RESEARCH COMMUNICATION

Cancer Incidence Rates and the Problem of Denominators - a New Approach in Indian Cancer Registries

Ramnath Takiar, B Shobana

Abstract

In India, the national census provides population figures once every 10 years. However, since cancer incidence data provide various rates for five year age groups, the calculation of the relevant population estimates for a given year between any two-census years, serving as denominators, assumes importance. The Individual Exponential Growth Rate Method is in current use by various Indian cancer registries to estimate the population by five yearly age groups. Using the five yearly age group estimates by the same method, various rates like the Crude rate, Age Standardized Rate and Cumulative Rates, are reported in Cancer Incidence in Five Continents, Vol. VIII. However, this approach has been shown to suffer from bias and often results in sacrificing the overall growth rate and corrections become necessary in five yearly age group populations to maintain it. We here show that the proposed Difference Distribution Method is able to maintain the overall growth rate and overcomes the bias in estimation of different five yearly age group populations. Further, for population projections, this method scores over the Individual Exponential Growth Method, serving as a new methodology for population estimation by five yearly age groups for inter-census years for Indian cancer registries.

Key Words: Census - cancer incidence data - denominator - inter-census population estimates - growth rate

Asian Pacific J Cancer Prev, **10**, 123-126

Introduction

The National Cancer Registry Programme (NCRP) at Bangalore, India, is a long-term activity of the Indian Council of Medical Research. It is receiving data continuously on cancer incidence cases from the fourteen Population Based Cancer Registries, shortly termed as PBCRs. The data so received for incidence cases are then analyzed and expressed in the form of various rates like the crude rate, age specific rates and the truncated rates (NCRP, 2001; 2005). All the rate calculations essentially require a knowledge of the total population (mid year), sizes of five yearly age group populations and populations for the truncated age groups.

In India, the census has been providing population figures once in every 10 years starting from the year 1951. The latest census was conducted in the year 2001. Thus, for cancer incidence data, to provide various rates for a given year, lying between any two-census years, the calculation of population estimates assumes importance. The exponential growth rate method (NCRP,2001; 2005) is in use to estimate the total population for the given year. However, to estimate the population by five yearly age groups, the method which is in current use is the Individual Exponential Growth Rate Method. The cancer incidence data reported for Indian registries of Ahmedabad, Bangalore, Chennai, Delhi, Karnuagappally, Mumbai, Nagpur, Poona and Trivendrum in IARC Publication (Parkin et al., 2002) testify the use and acceptance of the same method for calculation of five yearly populations for all the Indian registries.

The present paper attempts: 1) To highlight the limitations of the Individual Exponential Growth Rate Method currently in use by the Indian registries for the estimation of the five yearly age group populations between any two Census years; 2) to describe a new method to estimate the five yearly age group populations for inter-census years; and 3) to show that the proposed method scores over the existing used method in India in providing the population estimates for the five yearly age groups for a given year.

Materials and Methods

Consider the formulae involved in estimation of the population for a given year lying between any two-census operational years: Assume that a population P_0 after a period 't' grows to a population of P_t. Then the formula for the annual growth rate 'r' and for the population estimate after the time period ' x ' can be given as follows:

Exponential Growth Method:

$$r = (P_t / P_0)^{1/t} - 1,$$
$$P_x = P_0 * (1 + r)^x \quad \text{when } 0 < x < t \;\&$$
$$P_x = P_t * (1 + r)^x \quad \text{when } x > t.$$

*National Cancer Registry Programme, Indian Council of Medical Research, Bangalore, India *For Correspondence: ramnath_takiar@yahoomail.co.in*

The population growth rate is assumed to remain constant after the time x. However, based on the availability of the data, it can be changed suitably. The above calculations may be illustrated with the help of an example. Assume that for a certain area, the population of 1991 was 1,986,270 and which has grown to 2,269,380 in the year 2001. Now, suppose that the aim is to provide the population estimate for the year 1996. Then proceed as follows:

Here $P_0 = P_{91}$ =1,986,270; t = 2001-1991=10 years;
$P_t = P_{01}$ = 2,269,380 and x =1996-1991=5 years.
Then $r_{0191} = (2,269,380/1,986,270)^{(1/10)} - 1 = 2.5090$;
$P_{96} = P_{91} * (1+2.5090)^5 = 2,251,945$

In cancer epidemiology, the following five yearly age groups are usually considered to calculate the various age specific rates: 0-4; 5-9; 10-14; 15-19; 20-24; 25-29; 30-34; 35-39; 40-44; 45-49; 50-54; 55-59; 60-64; 65-69; 70-74; 75 & above. These age groups are referred with the help of i = 1,2,3… 16. For example for i = 1 => the 0-4 years age group is referred; for i = 2 the 5-9 years age group is referred and for i = 16 the 75 years & above age group is referred.

