RESEARCH ARTICLE

Cancer: Scenario and Relationship of Different Geographical Areas of the Globe with Special Reference to North East-India

Jagannath Dev Sharma[1], Manoj Kalit[1]*, Tulika Nirmolia[2], Sidhartha Protim Saikia[2], Arpita Sharma[1], Debanjana Barman[1]

Abstract

Background: Cancer is becoming the most important public health burden around the globe. As per the GLOBOCAN 2008 estimates, about 12.7 million cancer cases and 7.6 million cancer deaths were estimated to have occurred in 2008. The burden of cancer cases for India in the year 2020 is calculated to be 1,148,757 (male 534,353; female 614,404) compared to 979,786 in 2010. The pattern of cancer incidence is varying among geographical regions, esophageal cancer for example being high in China, lung cancer in USA, and gallbladder cancer in Chile. The question remains why? Is it due to the diversity in genome pool, food habits, risk factor association and role of genetic susceptibility or some other factors associated with it? In India, the North East (NE)-India region is seeing a marked increase in cancer incidence and deaths, with a very different cancer incidence pattern compared to mainland India. The genome pool of the region is also quite distinct from the rest of India. Northeastern tribes are quite distinct from other groups; they are more closely related to East Asians than to other Indians. In this paper an attempt was made to see whether there is any similarity among the pattern of cancer incidence cases for different sites of NE-India region to South or East-Asia. Materials and Methods: Principal Component Analysis (PCA), Hierarchical Cluster Analysis (HCA), Pearson Correlation coefficient test was assessed to evaluate the linkage of North-East India region to other regions. A p value <0.05 was considered as statistically significant. Results: The results clearly shows that there are similarities in occurrence of cancer incidence patterns for various cancer sites of NE-India with South and East-Asian regions, which may lead to the conclusion that there might be a genetic linkage between these regions.

Keywords: Cancer - oesophagus - lung - stomach - hypopharynx - nasopharynx - TRC

Asian Pac J Cancer Prev, **15 (8)**, 3721-3729

Introduction

Neoplasia (Greek for "new growth") is the abnormal and uncontrolled proliferation of cells in a tissue or organ (Kalssen et al., 2003). The term mutation is the abnormal change in gene which is of two types, one is inherited and other is acquired (somatic) (wwwcancerorg/cancer/cancercauses/geneticsandcancer/heredity-and-cancer). Cancer is becoming the devastating disease of all at present now. It is the most important public health burden around the Globe. Global burden of Cancer incidence is continuously increasing and increasing. As per the GLOBOCAN 2008 estimates, about 12.7 million cancer cases and 7.6 million cancer deaths are estimated to have occurred in 2008 (Jemal et al., 2011). According to World cancer Report 2008, by the year 2030 the global burden of cancer will increase to 26.0 million with 17.0 million deaths (with 1% annual increase in rates compared to 2008).

The most common cancers in the world in term of incidence were lung (1.52 million cases), breast (1.29 million) and colorectal (1.15 million). Lung cancer was the most common cause of death (1.31 million), followed by stomach cancer (780,000 deaths) and liver cancer (699,000 deaths) (World Cancer Report, 2008). One in 4 deaths in the United States is due to cancer (Siegel et al., 2013). As Estimated the burden of cancer cases for India in the year 2020 will be 1148757 (male 534353; Female 614404) compared to 979786 in 2010 (Takiar et al., 2011). In India, the International Agency for Research on Cancer estimated indirectly that about 635 000 people died from cancer in 2008, representing about 8% of all estimated global cancer deaths and about 6% of all deaths in India (Dikshit et al., 2012). In India carcinoma of lung was responsible for maximum death in men (Ferlay et al., 2007).

The pattern of cancer incidence is varying among geographical regions (Ferlay et al., 2010). Like esophagus is high in China, Zhongshan for both male and female (AAR-26.9, 10.1) followed by China Guangzhou City (AAR-22.2, 9.8), Gallbladder in Chile, Valdivia (AAR-12.3, 27.3), Lung in USA, USA Louis., New

*[1]Population Based Cancer Registry, Indian Council of Medical Research, Dr. B. Borooah Cancer Institute, Guwahati, [2]Department of Biotechnology, Assam Medical College and Hospital, Dibrugarh, Assam, India *For correspondence: manojkalita5354@gmail.com*

Orleans: Black, Louisiana: Black in Male (AAR 96.6, 91.7) and in Female USA, Kentucky (AAR 50.3), USA Pennsylvania: Black (AAR 46.8) while Stomach was high in Japan Hiroshoma (AAR-80.3, 30.2), followed by Japan Yamagata Prefecture (AAR-79.4, 31.3) as per IARC, Cancer Incidence in Five Continents Vol. IX report (Ferlay et al., 2007). The question remains why? Was it due to the diversity in genome pool, food habits, risk factor association and role of genetic susceptibility or some other factors associated with it? For every cancer, there seems to be at least one inherited form, i.e. All Cancers exist in both heredity and non heredity form (Alfred et al., 1985). As per American Cancer Society, of all Cancers 5% to 10% Cancers were inherited from their parents (wwwcancerorg/cancer/cancercauses/geneticsandcancer/heredity-and-cancer). Demographic, ecological, environmental, cultural, and genetic variables all contribute to the heterogeneity of cancer incidence (Wang et al., 2012).

In this paper an attempt was make to observed whether really incidence pattern is different for different geographical region based on Age Adjusted Rate (AAR/ASR) data's and also tried to evaluate whether genomic linkage is really available or not based on GSTM1 and GSTT1 available dat.

In India cancer incidence and mortality data were collected by Population Based Cancer Registries (PBCRs) and were estimated and reported in terms of CR (Crude rate), AAR (Age Adjusted rates) and AAMR (Age Adjusted Mortality Rate) by National Cancer Registry Programme (NCRP), Indian Council of Medical Research (ICMR). Recently NCRP has published its Three Year Report of Population Based Cancer Registries 2009-2011, (NCRP, 2013). As per the report the highest AAR and AAMR for all cancer sites per 100, 000 inhabitant of the Indian region was observed in NE region States. These rates were age-standardized or Age Adjusted rate (ASR or AAR) (per 100,000 person-years) using the World Standard Population as proposed by Segi and modified by Doll et al. (1966), (Segi 1960 and Doll et al., 1966).

Out of 10 high incidence regions in the country, 8 in male and 5 in female were observed from NE-State compared to mainland India. Aizwal District of Mizoram shows a very high incidence (AAR 273.4, 227.8) and mortality (AAMR 154.1, 110.1) for both sex male and female followed by Mizoram State (AAR 189.5, 153.7) and (AAMR 110.3, 76.5) with other NE-States compared to mainland India (Table 1) (NCRP, 2013). In global scenario as it compared to other parts of the world as per cancer Incidence in Five continents VOL IX recoded data for incidence (AAR/ASR), observed that the in Male Aizwal District, East Khasi Hills, Mizoram State, Kamrup Urban Districts, Meghalaya secures 11th, 12th, 14th, 15th and 17th position respectively (Figure 1). Whereas in Female Aizwal District, Kamrup Urban District, Mizoram State, Bangalore and Chennai occupies 11th, 14th, 15th, 16th, 17th number position respectively as based on AAR of all sites around the globe (Figure 2).

