

# INTRODUCTION

Cancer is a disease in which there is unregulated cell growth in any organ systems occurring in humans of all age groups, irrespective of age, sex, nationality, ethnicity, economic status, educational strata, geological and geographic distributions. In its presentation, it could be acute (sudden onset), sub-acute (slow onset), or chronic (long period of time). In its symptoms, it is known to present itself in the most insidious non-specific presenting symptoms like fever, diarrhoea or weight loss to the symptoms like bleeding, obstructive symptoms, growths. As a disease it has the potential to restrain a person from achieving full physical, physiological, psychological and economic potential. It's a major concern for the patient, his/her family, the clinician, the healthcare provider and the taxpayer.

The aetiology of cancer is multi-factorial: genetic predisposition, exposure to tobacco, certain chemicals, infections, radiation, inappropriate lifestyle factors (alcohol, inappropriate diet, physical inactivity, high body mass index, diabetics and metabolic syndrome) have all been implicated in the causation of cancer. Yet, the cause remains undetermined in a large proportion of patients. This is important since much of the preventive strategies are based on some of the known factors. Associations have also been made on the degree of exposure, dose of exposure, duration of exposure, age (vulnerability) of the exposed and the like. All these factors have come from deductive reasoning of epidemiological data and an insight into the possible causative mechanisms. There have been remarkable successes in the evolving treatment modalities which have strengthened the fight against cancer.

## **The National Cancer Registry Programme - An Overview**

Cancer registry is an organization of systematic collection, storage, analysis, interpretation and reporting of data on patients with cancer (IARC). A proper analysis and interpretation of data provides insights with inputs for its prevention, control and management.

Time-trend studies are also possible when data have been accumulated over long periods of time. In addition to incidence figures, population-based cancer registries who conduct follow-up of their patients are able to estimate the prevalence of cancer. Prevalence figures give an indication of the existing burden of the disease in the community.

A cancer registry provides an economical and efficient method of ascertaining cancer occurrence rather than intervention trials and cohort studies.

In India, the National Cancer Registry Programme (NCRP) under the Indian Council of Medical Research (ICMR) with its network of cancer registries was started in December 1981 with the co-ordinating centre at Bengaluru. Presently it is operated by the ICMR-NCDIR, Bengaluru. This provides the data on cancer incidence, mortality, pattern, trend and geo-pathological distribution of cancers. It also helps to formulate and implement policies and programmes, monitor and evaluate the cancer control activities.

There are two types of cancer registries under the programme. Population Based Cancer Registries (PBCRs) record all the new cancer cases occurring in a defined population within a geographic area. The Hospital Based Cancer Registries (HBCRs) record information on cancer patients attending a particular hospital, with focus on clinical care, treatment and outcome. Cancer Atlas approaches have also been used for specific short-term purposes.

The main objectives of the programme are:

1. To generate reliable data on the magnitude and patterns of cancer.
2. Propose further epidemiological studies based on results of registry data.
3. Help in designing, planning, monitoring and evaluation of cancer control activities under the National Programme for Prevention and Control of Cancer, Diabetes, Cardiovascular Diseases and Stroke (NPCDCS).
4. Develop training programmes in cancer registration and epidemiology.

NCRP started with a network of three PBCRs in Bangalore, Chennai and Mumbai and three HBCRs in Chandigarh, Dibrugarh and Thiruvananthapuram. The number of registries working under the programme have expanded greatly from the time of inception and presently there are 36 PBCRs and 236 HBCRs registered under NCRP.

Since cancer is not a notifiable disease, cancer registration in India is active and staff of all registries visit hospitals, pathology laboratories and all other sources of registration of cancer cases on a routine basis. Death certificates are also scrutinized from the local government units like municipal corporation and panchayat raj institutes and information is collected on all cases where cancer is mentioned as a cause of death on the death certificates.

The information that is collected on a core form is entered into a software provided by ICMR - NCDIR. The data is further transmitted to ICMR - NCDIR. Over the years, the registries and the office of the NCRP have used modern advances in electronic information technology to enter the data, checking of the data, verification of duplicates and matching of mortality and incidence records. The software applications developed by NCDIR have further evolved and so has the data submission methodology and overall support. Data quality is assessed at the coordinating unit under different dimensions like comparability, validity, timeliness and completeness. Frequent training and re-training programs are conducted for cancer registry investigators and staff to maintain quality of work. Interaction with local health and other stakeholders is undertaken by the registries to keep them informed and to firm up partnerships.

