Cancer Atlas of India
Nandakumar et al., pp. 740–754

Documentation of cancer incidence in a country as vast as India is not a trivial undertaking. To date, the Indian National Cancer Registry Programme (NCRP) comprises only 6 population-based cancer registries (PBCR).

In an effort to collate cancer incidence from a broader sample, Nandakumar et al. collected data directly from 501 pathology laboratories serving the country's 593 districts. Data on 217,174 microscopically diagnosed cancers over the period 2001–2002 were submitted largely via the internet.

In order to determine how complete the data from various districts were, cancer incidence from the districts was compared to the incidence recorded by a past PBCR study. In all, 82 districts exceeded the PBCR incidence levels and thus were deemed to be representative.

The data generated confirmed some previous results but also gave new insights into national cancer incidence. These include a high rate of the following cancers: mouth cancer in southern states, nasopharynx cancer in the northeast, and stomach cancer in the Mizoram state.

The authors concede that districts with low incidence rates may simply reflect insufficient data collection and, thus, their cancer incidence is not interpretable. Additionally, it is acknowledged that about 15% of cancers are not diagnosed by pathology. However, collecting data directly from pathology laboratories compliments population-based data. It is also an efficient and cost-effective means of generating a cancer atlas that can offer new clues to cancer etiology.
Geographic pathology revisited: Development of an atlas of cancer in India

Ambakumar Nandakumar, Prakash Chandra Gupta, Palath Gangadharan, Rudrapatna Narayanrao Visweswara and Donald Maxwell Parkin

1National Cancer Registry Programme, Indian Council of Medical Research, Bangalore, India
2Indian Institute of Fundamental Research, Mumbai, India
3IKNP Centre, Vidyanagar, India
4Vaidhika Institute of Medical Sciences, Bangalore, India
5Descriptive Epidemiology Group, International Agency for Research on Cancer, Lyon, France

Information on 217,174 microscopically diagnosed cancers diagnosed in 2001–2002 was collected from pathology laboratories in 68 districts across India. Data collection took place primarily via the Internet. Average annual age-adjusted incidence rates for microscopically diagnosed cases (MAAR) by gender and site were calculated for each of the 68 districts in the country. The rates were compared to those from established population based cancer registries (PRCR). In 82 districts, the MAAR for all cancers was above a "completeness" threshold of 36.2 (100/100,000) based on results of a rural PBCR. The rates confirmed some known features of the geography of cancer in India, and brought to light some new ones. Cancers of the mouth and tongue are particularly frequent in both genders in the southern states. Very high rates of nasopharynx cancer were found in the northeastern states (Nagaland, Manipur). There was clear geographic correlation between the rates of cervical and penile cancer, and a high rate of stomach and lung cancer (in both genders) in many districts of Mizoram State. The area of high risk for gallbladder cancer seems larger than suspected previously, involving a wide band of northern India. There is a belt of high incidence of thyroid cancer in females in southwest coastal districts. Other than identifying possible existence of high-risk areas of specific cancers, our study has recognized places where PBCR could be established. The study was remarkably cost-effective and the electronic data-capture methodology provides a model for health informatics in the setting of a developing country.

© 2005 Wiley-Liss, Inc.

Key words: maps; incidence; cancer; India; district

Geographic pathology, comparisons of disease rates or risk of individuals living in different areas, has been one of the longest established and productive types of descriptive study for more than a century. The information can be nicely conveyed by the use of maps that, as well as conveying the actual value associated with a particular area, gives a sense of the overall geographic pattern of the mapped variable, and allows comparison between the patterns on different maps. This is especially valuable when used to suggest possible causative hypotheses. Atlases of cancer may draw attention to geographic variations within countries, or across wider geographic areas. The mortality statistics of China, the United States and the European Union are perhaps the best-known examples. There are fewer atlases of cancer incidence, because cancer registries generally cover more limited geographic areas than mortality statistics. Examples are the cancer atlases of Scotland, Canada and the Nordic countries.

