends, in past three years, are related to Breast cancer (Moradpour et al., 2013; Wu et al., 2014), Cervix cancer (Boo et al., 2011), Ovarian cancer (Wang et al., 2014), Esophageal cancer (Kiadaleri, 2014), Lung cancer (Moradpour et al., 2013; Hashimi et al., 2014), Stomach cancer (Moradpour et al., 2013). In most of these studies the duration for which the time trend was assessed was more than 10 years while in few studies (Li et al., 2011; Kiadaleri et al., 2014; Wu et al., 2014), the duration was between 5-7 years. Thus, the present paper attempts to examine the suitability of different time periods to predict the future rates so that the valid projections of cancer burden can be done for India.

¹Statistics, Indian Council of Medical Research, ²ITC Ltd, ³National Cancer Registry Programme (ICMR), Bangalore, India *For correspondence: ramnath_takiar@yahoo.co.in

Objectives: *i)* To assess the time trends in selected five cancer sites, each, for males and females by three different time periods namely 1991-2005; 1996-2005 and 1999-2005; *ii)* To estimate the projected cancer cases for the years 2006, 2007, 2008 and 2009 for selected cancer sites by the APC of the above three time periods; *iii)* To compare the projected incidence rates with that of actually observed by selected cancer sites and selected years (2006-2009); *iv)* To assess the time period that gives the least error in terms of projections so that it can be taken as the most valid time period to assess the time trends for selected cancer sites for projections.

Materials and Methods

The cancer incidence data of selected cancer sites of Bangalore, Bhopal, Chennai, Delhi and Mumbai PBCR for the period of 1990-2009 was utilized. The five cancer sites selected for males were: Lung, tongue, mouth, prostate and NHL. The five sites selected for females were: Breast, cervix, ovary, gall-bladder and lung. The three time periods selected were namely 1991-2005; 1996-2005, 1999-2005. For the five selected sites, each for males and females and for each of the registry, the time trend was assessed and the linear regression equation was obtained to give predictions for the years 2006, 2007, 2008 and 2009. Thus, in total 50 predictions were available for each of the selected time periods. These predictions were compared with actual incidence data and errors were obtained in each case. For each site, based on the sum of squared errors, the time periods are ranked as 1, 2 and 3. The time periods which corresponded to the least squared error was ranked as 1 and the time period which gavethe maximum squared error was given the rank as 3. The ranks were pooled for all the 50 predictions. Based on pooled ranks, the time periods were judged as 1, 2 or 3.

Results

The Comparison of Predicted Incidence Rates with Actual Incidence Rates for selected cancer sites and years (2006-2009) based on the incidence rates of 15 years period (P15) is shown for Bangalore Registry in Table 1. The difference between actual and predicted incidence rates for selected 4 years indicates the extent of errors that is occurring by projections. The squared sum of errors for Breast (209.0), Cervix (155.1), Ovary (8.6), Gall bladder (0.4) and Lung (7.3) is shown in the last row of the Table 1.

The Comparison of Predicted Incidence Rates with Actual Incidence Rates for selected cancer sites and years (2006-2009) based on the incidence rates of 10 years period (P10) is shown for Bangalore Registry in Table 2. The squared sum of errors for Breast (63.3), Cervix (40.9), Ovary (7.0), Gall bladder (0.2) and Lung (1.2) is shown in the last row of the Table 2.

The comparison of predicted incidence rates with actual incidence rates for selected cancer sites and years (2006-2009) based on the incidence rates of 7 years period (P7) is shown for Bangalore Registry in Table 3. The squared sum of errors for breast (59.1), cervix (15.7), ovary (7.7), gall bladder (0.7) and lung (0.7) is shown in

the last row of the Table 3.

In order to know which time period is most suitable for prediction, the sum of squared errors for selected 4 years is tabulated for five cancer sites and three periods and shown in Table 4. For Bangalore registry, the period P7 (7 years

Table 1. Comparison of Predicted Incidence Rates with Actual Incidence Rates for Selected Years (2006-2009) -P15 period-Bangalore

Year	Incidence	Breast	Cervix	Ovary	Gall bladder	Lung
2006	P15 Predicted	29.3	15.2	6.4	1.2	3.0
	Actual	35.0	21.6	7.4	1.5	3.6
	Difference	5.7	6.4	1.0	0.4	0.6
2007	P15 Predicted	29.9	14.4	6.5	1.2	3.1
	Actual	39.9	21.5	8.7	1.6	4.3
	Difference	10.0	7.1	2.1	0.4	1.2
2008	P15 Predicted	30.5	13.7	6.7	1.2	3.2
	Actual	37.6	19.2	8.4	1.4	5.1
	Difference	7.1	5.5	1.7	0.2	1.9
2009	P15 Predicted	31.1	12.9	6.8	1.3	3.3
	Actual	36.2	18.7	7.0	1.7	4.7
	Difference	5.1	5.8	0.2	0.4	1.4
Pooled square Difference		209.0	155.1	8.6	0.4	7.3
*P15-(1991-2005)						