To improve the mortality data, all-cause mortality data is being collected in electronic form under NCRP. The same is being formatted, coded, checked and imported at NCRP to run the matches with the incidence.

The data from the NCRP has contributed significantly for improving public health and clinical patient care. Data from the NCRP registries is used as a basis for several research studies. Data is also regularly published in successive volumes of Cancer Incidence in Five Continents (CI 5) published by the International Agency for Research on Cancer - the cancer research arm of the World Health Organization (IARC-WHO). The incidence data from 15 PBCRs of India have been published in CI 5 - Vol XI published by IARC-WHO.

India as a country has demographically been known to have large proportions of younger population. If a comparison is drawn with some of the developed countries, the Indian age pyramid has a broader base (among lower age groups). The NCRP has witnessed a steady rise in incidence of cancer over the years and with larger number of populations in higher age groups, one of the reasons of rising incidence is the increasing life expectancy.

The proportion of population of India which has developed cancer would have great interest in knowing the stage-based survival of the type of cancer and also how much the advances in medical sciences could have controlled it.

The policy makers of the healthcare delivery system would like to know about the benefit of the availability of primary, secondary and tertiary health care and its impact in improving the survival and quality of life of cancer patients.

The clinicians treating it would be interested in knowing as to what the general trend of cancer has been, how effective is a particular modality of treatment, what are the average survival rates, any changes in the occurrence as per site and the like.

While all these three issues are directly or indirectly addressed by cancer registries, the possibilities of using the data in conjunction with other ongoing health plans are endless. The integration of survival data, hospitalisation data, morbidity data with preventive strategies, health education, provisioning of basic anti-cancer medications, provisioning of tertiary healthcare facilities to cover untouched areas are all potential areas where data driven knowledge can be of immense help.

Cancer registration in India face several challenges. Cancer is not a notifiable disease, and these poses data collection challenges. A few states have issued administrative notifications for the same. The mortality registration system has several gaps in the way mortality data is recorded affecting the coverage and completion of cause of death information. Cancer registries need to be linked to several other databases at national and local levels for seamless improvement of cancer statistics (Ayushman Bharat, other insurance scheme, mortality databases, Health Management Information System).

Cancer registries form the backbone of cancer prevention and control activities in India. Strengthening it will yield much improved information to track and monitor population and hospital level measures to track cancer.

## Definitions, Statistical Terms and Methods

**Cancer Registration** may be defined as the process of continuing, systematic collection of data on the occurrence and characteristics of reportable neoplasms with the purpose of helping to assess and control the impact of malignancies on the community.

**Cancer Case** refers to all neoplasms with a behaviour code of '3' as defined by the International Classification of Diseases - Oncology, Third edition (ICD-O-3) are considered reportable and are registered in NCRP.