India is a vast country, with a huge population of diverse cultures, habits and dietary practices. The socioeconomic status shows variation, as does the way of living and environment of the urban and rural populations. The study of geographic variations in cancer risk ought to be particularly fruitful in generating aetiological hypotheses. The available sources of information that can be used for this purpose, however, are limited. Systematic certification of death by qualified practitioners, that includes adequate information on cause of death, is confined to a few urban centres where in addition they are of variable quality. A national cancer registry programme (NCRP) was developed in 1982, under the auspices of the Indian Council of Medical Research (ICMR). Until 2003, this programme comprised only 6 population-based cancer registries (PBCR), covering a total population of 35.7 million (3.5% of the Indian total), including one registry serving a rural area.

The project to develop a cancer atlas of India was undertaken against this background, to try to make knowledge of the geographic patterns of cancer in the country. The registries under the NCRP have shown geographic differences in the patterns of cancer. For example, cancer of the gall bladder has a comparatively much higher incidence rate in the PBCR at Delhi and Bhopal, whereas cancer of the stomach has been a consistent leading site of cancer in males in the PBCR at Chennai and Bangalore. These observations have provided opportunities to hypothesise and look for various aetiological clues. Under the circumstances, coverage of more areas of the country through cancer mapping could open the doors for investigation of one or more cancers. The most readily available, comprehensive source of information on cancer occurrence nationwide is the network of pathology laboratories keeping adequate records on cancer cases diagnosed by cytology or histopathology. Although it is well known that many cancer cases are diagnosed without histology (15–20% of cases in the NCRP database), and that these are by no means a random sample of the total (by age and cancer site), these considerations were outweighed by the practical one of availability of information.

Material and methods

All medical colleges (listed by the Medical Council of India) in India and their respective departments of pathology, and several major hospitals were contacted. Letters were sent to 264 institutions. Those that agreed to participate were asked to begin collecting certain core items of patient information on all malignant neoplasms diagnosed between 1 January 2001–31 December 2002. On-site visits were made to these centres to examine the feasibility of obtaining accurate and complete information on place of residence of the patients reported, to check on computing and internet facilities and to assess the degree of technical support that would be required to successfully carry out the work. Several regional workshops were held for training purposes; they included principles of cancer registration, methodology of data collection, coding according to International Classification of Diseases (ICD) and an overview of web-site development and on-line transmission of data.

The methods for identifying cancer cases varied somewhat between pathology labs and the hospitals (cancer centres, government and private hospitals). The core information included the name of the patient, name of the father, mother, spouse and that of son and daughter (mainly to help in checking for duplicate registrations), age, sex ('gender'), address (at least name of the

*Correspondence to: National Cancer Registry Programme, No. 557, Sriniwasa Nikaya, New BEL Road, Dollars Colony, Bangalore 560004, India. Fax: 91 80 2351 2186; E-mail: arch@blc.vsnl.net.in
Received 27 September 2004; Accepted after revision 17 January 2005
DOI: 10.1002/j.2110
Published online 22 April 2005 in Wiley InterScience (www.interscience.wiley.com).
district of permanent residence), the primary site and morphology of the tumour and the date of diagnosis. Site and histology were coded according to the ICD-O-3.\textsuperscript{1,2} The Principal Investigator in each centre (usually a pathologist or radiotherapist) supervised and checked the diagnostic information and coding. Whenever the exact primary site of tumour was unknown, efforts were made to contact the treating clinician to obtain the details.

Data collection took place primarily via the Internet, with the 'core form' for recording collected data hosted on the website (www.canceratlasindia.org) (Fig. 1). Collaborating centres were given an individual login-ID and password and detailed instructions for data entry and transmission. All data were encrypted for transmission, to prevent unauthorized access. Centres without an Internet connection sent photocopies of completed forms to the Coordinating Centre in batches.