North East India (NE-India) is the eastern-most region of India. NE-India formerly comprises of seven states commonly known as the "Seven Sisters". They

Table 1. Pattern of Cancer Scenario in India

Regions	Incident rate (AAR/ASR)		Childhood incident rate AARPM (0-14)		TRC proportion to all cancer site		Death rate (AAMR)	
	Male	Female	Boys	Girls	Male%	Female%	Male	Female
Aizwal District**	273.4	227.8	107.8	59.4	47.4	22.4	154.1	110.1
East Khasi Hills**	216	114.1	24.3	11.3	69.3	43	75.4	34.2
Mizoram State (MZ)**	189.5	153.7	64.5	45.9	41.6	23.2	110.3	76.5
Kamrup Urban District**	185.2	156.3	62.2	45.7	49.1	25.7	57.3	32.6
Meghalaya**	157.3	83.7	18.6	11.9	66.1	40	51.5	25.6
MZ-Excl. Aizwal**	145.6	112.1	45.9	40.1	35.6	24.1	87.7	57.8
Thiruvanthapuram	132.6	123.2	97.6	103.3	35.6	10.3	52.1	34.2
Cachar District**	129	98	36.9	25	43.6	22.9	17.3	6.7
Nagaland**	126.1	70.2	36.9	14.7	35.8	10.9	NA	NA
Delhi	125.2	120.6	149.4	77.7	39.6	11.6	16.7	12.9
Kollam	118.5	91.6	121	77.6	41.2	12.7	60.8	33.6
Chennai	118	123.8	159.6	112.4	38.6	14.1	34.2	28
Ahmedabad Urban	117.5	87.1	83	50.6	53.8	16.6	30.3	18.1
Banglore	113.7	137.2	91.6	69.9	35	15.2	42.4	37.3
Bhopal	105.9	105.6	70.7	61.9	50.4	16.9	40.3	32.8
Dibrugarh District**	99.4	71.8	38.8	19.9	53.2	23.7	31.2	16.8
Mumbai	98.4	105.5	89	59.2	38.6	13.9	48.3	43.7
Nagpur	96.4	103	117.4	84.6	43.6	14.3	19.3	16.4
Imphal West District**	94.3	90.3	57.2	69.3	38.2	19.4	17.2	16
Kolkata	92.8	99.4	111.1	67.7	44.7	14.1	43.6	39.6
Sikkim State**	82.6	94.2	46	41.9	31.4	20.5	38.1	39.5
Tripura State**	78.8	56.1	45.4	46.5	55.3	20.5	30.4	18.7
Manipur State (MR) **	74.7	73.9	45.6	30.8	36.4	21.1	14.6	10.6
Pune	74.3	75.7	47.8	35.3	39.4	16.2	31.6	29.8
Ahmedabad-Rural	74.2	51.6	60.7	51	56.2	19	29.3	18.5
Mr Excl. Imphal West**	69.5	68.9	43.6	23.5	35.7	21.8	13.9	9
Aurangabad	59.6	62.1	37.6	40.7	49.2	15.7	14.5	9.5
Wardha	57.8	64	40.2	49.5	45.5	22.1	27.7	24.9
Barshi Roral	51.8	62.6	72.8	38.7	40.3	12.9	38.2	39.4
Barshi Expanded	43.7	56.6	32.9	18	37.8	13.4	15	14.7

*NA=Data not available; **indicates NE-India region

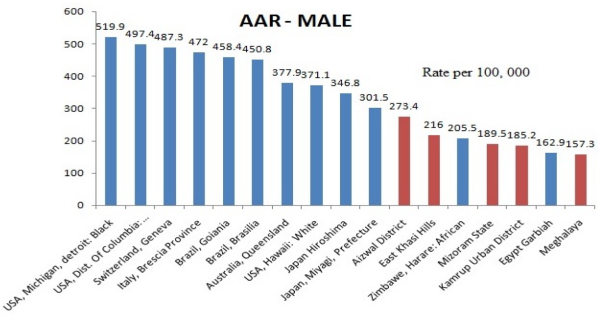

Figure 1. Comparison of AAR (Male)-India with Global Scenario

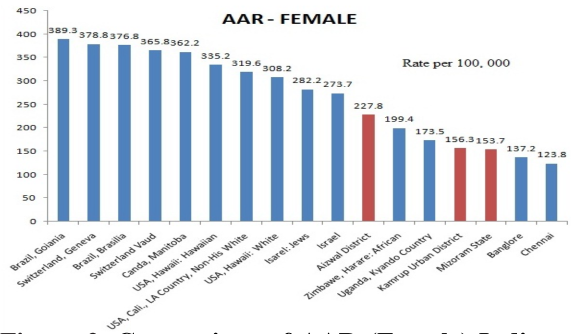

Figure 2. Comparison of AAR (Female)-India with Global Scenario

are MANIPUR, MIZORAM, ASSAM, MEGHALAYA, TRIPURA, NAGALAND, ARUNACHAL and SIKKIM is being included later on, which makes eight states now. The NE-India region is connected with mainland India through a narrow corridor called as 'chicken's neck' which constitutes barely one per cent of the boundaries of the region, while the remaining over 99 per cent of its borders are international-with Nepal, Bhutan, China, Thailand, Myanmar and Bangladesh. The region has over 160 Scheduled Tribes and over 400 other sub-tribal communities and groups. The genome pool of the regions is quite distinct from rest of mainland India. Within India, northeastern (NE-India) tribes are quite distinct from other groups; they are more closely related to East Asians than to other Indians (Cordaux et al., 2003). Northeastern region is distinguished by a preponderance of the Tibeto Burman languages, and the population here is thought to comprise migrating peoples from East and Southeast Asia (Kataki et al., 2011).

The NE-India population registries cover the area were MANIPUR [Manipur State (MR), Imphal West District, MR Excl. Imphal West], MIZORAM [Mizoram State (MZ), Aizwal District, MZ Excl. Aizwal], ASSAM [Dibrugarh District, Kamrup Urban District, Cachar], NAGALAND [Nagaland], MEGHALAYA [Meghalaya, East Khasi Hills], TRIPURA [Tripura State] and SIKKIM [Sikkim State].

NE-India is becoming the hub of Cancer incidence and deaths. A huge boom in cancer incidence cases was observed in these regions of India. NE-States of India is showing a very peculiar type of cancer incidence pattern compared to mainland India. Most of the N.E. regions grab the top most positions for various cancer sites for both sex (male and female) in India and around the Globe in terms of incidence and mortality. Oesophagus, Stomach, Lung, Hypopharynx, Nasopharynx, Larynx, Tounge, Gallbladder, Thyroid were some of the top most leading cancers sites which were at most rise compared to other parts of the globe, which is a matter of concern. It may be due to genetic, life style, food habits or other factor associated with it.