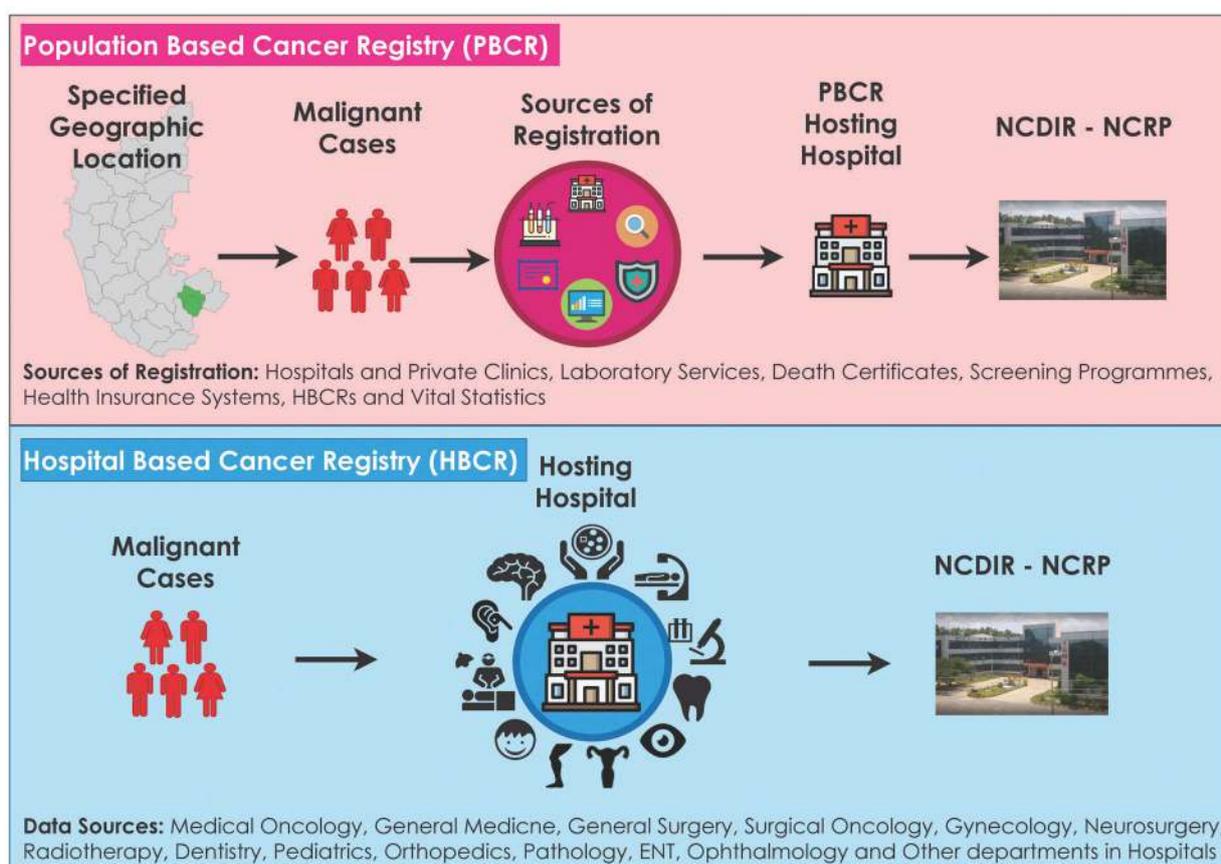
**Cancer Registry** is the office or institution which attempts to collect, store, analyse and interpret data on persons with cancer.

**Population Based Cancer Registries** (PBCRs) systematically collect information on an reportable neoplasms from multiple sources in a geographically defined population residing in the area for a period of one year.

**Hospital Based Cancer Registries (HBCRs)** are concerned with recording of information on the treatment, management and outcome of cancer patients registered in a particular hospital.

**Sources of Registration** will usually be hospitals or cancer centres but, depending on the local circumstances, a population-based registry will also involve private clinics, general practitioners, laboratories, health insurance systems, HBCRs, screening programmes and Vital statistics Department.

**Data Processing** Data Processing involves importing or downloading of data from the registries into the local database at ICMR-NCDIR. Quality of the data is checked for errors that may have been committed at data collection, abstraction or entry. Identification and elimination of duplicates is done through deterministic approach and by identifying names that are phonetically the same. Multiple combination of variables are used to generate the probable list. Duplicate deletion is done without any loss of information. Mortality data is linked/matched with incidence and the unmatched mortality cases are identified as either Death Certificate Notification (DCN)/ Death Certification Only (DCO). Clarification at each step is sought from each registry and the data is finalized for further analysis



**Age-Group** used for estimating populations as well as grouping cancer cases as per the WHO guidelines which is 0-4, 5-9, 10-14....75+.

According to the same definition the age group 0-14, 0-19 constitutes childhood cancer.

**Cancer Incidence** denotes new cases diagnosed in a defined population in a specified time period.

**Cancer Mortality** denotes number of cancer deaths occurring in a specified population during a specified time period.

**Rates** for cancer are always expressed per 100,000 population. For childhood cancer this may be expressed as per one million.

**Crude Incidence Rate (CR)** refers to the rate obtained by division of the total number of cancer cases by the corresponding estimated population (mid-year) and multiplying by 100,000.

$$CR = \frac{\text{New Cases of cancer of a particular year}}{\text{Estimated population of the same year}} \times 100,000$$

**Age Specific Rate (ASpR)** refers to the rate obtained by division of the total number of cancer cases by the corresponding estimated population in that age group and gender/site/geographic area/time period and multiplying by 100,000.

$$ASpR = \frac{\text{New Cases of cancer of a particular year in the given age group}}{\text{Estimated population of the same year for the give age group}} \times 100,000$$