A number of internal validity checks were carried out, including range checks on codes, impossible or unlikely combinations of age, gender, site and histology\textsuperscript{16}. Additional checks included unlikely district distribution of cases from a given centre, and unlikely centre distribution of cases in a given district. Coding of district was checked against a list of addresses, and ICD coding was checked against a printed listing of pathology diagnosis by a qualified pathologist (AN). When needed, clarification was sought from individual centres. Cases where the identity of district (of the place of permanent residence of the patient) was unknown, or those where the date of diagnosis was outside the 2-year calendar period (1 January 2001–31 December 2002) were excluded. A variety of duplicate checks were also done and cases found to be duplicate registrations were excluded.
TABLE I – PROPORTION OF MICROSCOPIC VERIFICATION (MV) AND MICROSCOPIC AGE-ADJUSTED INCIDENCE RATE IN POPULATION BASED CANCER REGISTRIES

<table>
<thead>
<tr>
<th>Sites</th>
<th>Average MV% (Males &amp; Females)</th>
<th>Microscopic age-adjusted incidence rate (per 100,000)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Males</td>
<td>Females</td>
</tr>
<tr>
<td>Tongue (ICD 10: C01-02)</td>
<td>90.1</td>
<td></td>
</tr>
<tr>
<td>Mouth (ICD 10: C03-06)</td>
<td>89.3</td>
<td></td>
</tr>
<tr>
<td>Nasopharynx (ICD 10: C11)</td>
<td>71.7</td>
<td></td>
</tr>
<tr>
<td>Oesophagus (ICD 10: C15)</td>
<td>78.4</td>
<td></td>
</tr>
<tr>
<td>Stomach (ICD 10: C16)</td>
<td>73.9</td>
<td></td>
</tr>
<tr>
<td>Gall bladder (ICD 10: C23-24)</td>
<td>62.9</td>
<td></td>
</tr>
<tr>
<td>Lung (ICD 10: C33-34)</td>
<td>68.3</td>
<td></td>
</tr>
<tr>
<td>Breast (ICD 10: C50)</td>
<td>88.9</td>
<td></td>
</tr>
<tr>
<td>Cervix Uteri (ICD 10: C53)</td>
<td>89.1</td>
<td></td>
</tr>
<tr>
<td>Penis (ICD 10: C60)</td>
<td>87.8</td>
<td></td>
</tr>
<tr>
<td>Thyroid (ICD 10: C73)</td>
<td>85.7</td>
<td></td>
</tr>
<tr>
<td>All sites (ICD 10: C00-C96)</td>
<td>82.9</td>
<td></td>
</tr>
</tbody>
</table>

Figure 2 – Distribution of collaborating centres (●) and centres contacted but not responded (●) with names of States.

The final database included only malignant tumours (behaviour code ≥3) with a microscopic basis of diagnosis (histology or cytol- ogy [including peripheral blood film]). The ICD-O-3 codes were con- verted to ICD 10 codes for tabulation and display purposes.

Population data for district, and for the areas covered by 6 PBCRs in the NCRP were available by gender and 5-year age
group from the census of 1991, and by gender from the census of 2001. These data were used to estimate the mid-year populations by gender and age group for 2001–2002 for districts and 1997–1999 for the PBCR. Incidence (per 100,000 person-years) was calculated for each of the districts, and age standardized/adjusted rates calculated using the direct method and the world
The rates so calculated relate only to microscopically diagnosed cancers and, in addition, may be incomplete for a given area. These incidence rates are referred as microscopic crude (MCR) and microscopic age adjusted (MAAR) incidence rates.

The MAAR of the most recent data (1997-1999) from the 6 PBCRs of the NCORP provide a benchmark for analysis and a baseline for comparison with that of the districts. Table 1 shows the proportion of cases with microscopic verification (MV%) in the PBCR, and, for each, the MAAR for all sites combined and for selected sites of cancer in males and females. For most sites, the lowest MAAR are in the rural PBCR of Bhubaneswar, in Maharashtra state. As most of the 593 districts in the country are predominantly rural, the MAAR for all sites combined in males (36.2/100,000) for this registry was taken to represent a reasonable minimum expected value, if case finding in a district were fairly complete. Registration in districts where the MAAR was less than this was assumed to be incomplete.

Most of the centres contributed data of both the years 2001 and 2002. A few centres provided data for 1 year only. Accordingly, where the variation in MAAR of districts between each of the years was no >10% the average annual MAAR was taken. Where the variation in MAAR was >10% between each of the years, the higher MAAR was used.