North East India follows a totally different kind of food and Dietary habits compared to mainland India. Some of the food habits (including tobacco and alcohol) were like Smoked meat, Smoked fish, Dry meat, Dry fish, Banana Plants (in Assam Called as Posola) and Banana Flower (in Assam Called as Koldil) used heavily all over NE-India whereas in Assam it is consumed from some selected banana trees but except Assam all other states consumed all types of banana trees and flower including wild ones from the forest side, Kalakhar (mostly in Assam-a product made from banana), Bamboo shoots (Consumed hugely in all parts of NE), wild strong flavor of exotic herbs (all over NE), fermented rice beer and millet beer consumed heavily all over the NE, most of the tribes in NE have a tradition to used fermented rice beer for all their ethnic occasions-like Opo or Apong, Sai Mod in Arunachal Pradesh), Haj and Sai Mod in Assam. Tuibur (water filtered smokeless), Meiziol (local cigarrate) used in Mizoram, Tamul (raw betel-nut) and Paan (betel leaf) combined or alone consumed heavily all over the NE. Some of the selected sites were discussed below which were severely high in these regions.

In North East India region incidence of eesophagus cancer was much higher than any other parts of the world for both male and female, AAR of Oesophagus cancer for male in East Khasi Hills was observed as (71.4), followed by entire state of Meghalaya (46.2), Aizwal District (42.0), Kamrup urban District of Assam (27.0) and entire Mizoram state (26.0) while for female it was East Khasi Hills (30.2), Meghalaya (19.8), Kamrup Urban District of Assam (18.3). Esophageal cancer ranks as the sixth most common cancer among males and ninth most common cancer among females globally (Kumar et al., 2006). In East Khasi Hills it was observed that 323 (34.2%) out of total 944 in male cases and in female 140 (24.9%) out of total 563 female cancer cases were Oesophagus cancer cases (NCRP, 2013). Oesophagus cancer was rise in the peak at these NE-India parts of the world. A global intervention was needed therefore to find out the actual cause of high incidence for this NE region of India. Kalakhar (mostly used in Assam) is a high alkaline substance made from the charred false stem or from the skin of a particular variety of banana that is used as coffee decoction or during the preparation of curry or "dal". Daily use of kalakhar greatly increases risk (OR=8.0) of Oesophegal cancer (Phukan et al., 2001). Difference in the incidence of esophageal cancer was observed between developed and developing countries and a 50-fold difference have been observed between high- and low-risk populations (Ishan et al., 2001).

After oesophagus cancer, stomach cancer is the second top most leading sites for NE Region of India with a high incidence rate. Stomach cancer is the second leading cause of cancer death and forth most common malignancy in the world. Globally a total of 989,600 new stomach cancer cases and 738,000 deaths are estimated to have occurred in 2008, accounting for 8% of the total cases and 10% of total deaths (Jemal et al., 2011). The state of Mizoram is a high risk region of stomach cancer in India (Rao et al., 1998; Phukan et al., 2011). In Aizwal district of Mizoram, 64 males and 31 female's per 100,000 inhabitants of that region were diagnosed with Stomach cancer, whereas entire Mizoram State with [AAR 47.6 (male), 22.7 (female)] and Mizoram Excluding Aizwal [AAR 38.9 (male), 18.1 (female)]. In India 8 out of top 10 regions for both sexes (male and female) as per incidence cases of Stomach cancer were from NE-India region. Globally incidence of Stomach cancer was very high in Japan followed by Mizoram for both male and female. Why so high incidence of Stomach cancer in this region of NE-region compared to national and international figure was a matter to look after. It may be due to the some of the popular ethnic food habits like consuming smoked meat or fish and high rate of consumption of alcohol and smoking habits. Meiziol (local cigarette) and consumption of Tuibur (smokeless tobacco) is popular among inhabitants of Mizoram. Tuibur is the popular locally made water filtrated smokeless tobacco which is even sold in the open market. Studies have suggest that Combination of Smoking, Meiziol and Tuibur consumption habit is found to be highly associated with lung cancer in the region of Mizoram and greatly increased the risk (OR=23.02) (Malakar et al., 2013).

Nasopharynx is another cancer site which is much higher in NE- State compared to mainland India and also leading sites globally after Chinese registries, numerically third for male and female both. Nasopharyngeal carcinoma is a rare cancer worldwide except in South East Asia, Southern China (Sharma et al., 2012). For Nasopharynx cancer 10 out of 10 leading cancer regions in India were from NE states for both male and female. This is really a serious matter, which should be look with a great deal. Northeastern region is distinguished by a preponderance of the Tibeto Burman languages, and the population here is thought to comprise migrating peoples from East and Southeast Asia, who are presumed to have brought with them the risk for nasopharyngeal carcinoma to this region (Kataki et al., 2011).

In Male Nagaland with AAR (21.0), followed by Aizwal (6.7) and in female it was Aizwal (5.2), followed by Nagaland (4.6) (NCRP 2013). While it is observed rare among mainland Indian region, Except in Chennai (Male AAR 1.1) all mainland Indian regions for both male and female incidence was observed as (AAR<1) which shows that Nagaland is 21 times higher incidence in male then compared to national scenario and 10 times in Female compared to Thiruvananthapuram (Female AAR 0.4) and mainland Indian region. It may be due to consumption of smoked meat or fish, salted fish, Wild herbs used, household burning of firewood's and other environmental factors. Few studies already had shown the carcinogenicity of smoked meat which contains nitrosamines like in salted fish. A case control study in nasopharyngeal carcinoma was studied in Nagaland reveals an association of nasopharyngeal carcinoma with consumption of smoked meat. The use of herbal nasal medicine seems to be an additional risk factor for nasopharyngeal carcinoma in Nagaland. Consumption of smoked meat was found to be the risk factor for nasopharyngeal carcinoma (adjusted odds ratio=10.8; 95% CI 3.0-39.0). History of using herbal nasal medicine was also found to be associated with nasopharyngeal carcinoma (OR=21.9, CI=6.8-71.4) (Chelleng et al., 2000).

The incidence of hypopharynx cancer is also increasing in males and females of NE States region. In India it was estimated that a total of 19397 Hypopharynx cases was reported in 2010 which will go up to 22568 by the year 2020 (Takiar et al., 2010). Hypo pharynx was mostly high in East Khasi Hills (21.5) district of Meghalaya, followed by entire Meghalaya (17.4) state, Aizwal District (15.4), Kamrup Urban District (14.7) and Dibrugarh (11.7) of Assam in Male, While in Female it was found high in the state of Assam in Kamrup Urban District (3.6) followed by Cachar District (2.6), Dibrugarh District (2.0) and in Meghalaya, East Khasi Hills (2.5), Entire Meghalaya (1.9) of NE Region(NCRP 2013). Tobacco and Dietry habits of these region may be linked up with high incidence of hypopharyengal carcinoma (Sharma et al., 2013).

