RESEARCH COMMUNICATION

A Cohort Study to Address the Natural History of HPV and Cervical Dysplasia in Trivandrum, South India: Methodological Issues and Initial Results

Cherian Varghese[1], MC Kalavathy, NS Amma, MK Nair

Abstract

Epidemiological studies, especially cohort studies have many limitations in countries like India. The present population based molecular epidemiolohgical cohort study was planned to address the prevalence, risk factors and natural history of uterine cervical HPV infection in women in an area of suburban dwellings in South India. Epidemiological data, blood samples and cervical scrape smears were collected from women after obtaining an informed consent. A compliance rate of 38% for the first round of screening was noted. More than 70% of women demonstrated evidence of inflammation in their smears.

Key words: cervical cancer - developing countries - molecular epidemiology - human papillomavirus

Asian Pacific J Cancer Prev, **2**, 63-67

Introduction

Improvements in living standards and control of communicable diseases have led to cancer and other chronic diseases emerging as major public health problems in certain quarters of the Indian sub-continent (Nair and Sankaranarayanan, 1991). India is in a phase of transition and the economic and social liberalisation that is taking place will lead to an epidemiological transition (Murthy et al., 1990). Lack of reliable data on the magnitude of these diseases have been a major limitation in assigning priorities in medical research for chronic diseases. The Indian Council of Medical research has realised this and established population based cancer registries since 1982. Reports from the National Cancer Registry Programme of India reveal that the rate of occurrence of cancer is low in India compared to western countries, but the huge population in India will generate substantially large number of cases. The pattern of cancer is also different with a predominance of tobacco related cancers (Indican Council of medical Research, 1990).

Epidemiological research in cancer has been mainly undertaken in western populations. Even though there are

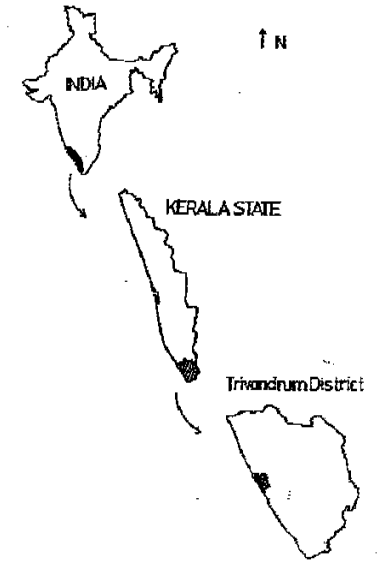

Figure 1. Location of the Regional Cancer Centre

[1]Address for correspondence :C VargheseMD PhD, Associate Professor, Cancer Epidemiology & Clinical Research Division, Regional Cancer Centre, S India. MC Kalavathy MBBS, MPH, Medical Officer, Regional Cancer Centre, S India. N S Amma MD, Former Professor of Cytopathology, Regional Cancer Centre, S India. M K Nair MD, FRCR, Director, Regional Cancer Centre, Trivandrum, S India.

similarities in the biology of the disease, the findings of these studies are not always applicable to the Indian population. Hospital based case-control studies have been undertaken by some centres in India, but they suffer from the lack of a temporal relationship between exposure and outcome (Sankaranarayanan et al., 1990; 1994). Prospective cohort studies can overcome this limitation and are very powerful designs for aetiological research in lifestyle related diseases (Doll and Hill, 1964).

Kerala, a southern sate of the Indian Union has achieved high rates of literacy and vital statistics which are comparable to the industrialised countries of the world. The Regional Cancer Centre (RCC) in Trivandrum, is a comprehensive cancer centre with a division of cancer epidemiology and clinical research. The centre has considerable experience in field work and clinical and epidemiological cancer research and has initiated collaborative research programmes with national and international agencies.

Cancer of the uterine cervix is the most frequent cancer of women in all the developing countries and the second most common cancer among females all over the world. The number of new cases of cancer of the uterine cervix, globally, is estimated to be 437,000 in one year. The Southern Asian region contributes just over one quarter of the total load. In developed countries the incidence rates are generally low, with age standardised rates being less than 15 per 100,000. In Southern Asia the incidence rate is 21 - 22 per 100,000 (Parkin et al., 1993). The developed countries of the world have experienced a steady decline in the morbidity and mortality due to cervical cancer over the years. Changes in the lifestyle and the widespread practice of organised pap smear programmes have been the major reasons for this decline (Beral et al., 1994). There are no organised community-based screening programmes for control of cervical cancer, mainly due to the lack of resources and due to the lack of awareness in the community.