Incidence rates (age-adjusted or age-standardized [AAR]) from Cancer Incidence in Five Continents for the period around 1995 [2] were used to compare with the MAAR observed in the Indian districts.

Results

The full results of the project have been published in 2 volumes as the development of an Atlas of Cancer in India. [2],[22]

One hundred and five centres, located in 68 different districts took part in the project (Fig. 2), recording data on 217,174 microscopically diagnosed cancers during the 2-year period (103,081 cases in 2001 and 114,093 cases in 2002). Age-adjusted incidence rates were calculated for all the 593 districts in the country, and in 82 districts, the MAAR exceeded the minimum “completeness” threshold of 36.2/100,000 in males (Fig. 3). For females, there were 10 districts in which the MAAR for “all cancer sites” exceeded the highest value in the Indian PBCR (Delhi, where the MAAR was 102.9 per 100,000), and 4 districts with female “all sites” rates higher than that in Delhi (113.9 per 100,000). The states with the
largest number of districts with rates exceeding the threshold were in the northeast (Mizoram, Manipur, Sikkim and Assam), and the south (Kerala, Karnataka, Tamil Nadu and Goa). There were also 5 or more districts from Maharashtra (including Goa) and Gujarat in the west and Punjab in the north. This distribution reflects, in part, that of the centres participating (Fig. 2), because these are likely to obtain relatively complete coverage of their local populations. The overlap is not exact. There were only 7 participating centres in the northeastern states, whereas some 20 districts of this region had above-threshold rates in men (6 in the state of Mizoram alone), and 17 districts in women.

The proportion of cases for which the district of residence was not known was <1% in the data from 94 centres, and <5% in 9 of the remaining 11. It was 6.2 and 9.4% in 2 centres in Jaipur. The cases analyzed (217,174) exclude those where the district of residence was unknown. The median percentage of cases from a given centre that were residents of the same district was 51.7% (with 48.3% from other districts). The percentage varies according to geographic factors (a centre may be at the periphery of the district and so receive most cases from elsewhere) or factors such as accessibility, treatment facilities available, and the availability of other cancer diagnostic or therapy centres in the same district.

Figures 4–12 show maps for selected cancers, to illustrate the geographic patterns shown by the atlas with the names of the districts in which the MAAR was higher than that recorded in any of the 6 population based cancer registries. In the text, attention is drawn to districts where such rates were based upon at least 10 cases. The rates recorded by the Indian PBCR are cited for comparison, and, when appropriate, other international data also.

Tongue (ICD-10: C01–C02) males

The highest MAAR was in Aizawl in Mizoram State (10.2 per 10^5), a rate rather higher than that in the PBCR of Bhopal (10.0 per 10^5), which was the highest in India (Fig. 4). Several other districts have similar high incidence rates, especially in the state of Gujarat in the west. The age standardized/adjusted rate (AAR) in Ahmedabad in this state was 9.3 per 10^5 in 1993–1997.

Mouth (ICD-10: C03–C06) females

The highest MAAR in the Indian PBCR was 6.4 per 10^5 (Bhopal) (Fig. 5). Several districts had rates in excess of this. The Kolar district (10.7 per 10^5) in Karnataka state was the highest, also Bangalore Rural and Kodagu districts of Karnataka state, Kollam and Thrissur districts in Kerala state, Villupuram district in Tamil Nadu state and Pondicherry district (under the Union Territory [UT] of Pondicherry).

Nasopharynx (ICD-10: C11) males

Cancer of the nasopharynx is generally considered to be uncommon in India (Fig. 6). The highest rate (AAR) in Cancer Incidence in Five Continents is 0.8 per 10^5 in Chennai (Madras), Kohima district in Nagaland state recorded a MAAR of 19.4/100,000, which is similar to the high rates found in southern Chinese
populations, (e.g., Hong Kong, 21.4 per 10^5) [2] Imphal West district in Manipur State also recorded a MAAR of 7.4 per 100,000. Several other districts in the states of Mizoram and Manipur recorded high MAAR in males and females, but the numbers of cancers were in single digits.