Table 2. Comparison of Predicted Incidence Rates with Actual Incidence Rates for Selected Years (2006-2009) for the P10 period - Bangalore

Year	Incidence	Breast	Cervix	Ovary	Gall bladder	Lung
2006	P10 Predicted	32.2	17.7	6.5	1.3	3.7
	Actual	35.0	21.6	7.4	1.5	3.6
	Difference	2.8	3.9	0.9	0.3	0.1
2007	P10 Predicted	33.3	17.4	6.7	1.3	3.9
	Actual	39.9	21.5	8.7	1.6	4.3
	Difference	6.7	4.1	2.0	0.2	0.4
2008	P10 Predicted	34.4	17.1	6.9	1.4	4.1
	Actual	37.6	19.2	8.4	1.4	5.1
	Difference	3.2	2.1	1.5	0.0	1.0
2009	P10 Predicted	35.5	16.7	7.1	1.4	4.3
	Actual	36.2	18.7	7.0	1.7	4.7
	Difference	0.7	2.0	0.1	0.2	0.3
Pooled square Difference		63.3	40.9	7.0	0.2	1.2
*P10-(1996-2005)						

Table 3. Comparison of Predicted Incidence Rates with Actual Incidence Rates for Selected Years (2006-2009) for P7 period - Bangalore

Year	Incidence	Breast	Cervix	Ovary	Gall bladder	Lung
2006	P7 Predicted	32.3	18.8	6.5	1.1	3.8
	Actual	35.0	21.6	7.4	1.5	3.6
	Difference	2.7	2.8	0.9	0.4	0.3
2007	P7 Predicted	33.4	18.7	6.6	1.1	4.1
	Actual	39.9	21.5	8.7	1.6	4.3
	Difference	6.5	2.8	2.0	0.4	0.2
2008	P7 Predicted	34.6	18.7	6.8	1.1	4.3
	Actual	37.6	19.2	8.4	1.4	5.1
	Difference	3.0	0.5	1.6	0.2	0.8
2009	P7 Predicted	35.7	18.6	6.9	1.2	4.6
	Actual	36.2	18.7	7.0	1.7	4.7
	Difference	0.5	0.1	0.1	0.5	0.1
Pooled square Difference		59.1	15.7	7.7	0.7	0.7
*P7-(1999-2005))						

Table 4. Pooled Square Differences According to Three Selected Time periods-Bangalore

Time Period	Breast	Cervix	Ovary	Gall bladder	Lung	Total
P15	209.0	155.1	8.6	0.4	7.3	-
P10	63.3	40.9	7.0	0.2	1.2	-
P7	59.1	15.7	7.7	0.7	0.7	-
Ranking						
P15	3	3	3	2	3	14
P10	2	2	1	1	2	8
P7	1	1	2	3	1	8

Table 5. Rank Based on Square Errors By Selected Sites, Periods and Cancer Registries - Females

Registry	Period	Breast	Cervix	Ovary	Gall bladder	Lung	TOTAL
Bangalore	P15	3	3	3	2	3	14
	P10	2	2	1	1	2	8
	P7	1	1	2	3	1	8
Bhopal	P15	3	1	3	3	1	11
	P10	1	2	1	1	3	8
	P7	2	3	2	2	2	11
Chennai	P15	1	2	2	2	3	10
	P10	2	3	1	3	2	11
	P7	3	1	3	1	1	9
Delhi	P15	2	3	3	1	3	12
	P10	1	2	1	2	2	8
	P7	3	1	2	3	1	10
Mumbai	P15	3	2	1	1	2	9
	P10	2	1	2	2	3	10
	P7	1	3	3	3	1	11
Pooled	P15	12	11	12	9	12	56
	P10	8	10	6	9	12	45
	P7	10	9	12	12	6	49

Table 6. Rank Based on Square Error By Selected Sites, Periods and Cancer Registries - Males

