

## RESEARCH ARTICLE

# Estimation of Denominators- a New Approach for Calculating of Various Rates in Cancer Registries

AS Haroon<sup>1</sup>, SM Gupta<sup>1</sup>, BB Tyagi<sup>2\*</sup>, J Farhat<sup>3</sup>

### Abstract

In this study, cancer incidence data were assessed to provide various rates of five year age groups for a given year, lying between two census years. The individual exponential growth rate method is most useful in both population-based and non-population based cancer registries in India to estimate the population by five yearly age groups and also find the rates of crude rates, age standard rates and cumulative rates. This method has been shown to endure from bias and often results sacrificing the overall growth rate and correction factor must be needful in five year age group population to maintain it. A second method, the difference distribution method is also able to maintain the overall growth rate and overcome the bias in estimation of five yearly age group populations. From this point of view these methods serving a new technique for population estimation by five yearly age groups for inter census years.

**Keywords:** Census - cancer incidence data inter census population estimation - growth rate and modified growth rate

*Asian Pacific J Cancer Prev*, 13, 3229-3232

### Introduction

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 2011. Thus for cancer incidence data to provide various rate for given year, lying between two census years, calculations for the population estimates assume importance. This data obtained forms a benchmark for framing the welfare developments, policies for billion plus human resources. The census data validation and scrutiny is completed by the all states/union territories under the overall supervision and monitoring of the census division of the office of the Registrar General of India.

The data is also useful in National Cancer Registry Programme (NCRP) at Bangalore, Bhopal, Chennai and Delhi and other areas, a long-term activity of the Indian Council of Medical Research (ICMR). It is receiving data continuously on cancer incidence cases from the fourteen Population Based Cancer Registries (PBCR). 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 rates calculation essentially requires a knowledge of the total population (mid year) size of five yearly age group populations and population for the truncated age groups.

The cancer incidence data reported for Indian cancer registries of Ahmadabad, Bangalore, Chennai, Delhi, Kurnuagappally, Mumbai, Nagpur, Poona and Trivandrum in IARC Publication (Parkin, et al., 2002) testified the used and acceptance of the same method for calculation of five yearly populations for all the Indian Cancer registries. Takiar and Shobana (2009) showed that Individual Exponential Growth Method and Difference distribution method are able to maintain the Overall Growth Rate and overcomes the bias in estimation of different years age group population and population projections. These method serving a new technique for a population estimation by five yearly age groups for inter-census year for Indian cancer registries.

### Materials and Methods

The population growth can be estimated by using exponential growth rate method as below

$$r = (P_x/P_0)^{1/x} - 1$$

Here  $P_x = P_0^*(1+r)^x$  if  $0 < x < t$   
and  $P_x = P_t^*(1+r)^x$  if  $X > t$

Where  $P_x$  is population at time  $x$  and  $P_0$  is present population and  $r$  is annual growth rate. The population growth rate is assume to remain constant after the time  $x$ , which based of the availability of the data, it can be change suitably. This method can be used to estimate the

<sup>1</sup>Department of Statistics, Meerut College, Meerut, <sup>2</sup>AIIMS, <sup>3</sup>Department of Zoology, Daulat Ram College, University of Delhi, New Delhi, India \*For correspondence: [btyagi51@gmail.com](mailto:btyagi51@gmail.com), [btyagi@yahoo.co.in](mailto:btyagi@yahoo.co.in)

five yearly age group populations for different years, lying between any two census operational years.

*Individual Exponential Growth Methods*

This method makes use of the five yearly age group population of immediately preceding two census year. As same that for a given area  $a_{i91}$  and  $a_{i01}$  denote the area population of the  $i^{th}$  five yearly age group for  $i = 1, 2, \dots, 17$  for the census of 1991 and 2001, respectively. For the age group, calculate the growth rate  $r_{i091}$  for the  $t=10$ . Then for the given year  $x$  and for the  $i^{th}$  five yearly age group, calculate the population estimates  $a_{ix}$  as follows. Let

$$r_{i091} = (a_{i01}/a_{i91})^{1/10} - 1$$

Case -I: when 1991  $X <$  2001

$$a_{ix} = a_{i91} * (1 + r_{i091})^x \quad \text{for } i = 1, 2, 3, \dots, 17$$

Case -II: when  $x >$  2001

$$a_{ix} = a_{i01} * (1 + r_{i091})^x \quad \text{for } i = 1, 2, 3, \dots, 17$$

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

*Some important term as below*

Estimated population of 2009 =  $a_{i01} * (1 + r_{i091})^{1/8}$

Adjusted population = correction factor \* estimated population of corresponding year.