Epidemiologic studies have identified early age at first sexual intercourse, multiple sexual partners, low social class and history of sexually transmitted diseases as the risk factors for invasive cervical cancer (Brinton et al., 1987). These risk factors suggest that a venereally transmitted infectious agent is the cause for cervical neoplasia and recent laboratory studies have provided a plausible mechanism for human papillomavirus (HPV) -induced cervical neoplasia (Werness et al., 1990).

Carcinoma of the uterine cervix is the second most common cancer in women in Kerala. The natural history of HPV as described by studies in the western population is not applicable to the women in Kerala due to the differences in sexual, marital and reproductive practices. We have initiated a cohort study in Trivandrum to address the natural history of HPV and cervical neoplasia in collaboration with the Institute of Cancer Research in UK. The specific objectives of the study were to obtain the age specific prevalence of cervical HPV infection and to study the factors associated with persistence of infection. This paper will address the methodological issues in generating a population based cohort study in India, the strategies used to overcome them and the initial results of our study.

Subjects and Methods

Generating a cohort of women from a stable population in a geographically defined region was the first priority. Kazhakuttom, a suburban Panchayat (smallest administrative unit in Government) in Trivandrum was selected as the study area (See the Map). This area has a good representation of all the three major religious groups (Hindu, Christian and Muslim) and is familiar to the researchers in RCC. The proximity to RCC was also considered as an advantage. A team of four social workers (three from the same locality), a social scientist and a medical officer were recruited.

The first task was to identify the eligible women and to assign unique identification numbers to them for future reference. The list of houses kept in the Panchayat office was not updated and the Panchayat numbers were not easily traceable on the houses. The addresses kept in the Panchayat office were also incomplete in certain cases. We were also not sure about the permanent residence status of the families. Without identifying the base population it will be difficult to identify the eligible population and to measure the compliance to the study.

The Panchayat was therefore divided into 12 electoral wards with ward boundaries available from the map in the office. An adjacent fishing village Maryanadu was also selected for the study. The population of Maryanadu has a unique life style and habits and is predominantly made up of Christian religious groups. The social workers of the project visited all the households in the wards and collected information on a structured format. The format included full address, name of the head of the household, members present, their age gender and marital status, religious group and average monthly income. In a brief interview their permanent residence status (residence at the given address for more than 12 months) in the area was also determined. All houses and individuals in the households were given a unique identification number. This number consisted of ward number (two digits), house number (three digits) and another two digits to identify the individual. This process took eight months to complete. All the forms were entered into a computer and from this database a list of all married women who are permanent residents of Kazhakuttom was generated.

As per the protocol of the study all eligible women had to be interviewed, and biological samples (blood sample and cervical smear) had to be collected. Once we had identified the potentially eligible study population, the workers invited the eligible women by ward to attend clinics. There is only one Government health centre in the Panchayat and this was not sufficient. Another building was hired, and clinics were started there, but many women failed to attend as they found it difficult to travel to the centre. It became obvious that in order to achieve maximum compliance it would be necessary to set up a number of smaller clinics close to where the women lived. This was achieved by holding the clinics in

peoples homes. Awareness about the project was generated by providing information to women's organisations (Mahila Samajam's) and places of worship. Wide publicity was given through local newspapers, slides in cinema halls and through individual letters.

On a typical clinic day, a team of researchers including the medical doctor, cytotechinicians and laboratory personnel would arrive at the designated clinic in a vehicle. Before organizing a clinic the social workers would have given the information to the community and women who responded to the invitation presented themselves at the clinic. They were then given additional information about the study and after obtaining an informed consent were interviewed and a blood sample and cervical swab were collected. The biological samples were stored in ice-packs and taken back to RCC for processing.

The analysis for HPV is to be conducted at the Institute of Cancer Research Laboratory in Sutton, UK and for this the samples are being shipped in dry ice. This study is planned to include annual screening which will provide information with regard to HPV infection and cytological abnormalities in the uterine cervix. Incidence and mortality from cervical cancer will be the long term outcome. Since there are no means of ascertaining these end points at present,a surveillance programme to monitor the incidence of cancer in Trivandrum was started with the partial assistance of the International Agency for Research on Cancer, Lyon, France and we hope that this mechanism will provide the incidence data for this cohort in the years to come. Three rounds of screening are over and the initial results are provided here.