Methods to estimate the five yearly age group populations for different years lying between any two census operational years are described below:

1) Individual Exponential Growth Method:

This method makes use of the five yearly age distributions of immediately preceding two census years. Assume that for a given area a_{i91} and a_{i01} denote the population of the ith five yearly age group for i = 1,2,3…16 for the Census of 1991 and 2001, respectively. For the ith age group, calculate the growth rate r_{i0191} for t=10. Then, for the given year x and for the ith five yearly age group, calculate the population estimates a_{ix} as follows:

Let $r_{i0191} = (a_{i01}/ a_{i01})^{1/10} - 1$
Case I: When 1991 < x < 2001
$a_{ix} = a_{i91} * (1+ r_{i0191})^x$ for i = 1,2,3… 16.
Case II: When x >2001
$a_{ix} = a_{i01} * (1+ r_{i0191})^x$ for i = 1,2,3…16 & x > 2001;

where the growth rate is assumed to remain the same after the year 2001.

The population estimation by the five yearly age groups by the Individual Growth Method, for the Chennai PBCR area for the year 2005, utilizing the age distribution of the Census 1991 and 2001, is demonstrated. The validity of the estimates are also demonstrated by testing whether $P_{05} = \Sigma a_{105}$ for i = 1,2,3… 16 with the same set of data. Ideally, the above relationship should hold good for a method, failing which it can be concluded that there is a bias in the method.

2) The Difference Distribution Method:

A new method is proposed to calculate the five yearly age group populations for Inter-census years for cancer registries. This method also makes use of the five yearly age distributions of immediately preceding two Census years. Assume that for a given area a_{i91} and a_{i01} denote the population of the ith five yearly age group for i= 1,2,3…16 for the Census years 1991 and 2001, respectively. Then, calculate the difference (d_{i0191}) in the population for each age group and express it as the proportion (p_{i0191}) of the overall change in the population (D_{0191}). Thus, in notations:

$d_{i0191} = a_{i01} - a_{i91}$ for i = 1,2,3… 16.
$D_{0191} = \Sigma d_{i0191}$ for i = 1,2,3… 16
$p_{i0191} = (d_{i0191}/ D_{0191})$ for i = 1,2,3… 16

To estimate the five yearly age groups populations for the year x, the Difference Distribution Method requires the knowledge of the following two populations 1) Base population (P_{1991}) and 2) The population at time x (P_x). Then, proceed as follows:

Case I: When 1991 < x < 2001
Let $D_{x91} = (P_x - P_{91})$ then
$a_{ix} = a_{i91} + (D_{x91} * p_{i0191})$ for i = 1,2,3… 16
Case II: When x > 2001
Let $D_{x01} = (P_x - P_{01})$ then
$a_{ix} = a_{i01} + (D_{x01} * p_{i0191})$ for i = 1,2,3… 16

It is assumed here that the difference distribution remains the same after the year 2001.

Considering the five yearly age distribution of Census for the Chennai Population Based Cancer Registry area for the year 1991 and 2001, the calculation of the five yearly age group data for the year 2005 utilizing the Difference Distribution Method is demonstrated.

Results

Table 1 provides the estimates of five yearly age group populations for the year 2005, for the PBCR area of Chennai, arrived, using the Individual Exponential Growth Rate Method. The five yearly age group populations for the years 1991 & 2001, provided by the Census, are shown in column 2 and 3 of the Table, with total populations of 1,986,278 and 2,219,539, respectively. The individual exponential growth rate is calculated for each of the five yearly age group and is shown in column 4. A large variation (-0.0089 to 0.0547) can be seen from one to another five yearly age group. The overall exponential growth rate was observed to be 0.0112.

The estimated five yearly age group populations are shown in column 5. The estimated total population for the year 2005 using the individual exponential growth rates comes out to be 2,329,977 while using the overall exponential growth rate; it comes out to be 2,320,341. This implies the use of the individual exponential growth rates, in this case, has resulted in overestimation of the total population as compared to that arrived using the overall exponential growth rate. Further, in this method, the individual age group populations are adjusted using the correction factor (2,320,341/2,329,977 = 0.996) to meet the overall exponential growth rate and are shown in column 6. The adjustment in the individual population estimates, so made, results in the modification of the individual growth rates, as shown in column 7.