**Oesophagus (ICD-10: C15) males**

The incidence of oesophageal cancer in India is moderately high (Fig. 7). Among the PBDR, the highest rates are observed in Bhopal (MAAR 9.5 per 10^5) and Chennai (7.2 per 10^5). Several districts recorded MAAR >6 per 10^5, especially in the northeast and southwest.

**Stomach (ICD-10: C16) males**

Stomach cancer has appeared consistently as the most common cancer of men in the Chennai and Bangalore PBDR since the commencement of the NCRP in 1982. Although the recorded rates (MAAR of 9-13 per 10^5) are considerably lower than in known high risk parts of the world risk areas (AAR in males >30 per 10^5), such as East Asia (China, Japan, Korea, Mongolia), Central Asia, Eastern Europe and parts of Central and South America (Costa Rica, Colombia, Chile). The Atlas has shown that, in addition to the raised incidence in districts in southern India (Karnataka, Tamil Nadu and Kerala), there is a focus of high MAAR in the northeast of the country, with the highest values in the districts of Serchhip (70.2 per 10^5), Kolath (56.1 per 10^5), Mamt (53.4 per 10^5), Aizawl (47.0 per 10^5) and Champhai (46.3 per 10^5) in Mizoram State (Fig. 8).

**Gall bladder (ICD-10: C23-24) females**

The incidence of gall bladder cancer in females in Delhi (AAR: 9.4 per 10^5; MAAR: 7.2 per 10^5) is the highest in Cancer Incidence in Five Continents. [2] The districts of Imphal East (MAAR: 6.9 per 10^5) and West (MAAR: 5.7 per 10^5) in Mizoram State and the Union Territory of Chandigarh (MAAR: 6.4 per 10^5) had rates similar to those in Delhi. Figure 9 suggests that the zone of high incidence involves several northern states (Punjab, Uttar Pradesh, Bihar, West Bengal), and this is all the more remarkable in that the overall incidence (all sites combined) in these areas is low (Fig. 3), presumably because registration was very incomplete.

**Cervix uteri (ICD-10: C53)**

Incidence of cancer of the cervix in India is high, with the highest rates observed in the south (Chennai PBDR has the highest incidence of cervical cancer among the Indian PBDR: AAR 30.6 per 10^5; MAAR 27.2 per 10^5) (Fig. 10). The map suggests that there is a belt of high incidence in Tamil Nadu State (Villupuram,
31.1 per 10^6; Cuddalore, 29.9 per 10^6; Thiruvallur, 28.6 per 10^6), adjoining Pondicherry (39.2 per 10^6) and Karnataka.

**Penis (ICD-10: C60)**

The incidence of cancer of the penis in the Indian PBCR is moderately high (Chennai and Barshi with AAR of 1.6 per 10^6), although this is lower than the highest rate in Cancer Incidence in Five Continents of 4.0 per 10^6 in Kyadondo in Uganda (Fig. 11). As for cervix cancer, there is a concentration of high incidence in the northeast of Tamil Nadu state; Villupuram district (MAAR 3.1 per 10^6), Cuddalore (2.2 per 10^6), Thiruvallur (2.1 per 10^6), Erode (2.0 per 10^6) and Kancheepuram (1.9 per 10^6).

The rank correlation coefficient of MAAR between cancer of the cervix in females and cancer of the penis in males in the 82 districts was significant at the 1% level (Pearson's coefficient = 0.473; Spearman's coefficient = 0.644).

**Thyroid (ICD-10: C73)**

There are several districts with relatively high incidence of thyroid cancer in the northeast (Fig. 12). There also seems to be an area of elevated incidence along the coast in the southwest of the country.

**Primary site unknown**

The proportion of cases with primary site unknown relative to all sites was <10% in 75 of 82 districts. Of the remaining seven districts Jaipur and Chandigarh districts had a high relative proportion of 11.4 and 13.5%, respectively.