Globally lung cancer was among top three leading cancers in both male and female (Hussain et al., 2012). Lung cancer is mainly a disease of modern era and probably one of the most important health problems today globally as well as for NE-India region (Mandal et al., 2013). Mandal's study also suggested that the gender gap has been narrowed such that about half of the patients diagnosed with lung cancer are women in this part of India. According GLOBOCAN 2008 report, in India a total of 47,010 new lung cancer cases among males and 11,557 new among females were reported (Ramshankar and Krishnamurthy, 2013). Incidence of Lung Cancer was also found severely higher in Aizwal District of Mizoram both in male and female compared to national figure of India. In Aizwal District 46 male and 45 female cases per 100,000 residential people of that region were with Lung cancer, which make this place nationally top (both for male and female) and 9th around the globe in male and 4th in female.(NCRP, 2013; Ferlay et al., 2013). Smoking is responsible for upwards of 80% of all lung cancers worldwide (www.who.int/tobacco/research/cancer/en). Non-smokers account for 15% of lung cancer cases and these cases are often attributed to a combination of genetic factors, radon gas, asbestos, pesticides and air pollution including passive and static smoking (Kirmani et al., 2010).

Tobacco growing in the field or in an uncured state is called green tobacco. Workers engaged in tobacco cultivation suffer from an occupational illness is known as .green tobacco sickness (GTS). The illness was first reported among tobacco workers in Florida, in 1970, as cropper sickness (Banerjee et al., 1993). Later, it was found to be caused by the absorption of nicotine from wet tobacco plants and reported as GTS (Merino et al., 1998). Form the starting of these GTS disease now tobacco is a major global thread to humankind causing different kinds of disease like cardiovascular to cancer and responsible for high mortality compared to all death caused by other factors. The International Agency for Research on Cancer (IARC) Monograph states that tobacco smoking is the major cause of lung cancer (all types) and is majorly associated with oral cancer, cancers of the oropharynx and hypopharynx, oesophagus, stomach, nasopharynx (IARC, 1987).

TRCs in all North Eastern region was very high and comprises more than 50% cases compared to all cancer cases. It was observed that in Female all the North Eastern States comprises high Tobacco Related Cancers (TRCs) to all Cancers incidence cases than national figure (Table 1). In Female all the high TRC related incidence regions in India were alone from N.E. region i.e. all 14 regions of N.E. India were at top. In most of the regions of NE-India, Female heavily smoked and chewed tobacco related substances compared to other parts in India. In Males Meghalaya states (Meghalaya and East Khasi Hills both) shows nearly 70% of all cancers was in the states were due to TRCs followed by Tripura, Dibrugarh and Kamrup comprising 50% TRC cases compared to total cancer cases (NCRP 2013). Cancer deaths occurred from TRCs in people aged 30-69 years in men for the geographical region Assam and for whole North-East (NE), the Age --Standarised death rate were 114.0 (Cumulative Risk-8.5 CI 6.5-10.4) and 115.0 (Cumulative Risk-11.2 CI 8.9-13.5) while for Women it is 37.5 (Cumulative Risk-5.5 CI 3.8-7.1) and 21.0 (Cumulative Risk-6.0 CI 4.6-7.4) per 100,000 (Dikshit et al., 2012). A case.control study on oesophageal cancer in Assam found that men chewing dried tobacco (chadha) had a nearly 5-fold greater risk of oral cancer compared to non-users (Ihsan et al., 2010). The consumption of tobacco in various forms such as smoking, bidi, betel quid (paan) along with alcohol are the major 'preventable' risk factors of cancer (Siddiqui et al., 2012). So most of the TRC related cancers are preventable, needs of some cancer control activities.

Childhood cancer incidence varies worldwide between 80-150/million children (Ferlay et al., 2003; Riberio et al., 2008) The childhood cancer incident in this region of North East was not as much as high compared to other regions of India as per the NCRP report (NCRP, 2013). In India Chennai reported a high childhood incident in both Boys and Girls. Age Adjusted Incidence rate (AARpm) in Boys is reported as 159.6 while as for Girls it is 112.4 per one million children followed by Delhi in Boys (149.4) and Thiruvananthapuram in Girls (103.3). It is quite significant to notice that AAR is much high in all the North Eastern region of India but AARpm is not. It is one of the striking questions? In East Khasi Hills of Meghalaya AAR (Male-216.0) is 9 times higher than AARpm (Boys-24.3) and AAR (Female-114.1) is 10 times higher than AARpm (Girls-11.3) followed by Entire Meghalaya AAR (Male-157.3) is 8 times higher the AARpm (boys-18.6) and in Female AAR (Female-83.7) is 7 times higher than ARpm (Girls-11.9) (NCRP 2013). Whereas all other region of North East likes Nagaland, Mz. Excl Aizwal, Kamrup Urban etc where AAR was near about 2 to 3 times high in male and 2 to 4 times in female (Table 1) then AARpm for Boys and Girls. Was it like that as the person grows he will got addicted to different carcinogenic, risk factors or role of genetic susceptibility comes in to play or what else, researcher should look into this matter. Note that AARpm was calculated per one million children, while AAR is per 100,000 persons.

Materials and Methods

Principal components analysis (PCA)

For evaluation of relationship based on available data in India region as compared to other geographical regions a correlation coefficient matrix of genomic abundance data was generated for extraction of principal components. Varimax rotated first two principal components were used to make a scatter plot. Geographical areas with similar genomic patterns tended to be together in the reduced multivariate data space without loss of information due dimensionality reduction. Firstly PCA was done on the basis of Genomic data then from available AAR data.

Based on available genomic data

Genomic data of GSTM1 and GSTT1 gene of Null and Non-Null genotypes from Eastern Asia, China, South Eastern Asia, Thailand, Indonesia, Singapore-Malay, Northern Europe countries-Denmark, UK, Southern Europe countries-Slovenia, Greece, Western Europe countries-France-North West Eastern Europe countries-Poland, Slovakia, Russia, and Indian registries considered were Tamil Nadu Pondicherry, Andhra Pradesh, Karnataka, Kerela, Lucknow, India-North1, India-North2, India-South1, India-South2 were subjected to principal component analysis (PCA) to elucidate major patterns of variation in genetic diversity among these geographical regions.

Data on GSTM1 and GSTT1 gene in various regions used in the above analysis was obtained from various sources like Lee et al. (1995); Conde et al. (1999); Garte et al. (2001); Miyakis et al. (2003); Abbas et al. (2004); Naveen et al. (2004); Gajecka et al. (2005); Pakakasama et al. (2005); Antha (2009); Agusa et al. (2010); Konwar et al. (2010); Kurose et al. (2012).

Glutathione-S-transferase (GST) is an important phase II xenobiotic compound metabolizing enzyme family, involved in tolerance to a particular drug or susceptibility to a disease (Malakar et al., 2012). The GSTM1 gene is classified into the mu class and the GSTT1 gene belongs to the theta class. Genes coding for GSTM1 and GSTT1 proteins are polymorphic in humans and GSTM1 is absent in 35-60% of individuals (Bell et al., 1993). Glutathione-s-transferases (GSTs), a supergene family of detoxification enzymes, provide protection against genotoxic and carcinogenic effects of numerous substances of both xenobiotic and endogenous origins. The phenotypic absence of GSTM1 and GSTT1 activity is due to homozygocity for an inherited deletion of these genes, termed the null genotype (Pemble et al., 1994).