Correction factor = Adjusted population of 2009 / Total population of 2009

Modified growth rate = (Estimated population of 2009 / Total population of 2001)<sup>1/8</sup> - 1

The population estimation for five yearly age-groups by using the individual growth method for the Meerut area for the year 2009, utilizing the age distribution of the census 1999 and 2001 is demonstrated. The validity of the estimates is also demonstrated by testing whether  $\sum P_{09} = \sum a_{i09}$  for  $i = 1, 2, \dots, 17$  with the same set of data. Ideally, the above relationship should held good for a method, failing which it can be concluded that there is bias in the method.

*Difference Distribution Method*

This is a new method is proposed to calculate the five yearly age group populations for inter-census year for cancer registries. This method also makes use of the five yearly age distribution of immediately preceding two census year. Assume that for a given area  $a_{i91}$  and  $a_{i01}$  denote the population of the  $i^{th}$  five yearly age group for  $i = 1, 2, \dots, 17$  for the census year 1991 and 2001 respectively. Then calculate the difference between both census is denote by the  $(d_{i01,91})$  in the population from each age group and express, it as the proportion is denoted by  $(P_{i01,91})$  in the population from each age group and express it as the proportion is denoted by  $(p_{i01,91})$  of the area all change in the population is denoted by  $(D_{i01,91})$ . Thus in nations:

$$d_{i01,91} = a_{i01} - a_{i91} \quad \text{for } i = 1, 2, 3, \dots, 17$$

$$D_{01,91} = \sum d_{i01,91} \quad \text{for } i = 1, 2, 3, \dots, 17$$

Then

$$P_{i01,09} = (d_{i01,91} / D_{01,91}) \quad \text{for } i = 1, 2, 3, \dots, 17$$

To estimate the five yearly age groups population for the year  $x$ , Then Difference Distribution Method require the knowledge of the following two populations.

1. Base – population as  $(P_{1991})$

2. The population at time  $x$  ( $P_x$ ) then, proceed as follows.

Case- I: when 1991  $< x <$  2001

Let

$$D_{x91} = (P_x - P_{91}) \text{ then}$$

$$a_{ix} = a_{i91} + (D_{x91} * P_{i01,91}) \text{ for } i = 1, 2, 3, \dots, 17$$

Case -II: when  $x >$  2001

$$D_{x01} = (P_x - P_{01}) \text{ then}$$

$$a_{ix} = a_{i01} + (D_{x01} * P_{i01,91}) \text{ for } i = 1, 2, 3, \dots, 17$$

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

**Results**

*Using the individual exponential growth method*

Estimates were made of five yearly age group male populations for the year 2009, for population based non cancer registries area of Meerut in Uttar Pradesh. The five yearly age group populations for the year 1991 and 2001 census populations for total males and females in Meerut city were 684200, 774670 and 592348, 677313 respectively. The individual exponential growth rate calculate for the five yearly age group demonstrated large variation (-0.0744, which is smallest negative value and 0.0315 is largest positive value and maximum growth rate is 0.0475 and minimum is - 0.0672). The overall exponential growth rate was observed to be 0.0125 in year 1991 to 2001.