Table 1. Population Estimation using the Individual Growth Method and Effects on Individual Growth Rates for Maintaining the Overall Growth Rate - Chennai

Age Groups	1991 (a_{i91})	2001 (a_{i01})	Growth Rate 1991-2001 $(r_{i0191}) = (a_{i01}/a_{i91})^{1/10}-1$	Estimated Pop'n 2005 $a_{i05} = a_{i01}*(1+r_{i0191})^4$	Estimated Pop'n adjusted to 2005 $a_{i05}*(P_{05}/\_a_{i05})$	Modified Growth Rate $(r_{i0191})'$
0-4	167,407	156,443	-0.0068	152,261	151,631	-0.0078
5-9	191,025	174,686	-0.0089	168,548	167,851	-0.0099
10-14	195,524	200,575	0.0026	202,632	201,794	0.0015
15-19	198,581	216,803	0.0088	224,552	223,623	0.0078
20-24	216,844	231,618	0.0066	237,806	236,823	0.0056
25-29	200,616	226,999	0.0124	238,499	237,513	0.0114
30-34	168,146	197,627	0.0163	210,820	209,948	0.0152
35-39	151,786	181,515	0.0180	194,978	194,172	0.0170
40-44	125,427	148,448	0.0170	158,799	158,142	0.0159
45-49	98,541	130,115	0.0282	145,417	144,816	0.0271
50-54	81,152	104,987	0.0261	116,378	115,897	0.0250
55-59	62,384	70,861	0.0128	74,566	74,258	0.0118
60-64	52,094	64,215	0.0211	69,820	69,531	0.0200
65-69	31,846	44,587	0.0342	51,012	50,801	0.0332
70-74	22,924	32,622	0.0359	37,566	37,411	0.0348
75+	21,982	37,437	0.0547	46,323	46,131	0.0535
All ages	1,986,278	2,219,539	0.0112	2,329,977	2,320,341	0.0112
			$P_{05} = P_{01}*(1+r_{0191})^4 =$	2,320,341		

Table 2. Population Estimation by Five Yearly Age Groups, using the Different Distribution Method - Chennai - 2005

Age Groups	1991 (a_{i91})	2001 (a_{i01})	Difference $d_{i0191} = (a_{i01} - a_{i91})$	Difference proportion $p_{i0191} = (d_{i0191}/D_{0191})$	Estimated Growth by 2005 $gi_{0501} = (D_{0501}*p_{i0191})$	Estimated Population $a_{i05} = a_{i01} + g_{i0501}$
0-4	167,407	156,443	-10,964	-0.0470	-4,738	151,705
5-9	191,025	174,686	-16,339	-0.0700	-7,061	167,625
10-14	195,524	200,575	5,052	0.0217	2,183	202,758
15-19	198,581	216,803	18,222	0.0781	7,875	224,678
20-24	216,844	231,618	14,774	0.0633	6,384	238,003
25-29	200,616	226,999	26,383	0.1131	11,401	238,400
30-34	168,146	197,627	29,482	0.1264	12,740	210,367
35-39	151,786	181,515	29,729	0.1274	12,847	194,362
40-44	125,427	148,448	23,020	0.0987	9,948	158,396
45-49	98,541	130,115	31,575	0.1354	13,645	143,761
50-54	81,152	104,987	23,835	0.1022	10,300	115,287
55-59	62,384	70,861	8,477	0.0363	3,663	74,524
60-64	52,094	64,215	12,122	0.0520	5,238	69,454
65-69	31,846	44,587	12,741	0.0546	5,506	50,093
70-74	22,924	32,622	9,697	0.0416	4,190	36,812
75+	21,982	37,437	15,455	0.0663	6,679	44,116
Total	1,986,278	2,219,539	233,261	1.0000	100,802	2,320,341
$D_{0191} = \Sigma(d_{i01})$	233,261	$P_{05} =$	2,320,341	$D_{0501} = P_{05} - P_{01}$	100,802	

The Table 2 provides the five yearly age groups population estimates, using the Difference Distribution Method. The five yearly age group populations for the Census years of 1991 and 2001 are shown in column 2 & 3, respectively. Their differences and the corresponding difference proportions (p_{i0191}) are shown in column 4 & 5, respectively. The estimated growths and the estimates of populations for five yearly age groups for the year 2005 are provided in column 6 & 7, respectively. The population estimates of 0-4 years and 5-9 years have shown a decreasing trend while the other age groups have shown an increasing trend as compared to the populations of the year 2001.