**Age Adjusted or Age Standardised Rate (AAR)** Cancer incidence increases as age increases.

Therefore, higher the proportion of older population, higher is the number of cancers. Most developed and western countries have a higher proportion of older population. So in order to make rates of cancer comparable between countries, a world standard population (given below) that takes this into account is used to arrive at age adjusted or age standardised rates. This is calculated according to the direct method (Boyle and Parkin, 1991) by obtaining the age specific rates and applying these rates to the standard population in that age group. The world standard population approximates the proportional age distribution of the world and is given below:

#### Age Distribution of World Standard Population (Segi.et.al)

Age Group	World Standard Population
00-04	12,000
05-09	10,000
10-14	9,000
15-19	9,000
20-24	8,000
25-29	8,000
30-34	6,000
35-39	6,000
40-44	6,000
45-49	6,000
50-54	5,000
55-59	4,000
60-64	4,000
65-69	3,000
70-74	2,000
75+	2,000
All Ages	100,000

$$AAR = \frac{\sum_{i=1}^A (a_i w_i)}{\sum_{i=1}^A w_i}$$

Where,  $a_i$  is the age specific rate (AspR) in age class  $i$ ;  
 $w_i$  is the standard population in age class  $i$ ;  
 $A$  represents the number of age intervals.

Or expressed in simpler terms thus:

$$AAR = \frac{\Sigma(ASpR) \times (\text{No. of persons in Std. world population in that 5 yr. age group})}{100,000}$$

**Cumulative Risk** refers to the probability that the person will develop a particular cancer during a certain age period in the absence of any other cause of death. The Cumulative Rate (CuR) is an approximation of the cumulative risk. It is obtained by adding the annual age-specific incidence rates for each five-year age interval (up to either 64 or 74 years of age or for whatever age group is to be used to calculate the cumulative risk) multiplied by 5 (representing the five-year age interval) times 100/100,000.

$$CuR = \frac{5 \times \Sigma(ASpR) \times 100}{100,000}$$

And cumulative risk is expressed as

$$\text{Cumulative Risk} = 100 \times [1 - \exp(-\text{cumulative rate}/100)]$$

$$\text{Possibility one in number} = (1 / \text{Cumulative Risk}) \times 100$$

**Truncated Age Adjusted Incidence Rate (TR)** - This is similar to the age adjusted rate except that it is calculated for the truncated age group 35-64 years of age.

**Sex Ratio** is used to describe the number of females per 1000 males.

**M/I Ratio Percent** is obtained by dividing the mortality count by the incidence count in a given year (%).

**Trends in Crude Rate or Age Adjusted Incidence Rates** - The significance of trend in CR or AAR was assessed based on Joinpoint regression.

**About Joinpoint Regression Program** - Joinpoint Regression Program, Version 4.7.0, is a statistical software for the analysis of trends using Joinpoint models, i.e, where several different regression lines are connected together at the "Joinpoints". The software takes trend data (e.g. cancer rates) and fits the simplest Joinpoint model that the data allow. The program starts with the minimum number of Joinpoints (e.g. 0 Joinpoint, which is a straight line) and tests whether more Joinpoints are statistically significant and must be added to the model (upto that maximum number). In this report we have seen Annual Percent Change (APC) of straight line for a specified period of time.

For example, if the APC is 1%, and the rate is 50.0 per 100,000 in 2000, the rate is  $50 \times 1.01 = 50.50$  in 2001 and  $50.5 \times 1.01 = 51.005$  in 2002.

Rates that change at a constant percentage every year, change linearly on a log scale. For this reason, to estimate the APC for a series of data, the following regression model is used.

$$\log(Ry) = \beta_0 + \beta_1 y$$

where  $\log(Ry)$  is the natural log of the rate in year  $y$ .

$$\begin{aligned} \text{The APC from year } y \text{ to year } y+1 &= \frac{(R_{y+1} - R_y)}{R_y} \times 100 \\ &= \frac{[e^{\beta_0 + \beta_1(y+1)} - e^{\beta_0 + \beta_1 y}]}{e^{\beta_0 + \beta_1 y}} \times 100 = (e^{\beta_1} - 1) \times 100 \end{aligned}$$

**Population Estimation** The census populations of 2001 and 2011 were used in this report to calculate the estimates of population for the years 2012 and 2016 (Difference Distribution method for estimation of populations by five yearly age groups)