Registry	Period	Tongue	Mouth	Lung	Prostate	NHL	TOTAL
Bangalore	P15	2	1	3	3	3	12
	P10	1	2	2	1	2	8
	P7	2	3	1	2	1	9
Bhopal	P15	1	1	1	2	2	7
	P10	2	3	3	1	3	12
	P7	3	2	2	3	1	11
Chennai	P15	1	2	2	1	1	7
	P10	2	1	1	2	2	8
	P7	3	3	3	3	3	15
Delhi	P15	3	2	2	3	1	11
	P10	2	3	1	1	2	9
	P7	1	1	2	2	3	9
Mumbai	P15	3	2	1	1	1	8
	P10	2	1	2	2	2	9
	P7	1	3	2	3	3	12
Pooled	P15	10	8	9	10	8	45
	P10	9	10	9	7	11	46
	P7	10	12	10	13	11	56

Table 7. Overall Rank Based on Pooled Ranks of Males and Females Cancer Sites

Period	Females	Males	Pooled	Over All Rank
P15	56	45	101	2
P10	45	46	91	1
P7	49	56	105	3

duration) is the best period for projections for the sites of Breast, Cervix and Lung while P10 period is found to be the best in the case of Ovarian and Gall bladder cancers.

Similar calculations were repeated for four other registries and together with the results of Bangalore registry is shown in Table 5. The sum of ranks for all the five registries is shown in last three rows of the Table. The sum of ranks suggests that P10 period is the best followed by P7 and P15.

Again, in the case of Males, P15 periods appear to be the best. However, P10 is not differing much in sum as compared to P15 (Table 6). Going by overall results, it appears that P10 period is the best period for projections (Table7).

Discussion

Projections of cancer cases are particularly useful in developing countries to plan and prioritize both diagnostic and treatment facilities. The two approaches which are commonly used now-a-days in assessment of trends in cancer incidence cases are joinpoint (Joinpoint Regression Program, 2009) and linear regression approach. Different authors tend to use different time periods for assessment of trend. While most of the studies use more than 10 years data to study the time trends while few studies also use lesser time periods ranging between 5-7 years (Li et al., 2011; Kiadaliri et al., 2014; Wu et al., 2014). These raises a question as to whether the period of 5-7 years is sufficient to assess the time trends and if so what is the method to validate it? If it is felt that the period of 5-7 years period is not sufficient then what is the minimum time period which should be adopted to study the trends so that meaningful projections are done from that.

The current study was designed to explore this question, using an empirical approach. Five sites each were chosen and using different time periods, predictions were done and compared with actual observations for four consecutive years (2006-2009). The errors were calculated for each period. It may be pointed out that five registries, five cancer sites and four different years give rise to 100 observations. Thus, for male and females, a total of 200 error observations were available to validate for each period so that the best time period can be chosen to assess the time trend.

It may be noted that for a given registry, a time period was adjudged as best if the sum of errors for predictions corresponded to the lowest value among the three time periods. However, while pooling for registries, the ranks were pooled and the decision was made as to which time period is best. In order to overrule the influence of errors observed in one registry over the pooled results, the rank approach was thought better and logical. In other words, the pooling of ranks indicates as to on an average which time period is better for predictions of incidence rates for the registries. Our results have shown that neither 15 years period nor the 7 years period is best for prediction. It is the 10 years period which is found to be the superior for prediction purposes.

References

- Boo Y, Kim W, Lee H, et al (2011). Incidence trends in invasive uterine cervix cancer and carcinoma in situ, South Korea. *Asian Pac J Cancer Prev*, **12**, 1985-88.
- Hashimi M, Wang X (2014). Trend analysis of lung cancer incidence rates in Ninawa Province, Iraq from 2000-2010. Decrease and recent stability. *Asian Pac J Cancer Prev*, **15**, 385-90.
- Joinpoint Regression Program, Version 3.4.2. September 2009; Statistical Research and Applications Branch National Cancer Institute, USA.
- Kiadaliri A (2014). Gender and social disparities in esophagus cancer incidence in Iran, 2003-2009: a time trend province level study. *Asian Pac J Cancer Prev*, **15**, 623-27.
- Li J, Song Q, Zhou X, et al (2011). Major cancer mortality and changes in Yanting 2004-2009. Introduction to cancer challenges in high risk area. *Asian Pac J Cancer Prev*, **12**, 409-13.
- Moradpour F, Fatemi Z, Marçay M (2013). Estimation of the Projections of the incidence rates, mortality and prevalence due to common cancer site in Isfahan, Iran. *Asian Pac J Cancer Prev*, **14**, 3581-85.
- NCRP (2013). Time trends in cancer incidence rates (1982-2010). national cancer registry programme, Indian Council of Medical Research.
- Wang B, Liu S, Zheng R, et al (2014). Time trends of ovarian cancer incidence in China. *Asian Pac J Cancer Prev*, **15**, 191-93.
- Wu L, Han R, Zhou J, et al (2014). Incidence and mortality of female breast cancer in Jiangsu, China. *Asian Pac J Cancer Prev*, **15**, 2727-32.