Results

Kazhakuttom Panchayat (KP), the study location has an area of 19.5 sq. Km. and is situated at a distance of 12 Km from RCC (Fig.1). There were 30,013 subjects in the 12 wards of the Panchayat at the initial registration.

Table 1. Total Population, Eligible Population and Compliance to Screening in the First Round.

Ward No.	Total Population	Eligible Individuals	1st Round	Compliance
I	2836	891	425	48%
II	2290	686	229	33%
III	1762	532	290	55%
IV	2319	675	191	28%
V	2453	710	368	52%
VI	2919	860	204	24%
VII	2950	881	201	23%
VIII	3348	1017	394	39%
IX	1968	594	153	26%
X	2624	805	318	40%
XI	2850	849	346	41%
Maryanadu	1694	414	228	55%
Total	30013	8914	3347	38%

The age distribution of the population in the study area showed that except for Maryanadu, the age distribution is relatively constant in all the wards and that 15-18% of the population is in the 0-9 age group. In Maryanadu, 23% of the population was below the age of nine years. The male:female ratio across the region was 1:1, except in Maryanadu which showed a slight male predominance.

Distribution by religious groups in the study area revealed Hindus to be the predominant religious group, followed by Muslims and Christians. The proportion of Muslims varies in the different wards and they tend to cluster in defined geographic areas around a place of worship.

The joint income of the family as stated by the head of the household is not fully reliable and there was no other objective evidence regarding the family income. Ward one in KP and Maryanadu village has a high proportion of families in the low socio-economic category. The high proportion of middle income groups in the study area overall, however, makes it a representative sample of the population of Kerala state.

Out of the 8914 eligible women in the base cohort, 3347 (38%) complied with the invitation for the first round of screening (Table-1). In the initial part of the study women were interviewed in their residence and then invited toaaaa the clinics. To assess the representativeness of the compliant subjects we compared the age distribution, distribution by religion and socio-economic status of the compliant group and the non-compliant group forn those subjects who were interviewed in their house (n=1748).

The women who complied with the invitation for screening tend to be in the younger age groups. The non-compliant group had a high proportion of Muslims and Christians, compared to the group who underwent screening. Table 2, shows the distribution of the subjects in the compliant and non-compliant group, by monthly family income. Lower middle class and very high income groups are less likely to attend the screening programme.

Educational status is a good surrogate measure for social class and Table 3 presents the educational status of the subjects in the compliant and non-compliant groups.

The age specific prevalence of cytological abnormalities is presented in Table-4. Results from the cohort in Trivandrum shows that across all age groups 70% of women feature inflammation on their cervical smears. The cervical

Table 2. Compliance to Screening and Joint Family Income

Income (Rs)/month	Compliant n (%)	Non-compliant n (%)
<1000	1549 (47.3)	583 (36.2)
1001-2000	1015 (31.0)	595 (37.0)
2001-3000	489 (14.9)	316 (19.6)
3001-4000	164 (5.0)	67 (4.2)
4000+	5 (1.7)	48 (3.0)
Total	3274	1609

Table 3. Educational Status and Compliance to Screening

Education	Compliant n (%)	Non-Compliant n (%)
Illiterate	453 (13.8)	212 (12.9)
Literate	375 (11.4)	321 (19.5)
Primary	640 (19.5)	314 (19.1)
Secondary	805 (24.5)	539 (32.8)
College	808 (24.6)	259 (15.7)
Professional	201 (6.1)	-
Total	3282	1645

cells are being tested for HPV in collaborating laboratories in the United Kingdom.

Discussion

This large cohort study with biological sample collection opened up many lacunae in the health infrastructure in India for epidemiological research. In western countries registration with the general medical practitioner, membership in insurance schemes (Klatsky et al., 1974), motor vehicle licence, telephone directory, etc can be used as the sampling frame with unique identification features, but no such systems are available at present in India. Electoral lists and wards are changing every 5 years and cannot be considered as reliable. The joint family system in India adds to the difficulty to identify an individual uniquely. The relatively small study area enabled us to do a house to house survey in order to establish the baseline demographic data. Computerisation is an essential element in data management and we had realised this from the beginning. Once the baseline database is generated listings of populations can be made in different formats.