**Discussion**

Cancer atlases have been particularly useful in highlighting variations in the risk of different cancers according to place of residence, between countries, and also within them. There are examples from the United States, Scotland and China. An atlas covering a wide area requires uniform data collection, and quality, from the geographic units concerned. Generally, this is in the form of mortality statistics deriving from death certification or cancer incidence data, from national cancer registration schemes. These options are generally not available in most developing countries, where deaths' certification is unreliable or uneven in quality, and cancer registries are confined to selected (usually urban) areas. The approach in our study was to use a widely available resource, the pathology laboratory, to try to extend the coverage of incidence data. Although well aware of the defects of pathology, based registration, both in
CANCER ATLAS OF INDIA

Figure 7 - Districtwise microscopic age adjusted incidence rates (MAAR) per 100,000. Oesophagus (ICD 10: C15) males. Year 2001–2002.

Comparisons with the rates for cancers diagnosed microscopically from the Indian population-based registries make it clear that for most districts in the country, registration (even of cases diagnosed via pathology) must have been very incomplete. This is not surprising. The estimates of new cancers in India (for the year 2000) show some variation among investigators ranging from 566,000, 614,000, 314,000, and 650,000. Based on 217,174 cases (for the 2-year period) included in our study this would constitute between 13–21% of the total. These estimates, based mainly on the NCRP data, are essentially from urban centres and only one rural registry that covers part of a district.

Even within the specific local area of the collaborating centres, only some centres made an active effort to collect information on microscopically diagnosed cancers from the other, non-participating, hospitals or laboratories. There were 82 districts in which the incidence rates exceeded those in the rural cancer registry of Barshi, suggesting that data collection was relatively complete in these districts.

No special attempt was made to enrol centres located in the districts served by population-based cancer registries (PBCRs), so as not to disturb their data collection procedures. It would be interesting to compare the validity of the data collected through the cancer atlas project with that available via the PBCR, especially in this developing country setting. It should be noted, however, that the data collection methods in the districts served by PBCR were not totally independent of those used by the latter.

Completeness (about 85% of all cancer cases in the areas covered by population-based registries in India), and in the selection bias introduced, it was felt that the results would be interpretable if judged against the standards set by the population based cancer registry network in India.

National pathology registers have operated in a number of countries where, for one reason or another, it had proved impossible to develop population based registries (Luxembourg, Greece, South Africa, and Brazil). In Brazil, however, the rapid expansion of the network of registries has made this exercise of less interest than in India, where, at present, only some 3.5% of the population is covered by population-based statistics.

Collection of data via the Internet greatly facilitated the speed of communication of data, as well as ensuring uniform procedures for validation of collected information. The successful implementation of this concept is reflected in 3 ways. The first way includes the data that was downloaded on a regular basis for the past 2.5 years. The core data of approximately 1,200–1,500 cancer cases was received every week. The second way is in the feedback received from the participating centres. Ninety-five percent of the respondents felt that the website was easy to use and 80% of them had a fairly stable Internet connection. Third, because most of the collaborating centres were able to transmit the required information as soon as the diagnosis was made, the results for 2001–2002 were already available in early 2004, a very quick schedule by cancer registry standards.
The choice of the Barshi male rate of 36 as a standard is entirely arbitrary. Although there are cancer registries in 2 other rural areas of India (Karanagapally, Kerala State and Ambikklak, Tamil Nadu State), these were set up as part of specific research projects and the populations served are not at all typical of other rural areas in population structure, education and in several health parameters. The use of a rate other than 36, and concentration on districts with rates superior to it would not make any major difference to the conclusions to be drawn from the results of the project.

Low incidence rates in our study are not interpretable, because they may simply represent incomplete coverage of the area. In some states such as Bihar, Uttar Pradesh, Jharkhand and Chhattisgarh, the few districts with “high” rates is obviously due to the very few participating centers in these states (Fig. 1). Even in areas where the overall coverage was high, low incidence of a particular site could be partly due to a biased referral pattern and therefore not interpretable as reflective of true incidence rate.