Based on available age specific incidence rate (AAR) data

Cancer Age Specific Incidence Rate (AAR) as reported and published by IARC in CI V-VOL IX and NCRP (ICMR) report 2009-2011 (NCRP 2013) were used to compare the relationship pattern of India and North Eastern region of India to other geographical regions of Asia.

The sites from which AAR were taken consideration for PCA were Tongue, Mouth, Nasopharynx, Hypopharynx, Oesophagus, Stomach, Colon, Gall Bladder, Lung, Brain, NHL, Liver (Male), Larynx (Male), Corpus uteri (Female), Ovary (Female), Thyroid (Female), Prostate (Male), Kidney (Male), Urinary Bladder (Male).

PCA was done within Indian population including NE states registries (Figure 3) then with Asian population registries to observe the pattern of variation among these Asian regions.

From PBCR 2009-2011 report we have took the mainland Indian regions were covered by PBCRs as follows Bangalore, Barshi Expanded, Bhopal, Chennai, Mumbai, Ahmedabad Rural, Ahmedabad Urban, Aurangabad, Kolkata, Kollam, Nagpur, Pune, Thiruvanthapuram, Wardha and the NE states regions as listed above.

While from IARC CI V-VOL IX regions which were considered were China Shanghai, Korea, Korea Daegu, Malaysia Penang, Malaysia Sarswak, Singapore, Singapore Malay, Thailang Chiang Mai, Thailand Lampang, Pakistan, Singapore Indian, Singapore Chinese, Japan Aichi Prefecture, Japan Hiroshima, China Guangzhou City, China Jiashan, China Nangang District Harbin City, Japan Fukui Prefecture, Japan Miyagi Prefecture, Japan Nagasaki Prefecture. Japan Osaka

Prefecture, Japan Yamagata Prefecture, Korea Busan, Korea Gwangju, Korea Incheon, Korea Jejudo.

Pearson correlation of co-efficient

A Bi-Variate Pearson correlation of Co-efficient test was used to observe the co relationship between NE states to Indian and Asian countries. A high Pearson Correlation coefficient describes a strong relationship between the variables while $p<0.05$ considered as significant. The value of r lies between-1 and 1. A value near 0 means no (linear) correlation and values near±1 means very strong correlation.

Hierarchical cluster analysis

A dendrogram based on AAR values of nineteen different sites of Cancer was generated by neighbour joining hierarchical cluster analysis to see the relationship among the countries studied.

Results

Principal components analysis (PCA) of the genomic and age specific incidence (AAR) data

The result of PCA based on Genomic data is given in Figure 4. Clustering of the geographical areas on the basis of similarity in the pattern of the genetic diversity from India and around the globe based on GSTM1 & GSTT1 genotypes clearly indicates that the genomic relationship of India do not resemble with neighboring South, East Asian countries like China, Thailand, Japan, Korea also with European Countries. While as Indian & Asian states both have not any relationship with European countries. This shows that genetic variation is available among different geographical regions, for which it may be incidence for different cancer site is different for different geographical regions and or may associated with other factors well as, focus was needed.

PCA was done based on AAR values within Indian population registries to observe the pattern of variation and dramatically found that the North Eastern (NE) population registries are showing quite different pattern from rest of mainland India (Fgure 4).

As from studies suggest that northeast India show closer genetic affinities with East Asian groups than with other Indian groups (Cordaux et al., 2013). A PCA was attempted to find out whether there exist any relationship between Indian population registries especially between NE states and South, East - Asian countries. It is observed that most of the NE states were resembles pattern with East, South Eastern Asian countries while only Pakistan resembles the pattern with the mainland India, This clearly signifies some relationship (Figure 5). It might be genomic or due to the historical background, as it is known that most of the tribes of North Eastern states of India were migrated from these Asian countries. So there may be a linkage of Genetic, Food habit, Ethnicity etc.

Pearson correlation test

Table 2 summarizes the results of PCT. The observed results suggest that most of the NE States have a strong relationship with South, Eastern Asian countries like Manipur state vs Thailand Chiang Mai (r=0.859), China, Nangang, District (r=0.859) and Thailad, lampang (r=0.799), whereas Imphal West District vs Thailand, Lampang (r=0.766), Thailand Chaing mai (r=0.826) and for most of the correlation was observed as statistically significant ($p<0.05$), while with mainland India it shows very poor relationship and was found not significant for most of the regions.

Hierarchical cluster analysis

Dendrogram based on genomic data of GSTM1 and GSTT1 shows a hierarchical cluster within the geographical areas, i.e. variations present for among