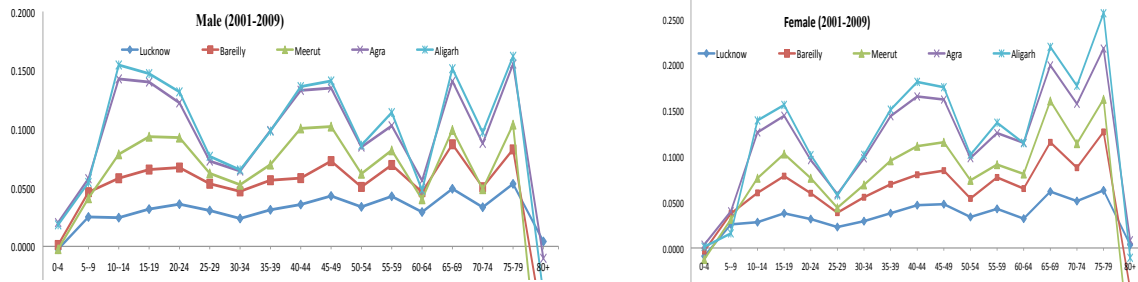
The estimated total populations for the year 2009 were 869,590 and 767,466 using the individual exponential growth rate method and the adjusted population for 2009 was 855,623 and 753,984. From the above the individual exponential growth rate method had over estimated the total population as compare to exponential growth rate method for adjusted population for 2009 so correction factors of  $855,623/869,590 = 0.9861$  and  $753,984/767,466 = 0.9824$  were required.

The adjustment in the individual population estimated result in the modification of the individual growth rate (2001-2009) and it was found to be same (0.0125 and 0.0135) as that individual growth rate during the period 1991- 2001.

*Difference Distribution Method*

The population was estimated by difference distribution method separately for male and female for year 2009 for five yearly age groups using the different distribution method in Meerut City. The growth rate and population estimate for 2009 had shown increasing trend in age group 0-4, 5-9 and 80+ in both the sexes. Whereas, the decreasing trend was shown in remaining age groups in both the sexes.

Similarly, for five non cancer registries area as, Meerut, Lucknow, Bareilly, Aligarh and Agra in Uttar Pradesh population were also estimated by using individual exponential growth rate method and difference distributed method had given approximately same results for male and female respectively (Figure 1A and B). Figure 1A indicates that male modified growth rate in five non cancer registries at Uttar Pradesh in the year 2001 and estimated population for year 2009. Agra City shown that maximum modified growth rate in 80 + year age group of



**Figure 1. Modified Growth Rates between 2001 and 2009 in Five Non Cancer Registries Area of Uttar Pradesh.**  
A) Male Population, B) Female Population

**Table 1. Comparison of Estimated Population by Sex in Five Non Cancer Registries Area of UP, Year 2009**

City	Male			Female		
	Individual	Overall	Ratio	Individual	Overall	Ratio
Meerut	869590	855623	1.016	767466	753984	1.018
Aligarh	481766	475147	1.014	422328	419494	1.007
Agra	1151870	1119029	1.029*	969138	945849	1.025
Bareilly	772556	764642	1.01	697698	688755	1.030*
Lucknow	155676	1542738	1.009	1412810	1391575	1.015

**Table 2. Crude, World Age Adjusted (AAR) & Truncated (35-64 years) Incidence Rates (TR) per 100,000 Persons, Site-Breast, Sex-Females in Meerut City, Year-2009**

Age Group	No.of Cases	Crude Rate	Age-Adjusted Rate	Truncated Rate
0-4	0	0	0	0.00
5--9	0	0	0	0.00
10--14	0	0	0	0.00
15-19	0	0	0	0.00
20-24	0	0	0	0.00
25-29	4	7.15	0.57	0.00
30-34	18	33.02	2.51	0.00
35-39	37	63.79	4.56	4.56
40-44	41	88.6	5.84	5.84
45-49	51	134.55	8.13	8.13
50-54	40	171.15	9.19	9.19
55-59	32	174.89	7.96	7.96
60-64	31	189.8	7.06	7.06
65-69	14	89.19	2.64	----
70-74	10	109.12	2.41	----
75-79	6	131.58	2	----
80+	1	30.13	0.45	----
Total	285	37.14	53.32	42.74

the population compare to other cities and Bareilly City show that minimum modified growth other rate of the 80+ year age group of the population compare to other cities. Figure 1B shows that 80+ years female age group in Meerut city was found minimum modified growth rate for year 2001 and for year 2009 compare to other non cancer registries in Uttar Pradesh and 80+ years female age group in Agra city was found maximum growth rate for respective year compare to other non cancer registries in Uttar Pradesh.

Table 1 shows that the total population for year 2009 by using individual exponential growth rate method was found over estimated adjusted population for year 2009 by using the individual exponential growth rate method. In case of male the maximum ratio of total population and adjusted population was found in Agra city and minimum

ratio was found in Lucknow compare to other non cancer registries in Uttar Pradesh. Similarly, in females the maximum ratio of total population and adjusted population was found in Bareilly and minimum ratio was found in Aligarh city compare to other non cancer registries of Uttar Pradesh.