Presentation of the results tends to focus on these 82 districts, or on those where the “Microscopic age-adjusted rate” (MAAR) was relatively high (compared to the results from the PBCR), knowing that no useful interpretation can be put on low estimated rates. For some districts (especially in the northeastern states), recorded incidence rates were high; indeed the MAAR exceeded the rates based on all diagnosed methods in the Indian PBCR. This must give rise to the suspicion of over-recording of new cancer cases, possibly due to inclusion of non-residents, to duplicate entries, or to inclusion of “prevalent” cases, first diagnosed before the study period, but returning for cytology or tissue biopsy within it. As described in Material and Methods, considerable care was taken to ensure that the place of residence was properly recorded, and the data file from each centre was carefully scrutinized for duplicate registrations, using printed lists (by centre and district), sorted by name, address, hospital and laboratory number and other identifying information. All cancer registries record a relatively high number of cancers during the early years of operation, because cases diagnosed in the preceding years return for follow-up. “Prevalent” registrations should be less in this project, where the pathologists should have been aware when a particular biopsy was a repeat, or carried out during follow-up, or for a recurrence.

The Atlas has confirmed some suspected features of the geography of cancer in India, as well as brought to light some new or little known facts.

The relative proportions and incidence rates of cancers of sites associated with the use of tobacco show variations for specific sites across different geographic areas. Cancers of the mouth and tongue seem to be particularly frequent in males and females in
the southern states. Cancers of the oropharynx are conspicuous along the West coast, cancers of the hypopharynx noticeable in the northeast, cancers of the pharynx striking in the state of Gujarat and the northeast states. This may be attributed to the diverse types of tobacco products and the manner in which they are consumed in different populations in India. 

Smokeless tobacco use in various forms, with or without other ingredients is a common practice in India. Smoking of bidis is more common than cigarette smoking, especially among the rural population. Women tend to chew (smokeless tobacco) rather than smoke, although this could be different in the northeast state of Mizoram.

The high incidence of nasopharyngeal cancer (NPC) in the northeastern states is quite clear. A high relative frequency has been reported in previous studies from the region, as well as appearing in data from the hospital-based cancer registry (HBCR) in Dehradun, Assam. It is clear from our results that the actual rate of incidence may be comparable with the high-risk populations of southern China with much more than that reported in Southeast Asia. NPC is generally linked to 3 anthropological factors: infection with EBV, dietary exposures to certain fermented foods and a strong genetic component. These have been studied in a limited way in the populations of the northeast. Chelleng et al. have reported a higher risk of NPC in those who consumed smoked meat and also in those who used herbal nasal medicine, and Kumar et al. found that positivity for anti-EBV IgG antibody was higher in NPC patients than “control” patients. The above factors notwithstanding, the probable genetic similarities between the inhabitants of the northeastern states, and populations in China and southeast Asia seems to offer a further clue to the etiology of NPC.

Stomach cancer is relatively rare in India. This is certainly not due to a low prevalence of infection with Helicobacter pylori, which has, in fact, been elevated in Indian populations. Quite possibly, some aspects of the Indian diet are protective, although the possible adverse effects of eating hot and spicy foods have been studied by several authors. The relatively high incidence of gastric cancer in northeastern India is a new finding. This may relate to distinct features of the diet such as smoked fish and meat that is consumed in these areas, and is another finding that would require specific study.

High incidence rates of gall bladder cancer have been reported from Delhi and Kolkata (Calcutta), and there are reports from centres in other northern states drawing attention to a relatively high frequency of gall bladder among clinical and pathological cases series. Our study suggests that there is a very wide
FIGURE 10 – Districtwise microscopic age adjusted incidence rates (MAAR) per 100,000, Cervix uteri (ICD 10: C53). Year 2001–2002.

genealogical band of high-risk, extending from the Punjab in the west, through to West Bengal in the east and on into the northeast. The aetiology of gall bladder cancer has been closely linked to that of gallstones, and these are also frequent in populations living in northern India. Gallbladder cancer is a common cancer in Indian migrants in the United Kingdom who originate predominately from northern states. Clearly, the reasons for this striking geographic pattern, be they genetic or environmental (e.g., diet), require elucidation.