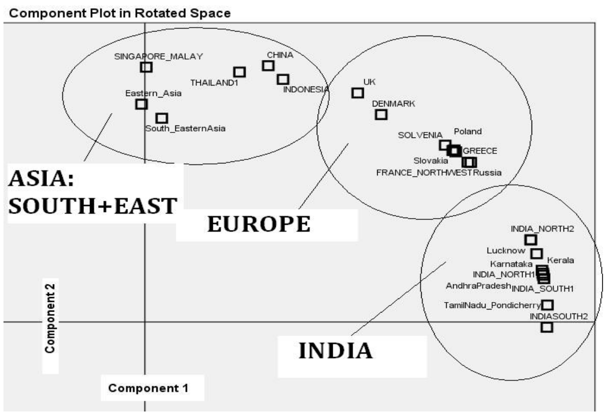

Figure 4. Principal Component Analysis of Genomic Data

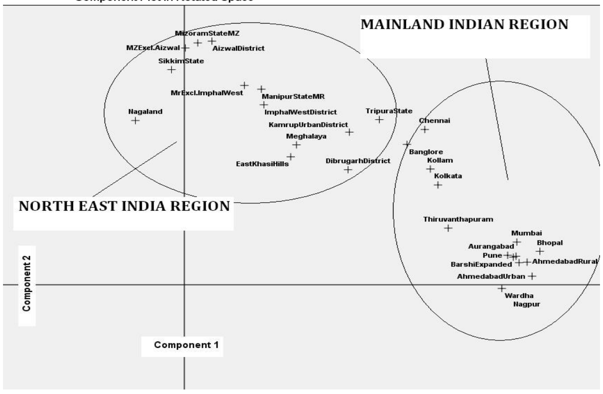

Figure 3. Principal Component Analysis of Age Adjusted Rate (AAR) Data of India

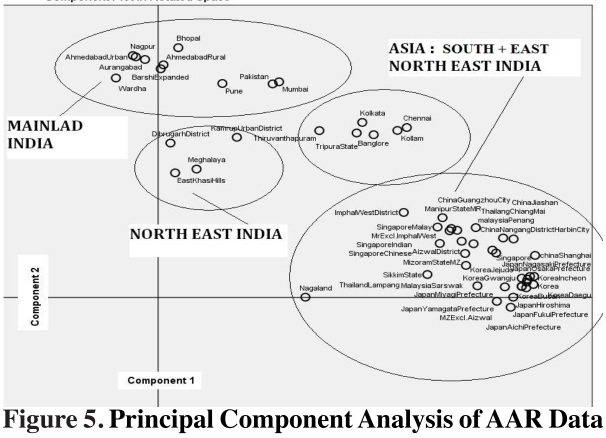

Figure 5. Principal Component Analysis of AAR Data of India and South East Asia

DOI:http://dx.doi.org/10.7314/APJCP.2014.15.8.3721

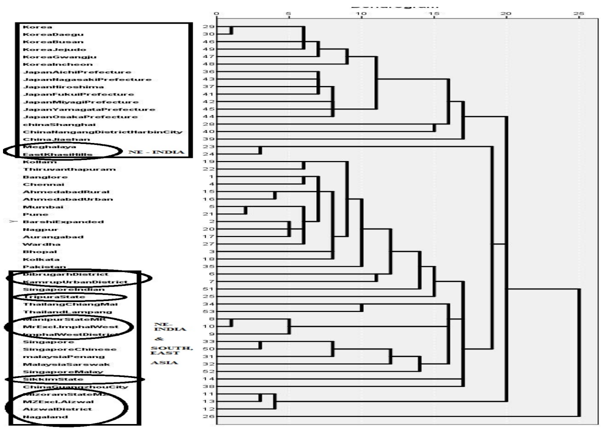

Figure 6. Hierarchical Cluster Analysis (HCA) Based on AAR Data

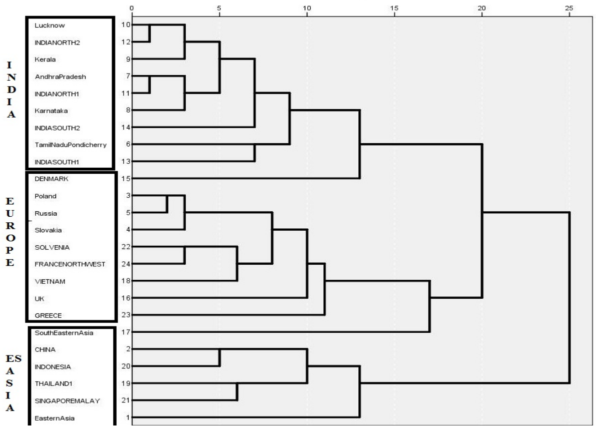

Figure 7. Hierarchical Cluster Analysis (HCA) Based on Genomic Data

different regions. This signifies that the genetic pattern is different for different geographical areas. Again based on incident values (AAR/ASR) of nineteen different sites of Cancer was generated by neighbour joining hierarchical cluster analysis and also find the relationship is different for different geographical regions. It that means due to the genomic differences pattern of incidence is different (Figure 6, 7).

Discussion

As per the report of NCRP (2013) and numerous researchers work on cancer it is evidently found that Cancer is becoming a family disease in this North East region of India. All the cancer sites of these region was at most high and was at top of the chart compared to national figure of India. State of Mizoram (Aizwal District & Entire Mizoram State), Meghalaya (East Khasi Hills & Entire Meghalaya) and Assam (Kamrup Urban District) accounting a very high cancer incident occurrence as on compared to international figure for all sites and for some sites they are the top leaders compared to global scenario. Oesophagus was at top most rising sites occupying top 5 positions in male and top 3 position in female by NE regions around globally. Male Oesophagus cancer in East Khasi Hills (AAR 71.4) (Meghalaya state) was 3.5 time higher then China, jishan (AAR 20.2) while in Female Oesophagus East Khasi Hills (AAR 30.2) was also 3.5 time higher then Pakistan, South Karachi (AAR 8.6). While the interesting fact is that for these regions as well as other NE regions the incidence of childhood

Table 2. Showing Vales of Correlation Coefficietn (r)

Regions	Indian Regions						
	Barshi Expanded	Bhopal	Mumbai	Ahmedabad Urban	Aurangabad	Nagpur	Pune
Manipur State (MR)	0.12	0.36	0.351	0.23	0.258	0.147	0.191
Imphal West District	0.104	0.385*	0.35	0.231	0.274	0.133	0.192
Mr Excl. Imphal West	0.083	0.3	0.332	0.163	0.223	0.109	0.178
Mizoram State (MZ)	0.144	0.133	0.214	0.087	0.17	0.169	0.186
Aizwal District	0.163	0.176	0.239	0.126	0.205	0.189	0.2
MZ-Excl. Aizwal	0.126	0.094	0.192	0.05	0.136	0.15	0.173
Sikkim State	0.117	0.087	0.131	0.001	0.065	0.122	0.133
Meghalaya	0.503**	0.288	0.167	0.339	0.407*	0.452*	0.277
East Khasi Hills	0.493**	0.265	0.144	0.324	0.380*	0.435*	0.266
Tripura State	0.445*	0.651**	0.459*	0.531**	0.517**	0.457*	0.34
Nagaland	0.137	-0.092	-0.082	-0.028	0.016	0.082	-0.02
Kamrup Urban District	0.542**	0.535**	0.383*	0.417*	0.433*	0.498**	0.441*
Dibrugarh District	0.595**	0.484**	0.296	0.482**	0.517**	0.567**	0.347

Regions	East- Asian South-Asian Regions						
	China Shanghai	China Guangzhou	Malaysia Sarswak	Thailand, Lampang	Thailang Chiang Mai	China Jiashan	China, Nangang District Harbin
Manipur State (MR)	0.758**	0.794**	0.761**	0.799**	0.859**	0.718**	0.859**
Imphal West District	0.676**	0.725**	0.640**	0.766**	0.826**	0.646**	0.792**
Mr Excl. Imphal West	0.786**	0.814**	0.800**	0.805**	0.862**	0.743**	0.876**
Mizoram State (MZ)	0.735**	0.511**	0.577**	0.438*	0.512**	0.744**	0.710**
Aizwal District	0.742**	0.546**	0.582**	0.480**	0.552**	0.756**	0.737**
MZ-Excl. Aizwal	0.723**	0.475**	0.569**	0.394*	0.467**	0.727**	0.677**
Sikkim State	0.642**	0.420*	0.421*	0.365*	0.389*	0.701**	0.610**
Meghalaya	0.135	0.074	-0.035	-0.007	-0.009	0.343	0.153
East Khasi Hills	0.087	0.048	-0.065	-0.029	-0.03	0.3	0.117
Tripura State	0.619**	0.605**	0.406*	0.574**	0.544**	0.663**	0.637**
Nagaland	0.265	0.125	0.381*	-0.11	-0.055	0.356	0.13
Kamrup Urban District	0.246	0.134	0.027	0.126	0.136	0.374*	0.253
Dibrugarh District	0.069	-0.037	-0.136	-0.062	-0.064	0.244	0.063

*Correlation is significant at the 0.05 level (2-tailed). $p<0.05$; **correlation is significant at the 0.01 level (2-tailed). $p<0.01$;

cancer is much lower than India and rest of the world, which signifies that probably most of the inhabitants of these regions were develop the condition later on; so the oblivious question is how, why and for what reason?