Ensuring adequate compliance to the study was the second major difficulty. The compliance achieved in this study cannot be compared to a study conducted in a Western population. Women had to attend to the daily chores of their household and in many instances they were earning members. Their list of priorities start with food and shelter and it was difficult to drive home the message that annual screening by cytology can help to prevent the possibility of cervical cancer 10 years later. This was the first organised programme in the region and we hope that as time goes by more and more people will participate in the study.

Since Kerala is a very literate state, we were able to use the print media to make the community aware of the study. The health benefits of participation were explained in small group discussions. Clinics were held as close to the homes of the eligible women as possible. The medical officer of the study offered health care to other members of the family and this provided an incentive for the women to come to the clinic. Subjects in the high educational levels and middle income categories were more likely to participate. This was not surprising as high income groups in India need not necessarily have high levels of education, which is the single most important factor in determining a healthy life-style. Health care workers in the community whose mandate is to provide health education, need to be made aware of the changing disease patterns and the importance of primary prevention strategies for chronic diseases. The services of the print and electronic media should also be tapped into for providing health education messages to the community.

End point ascertainment seems to be the most insurmountable obstacle in India for longitudinal studies. The sources of data collection available in a western country are illustrated in the Framingham study (Dawber et al., 1963). None of the resources mentioned in Framingham were available in Trivandrum. There are no means to identify one individual in any hospital as there is no unique identification number by which an individual can be traced in the health system. Multiplicity of health systems (Ayurveda, Homeopathy) adds complexity to the problem. Medical record keeping in the hospitals also leaves much to be desired.

Chronic diseases are not notifiable and there is no reliable source of incidence data. Death certification is not mandatory and there is no method to identify the cause of death. Even though there are competing priorities for the health sector in India, it is essential that death certification in the standard WHO format by a medical practitioner is made mandatory.

Table 4. Prevalence of Cytological Findings in Different Age Groups

	Cytological findings							
Age gp	Normal n(%)	Inflmn. n(%)	TV/Fungal n(%)	CIN1 n(%)	CIN 2 n(%)	CIN 3 n(%)	Inv Ca n(%)	Total
10-19	3 (23)	9 (69.2)	-	1 (7)	-	-	-	13
20-29	119 (19.6)	555 (72.8)	52 (6.8)	35 (4.5)	1 (0.13)	-	-	762
30-39	151 (13.1)	837 (72.3)	93 (8.0)	72 (6.2)	1 (0.08)	2 (0.17)	-	1156
40-49	114 (15.0)	513 (67.5)	72 (9.4)	55 (7.2)	2 (0.26)	2 (0.26)	1 (0.13)	759
50-59	97 (25.2)	222 (57.6)	11 (2.85)	52 (13.5)	2 (0.52)	-	1 (0.26)	385
60-69	74 (39.5)	100 (53.4)	4 (1.3)	8 (4.2)	-	1 (0.53)	-	187
70+	10 (25.6)	24 (61.5)	1 (2.56)	2 (5.1)	-	-	2 (5.1)	
Total	568	2260	233	225	6	5	4	3301

(TV- Trichomonas Vaginalis)

Establishment of disease registers like the cancer registries and making the major chronic diseases like diabetes, cardiovascular diseases and cancer, notifiable will go a long way in coping with the control of diseases of the next century.

A literature search in 'Medline' for the years 1975-96, using the key words 'cohort' and 'India' resulted in 70 citations. Fifty two of them (74%) of them were based on patients attending a hospital with very short follow up periods. There was only one study addressing precancer/ cancer in the general population (Gupta et al., 1992) . Two other studies were population based cohorts, a population based study in Delhi to examined the risk factors for coronary heart disease (Chadha et al., 1993) and another study on Bhopal gas victims (Andersson et al., 1990).

Cervical cytology of women showed a very high proportion of infections in the genital tract in all the age groups. This was the first time these women had a smear test as cytology facilities are not widely available in Kerala. Moreover, women are reluctant to present to the health care system with symptoms of cervico-vaginal infections. Organised screening programmes coupled with provision of appropriate treatment and awareness programmes among women can reduce the rate of infections in the genital tract of women.