Using the Census data of 1991 and 2001 for five Indian cancer registry areas, the population estimates are arrived by both the methods for the year 2005. Comparisons reveal that each time, the individual exponential growth method overestimate the total population as compared to the over all exponential growth rate (Table 3).

Discussion

It is quite evident from the data provided in Table 1 that the annual growth rate varies between age groups. Interestingly the population decreases in 0-4 years and 5-9 years suggesting that these age groups have registered a negative growth while all other age groups have registered a positive growth. Hence, to assume the uniform growth in all five yearly age group populations and thereby to use the same proportional distribution as seen in the

Table 3. Comparison of 2005 Population Estimates* by the Individual and Overall Exponential Methods

Registry	Males			Females		
	Individual	Overall	Ratio	Individual	Overall	Ratio
Bangalore	3,874,164	3,860,766	1.0035**	3,531,651	3,512,064	1.0056
Bhopal	1,102,322	1,099,885	1.0022	990,793	987,397	1.0034
Chennai	2,329,977	2,320,341	1.0042	2,253,489	2,242,322	1.0050
Delhi	8,901,916	8,888,097	1.0016	7,297,247	7,271,169	1.0036
Mumbai	7,167,323	7,150,170	1.0024	5,789,909	5,763,697	1.0045

*Based on 1991 and 2001 Census data; ** - 3,874,164/3,860,766 = 1.0035

previous Census to current estimated populations is not justifiable. Further, calculation of the individual exponential growth rate for each five yearly age group population presupposes a static situation, which is not true. In fact, the population in each five yearly age group is quite dynamic and by the end of the each five years, the entire population of 0-4 years moves on to the 5-9 year age group and so on.

It is shown in Table 1 that the estimated total population for the year 2005 using the individual exponential growth rates comes out to be 2,329,977 while using the overall exponential growth rate, it comes out to be 2,320,341 implying that the use of the individual exponential growth rates, in this case, has resulted in overestimation. This was further substantiated by the data provided in Table 3 whereby it get clear that the Individual Exponential Growth Method often results in over or under estimation of the estimates. Further, in this method, when the individual age group populations are adjusted to correct for the overestimation in the total population, results in reduction in the individual growth rates with a marked variation from one age group to another. This exhibits the inability of the Individual Exponential Growth Rate Method to maintain the overall growth rate. Further, in an effort to maintain the overall growth rate, by suggesting the corrections, results in modifications in the individual population growth rates. Thus it can be summarized that the Individual Exponential Growth Rate Method ultimately results in either sacrificing the overall growth rate or it fails in maintaining the individual growth rates, the assumption with which the populations are projected. Hence, this method is not appropriate for the estimation of the population by the five yearly age groups.

It is shown that the proposed Difference Distribution Method is able to maintain both the negative as well as the positive growth in different five yearly age group populations. It has maintained the negative growth in 0-4 years and 5-9 years and maintained the positive growth in other five yearly age groups. The proposed method also maintains the overall growth rate unlike seen in the case of Individual Exponential Growth Rate Method. Thus, the proposed method scores over the current method, which is in use in the projection of the five yearly age group populations.

Provision of the correct estimates for the five yearly age group populations has far reaching impact on various rates, which are calculated for cancer incidence data or other rates based on population like age specific fertility rates, age specific mortality rates and the age specific prevalence rates for the nutritional deficiency signs. To assess the time trends in the age adjusted rates, for the important sites of cancer (breast, cervix, oesophagus, lung), is an important exercise and unless the five yearly age group estimates are arrived properly it will not be possible to assess the time trends correctly. Hence, this proposed method will be useful not only in assessing the time trends, in calculation of various cancer related rates but also be useful in other age specific rates based on the different population groups. This method also serves as a new methodology in population estimation by different age groups.

Acknowledgement

The authors of the paper are grateful to the Officer-in-Charge, Dr. A. Nandakumar, National Cancer Registry Programme, Indian Council Of Medical Research, Bangalore, India, for his constant support and encouragement for pursuing the problem and writing of the manuscript.

References

National Cancer Registry Programme (ICMR)(2001). Consolidate Report of the Population Based Cancer Registries: 1990-1996. Bangalore, India.

National Cancer Registry Programme (ICMR) (2005). Two Year Report of the Population Based Cancer Registries: 1999-2000. Bangalore, India,

Parkin DM, Whelan Sl, Ferlay J, Teppo M, Thomas DB (2002). Cancer Incidence in Five Continents Vol. VIII, IARC Press, Lyon, France.