Later on as by the PCA it was observed that most of the NE-States regions shows a different kind of cancer incident pattern as compared to mainland national regions (figure 4) while found to be similar pattern with South, East Asian region (figure 5). While from various studies as cited above already revealed that the genomic linkage was present for NE-India to East, South – Asian region. Whereas Tripura of Ne-India region shows a nearby similar pattern with Kolkata of mainland India, Most of places in Tripura are dominant by Bengali race, from where maximum number of cases were reported and Kolkata was also dominant by Bengali race itself. So there might be some link up for the pattern of cancer cases, should look up into this matter. PCA based on AAR data shows a relationship between the geographical regions which were similar in race or ancestors were from same origin and or with the same kind of food habits, etc. A further elaborate investigation was needed therefore; which would help to stratify the pattern of occurrences of cancer cases.

As again by PCA based on the GSTM1 and GSTT1 genetic data it is evidently observed that pattern of genome pool is different for different geographical regions (Figure 4). So was it the reason for which the incident pattern of different cancer sites was different among geographical region. The similarity in genomic pool as observed by the PCA analysis also leads towards the conclusion that as it might be due to the different kind of genomic pool the pattern of incidence was different among geographical regions.

The correlation among the region shows NE-India have a strong correlation to South, East Asian countries compared to mainland India. And the correlation is also found as statistically significant ($p<0.05$) (Table 2, while a very low or moderate level relationship was found with mainland India.

Hierarchical Cluster Analysis (HCA) based on GSTM1 and GSTT1 genomic data, clustering all the similar genome pools together for their respective regions and was also found similar for HCA analysis for Age Adjusted Rate (AAR) data. It is also notice in HCA analysis the NE-India is clustering with South, East-Asian countries. As by most of the studies showed that the food habits, lifestyle, ethnicity and genome pool is quite different from rest of India and a close linkage is present to South, East-Asia; (Cordaux et al., 2003) whereas most of the studies suggested that risk factor association to cancer sites was due to some of the ethnic food habits (including daily consumption items) and locally made tobacco and alcohol is highly associated in developing the condition; like Kalakhar with 8 fold risk to Oesophagus cancer, (Phukan et al., 2001) combined consumption habits of Meziol and Tuibur shows 23 fold risk to Stomach cancer. (Malakar et al., 2012) So a depth study was needed to find out actual causes of high incident, was it due to the food habits (including tobacco and alcohol) and or role of genetic susceptibility, heredity or else needs to be find out.

In conclusion, these studies suggest that continuation of the monitoring and depth investigation is needed to control the high incidence pattern of cancer in this geographical region North East India.

Acknowledgements

The author thankful to Dr. A. Nandakumar, Deputy Director General, ICMR and Officer in charge of the National cancer Registry Programme of India and Dr. A.C Kataki, for necessary guidance and valuable advices. The author also thankful to Mr. JA Barbhuiya, Mr. Ranjan Lahon, Mrs. Barsha Deka and Mr. Kamal C. Deka for helping in the preparation of paper, field work and data entry.

References

Abbas A, Delvinquiere K, Lechevrel M, et al (2004). GSTM1, GSTT1, GSTP1 and CYP1A1 genetic polymorphisms and susceptibility to esophageal cancer in a French population: different pattern of squamous cell carcinoma and adenocarcinoma. *World J Gastroenterol*, **10**, 3389-93.

Accesed from the addresshttp://www.cancer.org/cancer/cancercauses/geneticsandcancer/heredity-and-cancer. Accessed on 15th March 2014.

Amtha R, Ching CS, Zain R, et al (2009). GSTM1, GSTT1 and CYP1A1 polymorphisms and risk of oral cancer: a case-control study in Jakarta, Indonesia. *Asian Pac J Cancer Prev*, **10**, 1-26.

Agusa T, Iwata H, Fujihara J, et al (2010). Genetic polymorphisms in glutathione S-transferase (GST) superfamily and arsenic metabolism in residents of the Red River Delta, Vietnam. *Toxicology and Applied Pharmacology*, **242**, 352-62.

Alfred G, Knudson, Jr (1985). Hereditary cancer, oncogenes, and antioncogenes. *Cancer Res,* **45**, 1437-43.

Arora RS, Eden TOB, Kapoor G (2009). Epidemiology of childhood cancer in India. Epidemiology of childhood cancer in India. *Indian J Cancer,* **46**, 264-73.

Banerjee A, Pakrashi A, Chatterjee S, Ghosh S, Dutta SK (1993). Semen characteristics of tobacco users in India. *Arch Androl*, **1**, 35-40.

Bell DA, Taylor JA, Paulson DF, et al (1993). Genetic risk and carcinogen exposure: a common inherited defect of the carcinogen-metabolism gene glutathione S-transferase M1 (GSTM1) that increases susceptibility to bladder cancer. *J Natl Cancer Inst,* **85**, 1159-64.

Chelleng PK, Narain K, Das HK, Chetia M, Mahanta J (2000). Risk factors for cancer nasopharynx: a case-control study from Nagaland, India. *Natl Med J India*, **13**, 6-8.

Conde AR, Martins G, Saraiva C, Rueff J, Monteiro C (1999). Association of p53 genomic instability with the glutathione S-transferase null genotype in gastric cancer in the Portuguese population. *Mol Pathol*, **52**, 131-4.

Cordaux R, Saha N, Bentley GR, Aunger R, SM Sirajuddin (2003). Mitochondrial DNA analysis reveals diverse histories of tribal populations from India. *Eur J Hum Genet,* **11**, 253-64.

Dialyna IA, Miyakis S, Georgatou N, Spandidos DA (2003). Genetic polymorphisms of CYP1A1, GSTM1 and GSTT1 genes and lung cancer risk. *Oncol Rep*, **10**, 1829-35.

Dikshit R, Gupta PC, Ramasundarahettige C, et al (2012). Cancer mortality in India: a nationally representative survey. *Lancet*, **379**, 1807-16

Doll R, Payne P, Waterhouse JAH (1966). Cancer incidence in fivecontinents. Vol I. Geneva: Union Internationale Contre le cancer.

Ferlay J, Shin HR, Bray F, et al (2010). Estimates of worldwide burden of cancer in 2008: GLOBOCAN 2008. *Int J Cancer,* **127**, 2893-917.

Ferlay J, Curado MP, Edwards B, et al (2007). Cancer incidence in five continents, VOL. IX, IARC Scientific publications No. 160, Lyon, IARC.

Ferlay J, Bray F, Pisani P, Parkin DM (2003). Cancer incidence, mortality and prevalence worldwide. GLOBOCAN 2008. Lyon: International Agency for Research on Cancer (IARC Press) 2003.