Chronic inflammation of the cervix is associated with cervical dysplasia. A doubling of the risk of cervical dysplasia in the presence of Trichomonas Vaginalis has been observed (Zhang and begg, 1994) . Fifty five percent of the smears had fungal or bacterial overgrowth in cultures. The bacterial flora consisted of a variety of microbes. Chronic infection of the cervix and vagina might be a co-factor for the persistence of HPV infection and this can be addressed in our cohort.

Even though cervical cancer is a major public health problem in developing countries, there have not been many studies addressing the risk factors for HPV and for cervical dysplasia. The lack of population based data and the limitations in the cytological and laboratory infrastructure could be the major reasons for the paucity of the studies. It is important to identify risk factors for HPV infection which offer the potential for primary prevention.

In establishing a population based cohort with biological sample collection we have demonstrated that it is feasible to set up studies in India. International collaborative studies of this nature for mutual benefit should be the way forward for medical research in a world with diminishing resources.

Acknowledgement

The authors would like to acknowledge Prof Julian Peto and Dr Judy Deacon at the Institute of Cancer Research, Sutton, UK for their support in the conduct of this study. Thanks are also due to Smt G Mathew, Sociologist of the project and other workers.

References

Andersson N, Ajwani M K, Mahashabde S, et al (1990). C D Delayed eye and other consequences from exposure to methyl isocyanate: 93% follow up of exposed and unexposed cohorts in Bhopal. *Br J Ind Med*, **47,** 553-8.

Beral V, Hermon C, Munoz N, Devesa SS (1994). Cervical Cancer. Cancer Surveys: Trends in Cancer Incidence and Mortality. (Eds) R Doll, JF Fraumeni Jr, CS Muir. *Cold Spring Harbor Laboratory Press*, pp 265-86.

Brinton LA, Hamman RF, Huggins GR, et al (1987). Sexual and reproductive risk factors for invasive squamous cell cancer. *JNCI*, **79**, 23-30.

Chadha SL, Ramachandran K, Shekawat S, Tandon R, Gopinath N (1993). A 3- year follow up study of coronary heart disease in Delhi. *Bulletin of the World Health Organisation*, **71**, 67-72.

Dawber TR, Kannel WB, Lyell LP (1963). An approach to longitudinal studies in a community: The Framingham study. *Ann N Y Acad Sci*, **107**, 539-56.

Doll R, Hill AB (1964). Mortality in relation to smoking: Ten years observation of British Doctors. *Br Med J*, **1**, 1399-410.

Gupta P C, Mehta F S, Pindborg J J, et al (1992). Primary prevention trial of oral cancer in India: a 10-year follow-up study. *J Oral Pathol & Med*, **21**, 433-9.

Indian Council of Medical Research, New Delhi 1990.Biennial report 1988-1989. National cancer registry programme of India.

Klatsky AL, Friedman GD, Siegelaub AB (1974). Alcohol consumption before myocardial infarction. *Ann Intern Med*, **81**, 294-301.

Murthy NS, Juneja A, Seghal A, Prabhakar AK, Luthra VK (1990). Cancer projection by the turn of the century: Indian scene. *Indian J Cancer*, **27**, 74-82.

Nair KM, Sankaranarayanan R (1991). Epidemiologic leads to cancer control in India. *Cancer Causes and Control*, **3**, 263-5.

Parkin DM, Pisani P, Ferlay J (1993). Estimation of the worldwide incidence of eighteen major cancers in 1985. *Int J Cancer*, **54**, 594-606.

Sankaranarayanan R, C Varghese, Duffy SW, et al (1994). A case-control study of diet and lung cancer in Kerala, South India. *Int J Cancer*, **58**, 644-9.

Sankaranarayanan R, Duffy SW, Padmakumary G, et al (1990). Risk factors for cancer of the buccal and labial mucosa in Kerala, India. *J Epidemiol Comm Health*, **44**, 286-92.

Werness BA, Levine AJ, Howley PM (1990). Association of human papilloma virus types 16 and 18 E6 proteins with p53. *Science*, **248**, 76-79.

Zhang ZF, Begg CB (1994). Is Trichomonas Vaginalis a cause of cervical neoplasia? Results from a combined analysis of 24 studies. *Int J Epidemiology*, **23**, 682-90.