Table 2 shows that all breast cancer cases were collected particularly in Meerut city i.e. Non Population Based Cancer Registry area in Uttar Pradesh during the study period from 1st January 2008 to 31st December 2009. The crude incidence rate (CR) was 37.14, world age-adjusted (AAR) rate was 53.3 and the truncated incidence rate (35-64 years) was 42.7 per 100,000 persons.

## Discussion

The population of five non cancer registries areas in Uttar Pradesh as on 1<sup>st</sup> July 2009 was estimated by assuming an exponential rate of growth for each group by taking the base of 1991 and 2001 census figures. The estimated population of Meerut city comes to 869,590 for males and 767,466 for females by using of exponential rate of growth, while by using Difference Distribution Method for year 2009 the estimated population as on 1<sup>st</sup> July 2009 comes to 855,623 for males and 753,984 for females respectively. The ratio of total population and adjusted population was found 1.016 and 1.018.

From the above results, by individual exponential growth rate method the estimated total population as compare to adjusted population for 2009 the correction factor was found 0.9861 and 0.9824 corresponding to male and female. Further by this method when the individual age group population results in reduction in the individual growth rates with a marked variation from one to another age group like as 0-4, 5-9, 60-64 and 80+ year. This exhibits the inability of the individual exponential growth rate method to maintain the overall growth rate, by suggesting the corrections, results in either sacrificing the overall growth rate or it fails in maintaining the individual 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 assumptions with which the population is projected, hence, this method is not appropriate for the estimation of the population for the five yearly age groups. And second method (Difference Distribution Method) is able to maintain both the negative as well as the positive growth in different five yearly age group populations. It have maintained the negative

growth in 0-4, 5-9 and 80+ years and maintain the positive growth in other five yearly age group. But this method is given same result of total male and female population by using individual exponential growth rate method. The adjustment in the individual population estimated result in the modification of the individual growth rate (2001-2009) and it was found to be same 0.0125 and 0.0135 as that individual growth rate during the period 1991-2001.

Provision of the correct estimation for the five yearly age group populations has for reaching impact on various rates. Which are calculated for cancer incidence rates based on population like age specific incidence rates, age specific mortality rates and age specific prevalence rates for the nutritional deficiency sign? To assess the time trends in the age adjusted rates for the important sites of cancer (breasts, cervix, esophagus, lungs) in 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 the 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.

## References

- Ahmad O, Boschi-Pinto C, Lopez AD, et al (2001) Age standardization of rates: A new WHO standard, GPE Discussion Paper Series : No-31.
- Census of India 1991, Office of the Registrar General of India.
- Census of India 2001, Office of the Registrar General of India
- ICMR(2001). National Cancer Registry Programme, Consolidated report of the population based cancer registries, 1990-1996 Bangalore India. Indian Council of Medical Research, New Delhi.
- National Cancer Registry Programme (ICMR)(2005). Two Year Report of the Population Based Cancer Registries: 1999-2000. Bangalore, India.
- National Cancer Registry Programme(2008): Two Year Report of the Population Based Cancer registries (2004-05).
- National Cancer Registry Programme(2009) Time Trends in Cancer Incidence Rates (1982-2005).
- NCRP Survey (2005) Comparison of Age Distribution in Area 1 and Area 2- Bhopal.
- Parkin DM, Whelan SI, Ferlay J, Teppo M, Thomas DB (2002). Cancer Incidence in Five Continents Vol. VIII, IARC Press, Lyon, France.
- Population Census of India 2011 (from internet).
- Takiar R, Shobana B (2009). Cancer incidence rates and the problem of denominators - a new approach in Indian cancer registries. *Asian Pac J Cancer Prev*, **10**, 123-6.
- Wald A, Wolfowitz J (1993). An exact test for randomness in the non-parametric case based on serial correlation. *Ann Math Stat*, **14**, 378-88
- Winberg, GG ( 1971). Methods for Estimation of Production of Aquatic Animals. Academic Press, London, p. 175.
- Weatherley, AH, Gill, HS (1987). The Biology of Fish Growth. Academic Press, London, p. 443.