A particularly high incidence of both cervix cancer in women and penile cancer in men was seen in the northeastern districts in Tamil Nadu state and the adjoining Union Territory of Pondicherry. In cancer registry data from India reported in the Cancer Incidence in Five Continents series, Chennai (Madras) has reported consistently the highest rates in the country, and some of the highest in the world. Furthermore, the rural cancer registry of Ambililiki in the south of the state recorded a very high incidence of 65.4 per 10^5 in 1996-1998. It is not really known why this area seems to be at high risk for this cancer; although a recent population-based survey found a high prevalence of infection with human papilloma virus (HPV) in Ambiliiki. The geographic co-incidence of cervical and penile cancer has been noted many times before, and relates presumably to a common aetiology involving HPV infection. Tamil Nadu state also has a relatively high prevalence of HIV infection.

Cancer of the thyroid in women showed relatively higher incidence in the northeastern states, but a further noteworthy feature of the geography is the belt of comparatively high incidence commencing from the southern tip of the country and extending upward along the southwest coast to Goa state. Thyroid cancer has been noted to be common in other coastal areas of the world with large populations of fishermen, but there is no evidence that thyroid cancer risk is elevated among people, male or female, with a relatively high intake of iodine from fish and seafood. Follicular carcinomas may be related to iodine deficiency; they are more common in populations living in areas where endemic goitre is common. This is the case in the northeast of India, where, despite a programme of salt iodization, the diet of local populations continues to be low in iodine and endemic goitre remains frequent. The southwest coast of the country is known for its high level of natural background radiation. Whether the occurrence of a relatively higher incidence of thyroid cancer has any relevance to such radiation needs to be investigated, although
radiation from medical sources, and in atomic bomb survivors is known to increase thyroid cancer risk.63

Conclusion
In conclusion, the Atlas project has, as we describe, provided clues as to the possible existence of areas of high risk for specific cancers, and in this respect, provides a valuable springboard for future epidemiological studies. It identified places or centres where population-based cancer registries could be established with minimal effort, and 6 new registries in the northeastern states, covering 0.6% of the population of India, have been started within the NCRP since January 2003. It was also remarkably cost-effective. The amount spent on cancer care under the population-based cancer registries of the NCRP is, on average, Rs. 350 (approximately equivalent $8.1 US) in urban areas and Rs. 200 ($4.70 US) in rural registries (unpublished data from NCRP). The cost per case recorded in our study was Rs. 24 ($0.5 US). The low cost is in part due to the use of data already recorded and stored in departments of pathology, and in part due to the use of electronic data capture methods. We believe that the methodology could pave the way for broad usage and opening the field of health informatics and that this concept can easily be transferred to estimate burden and assess cancer patterns in the setting of a developing country and possibly for certain other diseases of public health importance.

Acknowledgements
The project and the workshops were funded by the WHO-India Office, Ministry of Health, Government of India and their support is gratefully acknowledged. The Director General of the Indian Council of Medical Research (ICMR), Professor N.K. Gargal, is a constant source of encouragement and his support greatly helped in the success of the project. Dr. B. Shah, Chief of the Division of Non-Communicable Diseases of ICMR has been very supportive and so have all the members of the Steering and Monitoring Committee of the NCRP. Dr. R. Renganathan, then Chief of the Indo-Foreign Cell of ICMR was primarily responsible in making the crucial efforts at the Ministry of Health and the WHO to get the project cleared. Gratitude is also expressed to the Registrar General and Census Commissioner of India for providing the populations of the districts and the district-wise and state-wise country maps. Dr. C. Varghese (WHO-India) and Dr. J. Leowski
(WHO-SERCO) participated in the workshops that assisted in generating enthusiasm in the study. The Principal Investigators, Co-Principal Investigators and other concerned faculty of the collaborating centres (names listed below) have contributed overwhelmingly toward the success of the project. Their active participation in the workshops and their motivation and tireless efforts in sending data promptly has been of great help in bringing out a timely report. The authors record their appreciation. The administrative and secretarial staff of NCRP have amply facilitated in the preparation of the manuscript.

References