Gajecka M, Rydzanicz M, Jaskula-Sztul R, et al (2005). CYP1A1, CYP2D6, CYP2E1, NAT2, GSTM1 and GSTT1 polymorphisms or their combinations are associated with the increased risk of the laryngeal squamous cell carcinoma. *Mutation Res*, **574**, 112-23.

Garte S, Gaspari L, Alexandrie AK, et al (2001). Metabolic gene polymorphism frequencies in control populations. *Cancer Epidemiol Biomarkers Prev*, 10, 1239-48.

Hussain MA, Pati S, Swain S, et al (2012). Pattern and trends of cancer in Odisha, India: a retrospective study. *Asian Pac J Cancer Prev,* **13**, 6333-6.

Ihsan R, Chattopadhyay I, Phukan R, et al (2010). Role of epoxide hydrolase 1 gene polymorphisms in esophageal cancer in a high-risk area in India. *J Gastroenterol Hepatol*, **25**, 1456-62.

International Agency for Research on Cancer (IARC) (1987). IARC monographs on the evaluation of the carcinogenic risks to humans, Supplement 7. Overall evaluations of carcinogenicity: An updating of IARC monographs Volumes 1.42. Lyon: IARC Press: 357-61.

Jemal A, Bray F, Melissa MM, et al (2011). Global cancer statistics. *CA Cancer J Clin*, **61**, 69-90.

Kataki AC, Simons MJ, Das AK, Sharma K, Mehra NK (2011). Nasopharyngeal carcinoma in the Northeastern states of India. *Chin J Cancer*, **30**, 106-13.

Kirmani N, Jamil K, Naidu MUR (2010). Occupational and environmental carcinogens in epidemiology of lung cancer in South Indian population. *Biology Med*, **2**, 1-11.

Konwar R, Chaudhary P (2010). Glutathione S-transferase (GST) gene variants and risk of benign prostatic hyperplasia: a report in a North Indian population. *Asian Pac J Cancer Prev,* **11**, 1067-72.

Kumar A, Chatopadhyay T, Raziuddin T, Ralhan R (2006). Discovery of deregulation of zinc homeostasis and its associated genes in esophageal squamous cell carcinoma using cDNA microarray. *Int J Cancer,* **120**, 230-42.

Kurose K, Saito Y (2012). Population differences in major functional polymorphisms of pharmacokinetics/pharmacodynamics-related genes in Eastern Asians and Europeans: implications in the clinical trials for novel drug development. *Drug Metab Pharmacokinet*, **27**, 9-54.

Lee EJ, Wong JY, Yeoh PN, Gong NH (1995). Glutathione S transferase-theta (GSTT1) genetic polymorphism among Chinese, Malays and Indians in Singapore. *Pharmacogenetics,* **5**, 332-5.

Mandal SK, Singh TT, Sharma TD, et al (2013). Clinico-pathology of lung cancer in a regional cancer center in northeastern India. *Asian Pac J Cancer Prev*, **14**, 7277-81.

Malakar M, Devi KR, Phukan RK, et al (2012). Genetic polymorphism of glutathione S-transferases M1 and T1, tobacco habits and risk of stomach cancer in mizoram, India. *Asian Pac J Cancer Prev*, **13**, 4725-32.

Merino G, Lira SC, Martinez-Chequer JC (1998). Effects of cigarette smoking on semen characteristics of a population in Mexico. *Arch Androl*, **41**, 11-5.

Naveen AT, Adithan C, Padmaja N, et al (2004). Glutathione S-transferase M1 and T1 null genotype distribution in south Indians. *Eur J Clin Pharmacol*, **60**, 403-6.

National Cancer Registry Programme (ICMR) (2013). Three-year report of the population based cancer registries: 2009-2011. Bangalore, India.

Pakakasama S, Mukda E, Sasanakul W, et al (2005). Polymorphisms of drug-metabolizing enzymes and risk of childhood acute lymphoblastic leukemia. *Am J Hematol*, **79**, 202-6.

Pemble S, Schroeder KR, Spencer SR, et al (1994). Human glutathione S-transferase theta (GSTT1): cDNA cloning and the characterization of a genetic polymorphism. *Biochem J*, **300**, 271-6.

Phukan RK, Chetia CK, Ali MS, Mahanta J (2001). Role of dietary habits in the development of oesophageal cancer in assam, the north-eastern region of India. *Nutr Cancer*, **39**, 204-9.

Phukan RK, Zomawia E, Hazarika NC, et al (2004). High prevalence of stomach cancer among the people of Mizoram, India. *Curr Sci*, **87**, 285-90.

Ramshankar V, Krishnamurthy A (2013). Lung cancer detection by screening-presenting circulating miRNAs as a promising next generation biomarker breakthrough. *Asian Pac J Cancer Prev*, **14**, 2167-72.

Rao DN, Ganesh B (1998). Estimate of cancer incidence in India in 1991. *Indian J Cancer*, **35**, 10-7.

Ribeiro RC, Steliarova-Foucher E, Magrath I, et al (2008). Baseline status of paediatric oncology care in ten low-income or mid-income countries receiving my child matters support: a descriptive study. *Lancet*, **9**, 721-9.

Senthilkumar KP, Thirumurugan R (2012). GSTM1 and GSTT1 allele frequencies among various Indian and non-Indian ethnic groups. *Asian Pac J Cancer Prev*, **13**, 6263-9.

Sharma TS, Singh TT, Laishram RS, et al (2012). Nasopharyngeal carcinoma-a clinico-pathological study in a regional cancer centre of northeastern India. *Asian Pac J Cancer Prev*, **13**, 1583-7.

Sharma JD, Kataki AC, Vijay CR (2013). Population-based incidence and patterns of cancer in Kamrup Urban Cancer Registry, India. *Natl Med J India*, **26**, 133-41.

Segi M (1960). Cancer Mortality for Selected Sites in 24 Countries (1950-57). Sendai, Japan: Department of Public Health, Tohoku University of Medicine.

Siegel R, Naishadham D, Jemal A (2013). Cancer statistics 2013. *CA Cancer J Clin,* **63**, 11-30.

Siddiqui Md S, Chandra R, Aziz A, et al (2012). Epidemiology and histopathological spectrum of head and neck cancers in Bihar, a State of Eastern India. *Asian Pac J Cancer Prev*, **13**, 3949-53.

Tobacco Free Initiative (TFI). World Health Organization. http://www.who.int/tobacco/ research/cancer/en/. Accessed on 12th March 2013.

Takiar R, Nadayil D, Nandakumar A (2011). Projection of number of cases in India (2010-2020) by cancer groups. *Asian Pac J Cancer Prev*, **11**, 1045-49.

Wang Y-C, Wei L-J, Liu J-T, Li S-X, Wang Q-S (2012). Comparison of cancer incidence between China and the USA. *Cancer Bio Med*, **9**, 128-32.

World Health Organization. World Cancer Report 2008, edited by Peter Boyle, Benard Levin. Lyon, France: International Agency for Research on Cancer; Lyon.