Detecting spatial clusters of *Taenia solium* infections in a rural block in South India

M. Venkata Raghava a,∗, V. Prabhakaran b, T. Jayaraman b, J. Muliyila, A. Oommen b, P. Dornyc, d, J. Vercruyssec, d, V. Rajshekhar b

a Department of Community Health, Christian Medical College, Vellore 632 002, Tamil Nadu, India
b Department of Neurological Sciences, Christian Medical College, Vellore, India
c Department of Virology, Parasitology and Immunology, Ghent University, Belgium
d Institute of Tropical Medicine, Antwerp, Belgium

**A R T I C L E   I N F O**

Article history:
Received 18 November 2009
Received in revised form 11 June 2010
Accepted 11 June 2010
Available online 17 July 2010

Keywords:
Cysticercosis
clusters
epilepsy
geographic information systems
taeniasis

**A B S T R A C T**

Neurocysticercosis (NCC) is a major cause of seizures/epilepsy in countries endemic for the disease. The objectives of this study were to spatially map the burden of active epilepsy (AE), NCC, taeniasis, seroprevalence for cysticercal antibodies and positivity to circulating cysticercal antigens in Kaniyambadi block (approximately 100 villages comprising 100 000 population) of Vellore district and to detect spatial clusters of AE, NCC, taeniasis and seroprevalence. Using geographic information system (GIS) techniques, all 21 study villages with over 8000 houses (population of 38 105) were mapped. Clustering of different indices of *Taenia solium* infection was determined using a spatial scan statistic (SaTScan). There was a primary spatial cluster of AE with a log likelihood ratio (LLR) of 10.8 and relative risk (RR) of 22.4; however, no significant clustering for NCC was detected. Five significant spatial clusters of seropositivity for cysticercal antibodies, two clusters of seropositivity for cysticercal antigens and one for taeniasis were detected (LLR of 8.35 and RR of 36.67). Our study has demonstrated the use of GIS methods in mapping and identifying ‘hot spots’ of various indices of *T. solium* infection in humans. This spatial analysis has identified pockets with high transmission rates so that preventive measures could be focused on an intensive scale.

© 2010 Royal Society of Tropical Medicine and Hygiene. Published by Elsevier Ltd. All rights reserved.

1. Introduction

Taeniasis is an intestinal disease acquired by ingestion of viable *Taenia solium* metacestode larvae from pork, while cysticercosis is acquired by the ingestion of *T. solium* eggs. *Taenia solium* infection is widely endemic in the rural areas of developing countries in Asia, Africa, South and the Central Americas. In Asia, the exact geographic origin and the epidemiological factors associated with transmission are biased since most available data are from hospital based populations.1

Human cysticercosis frequently affects the central nervous system and is a major cause of acquired epilepsy.2–4 In Asia and Africa, the subcutaneous form of cysticercosis which is concomitant with intracerebral infection is the most common variety and accounts for more than 30% of cases.5 Neurocysticercosis (NCC) is a major cause of seizures/epilepsy in countries endemic for the disease. In African countries, studies have shown that in about 30 to 50% of cases, epilepsy had been documented to *T. solium* infection.1 Earlier community based studies carried out in Kaniyambadi block (a block is a sub district level
organization of rural areas under the Community Development Programme of the Government of India and comprises of approximately 100 villages and 100,000 population), a rural development block in Vellore district of the state of Tamil Nadu in India have demonstrated active and continuous transmission of cysticercosis. This is evident from the high prevalence of NCC as a cause of over a third of all cases of active epilepsy (AE).6 The prevalence of taeniasis in the community, as diagnosed by using a coproantigen assay, was 30.2 per 1000 population (Raghava et al., unpublished data). A high seroprevalence of cysticercal antibodies at 17.7% in the seizure-free general population in the same community also indicates high exposure to cysticercal antigens.7 This provided an impetus for us to study the spatial distribution of the disease in this community.

In the last decade, geographical information systems (GIS) have provided a powerful tool to display epidemiological data in a spatial format in the form of maps. Not only is a spatial display of the distribution of a disease possible, but also it is possible to perform statistical analyses to determine whether there is a disproportionate concentration of disease in some pockets of the geographical areas surveyed. These clusters or ‘hot spots’ may provide a target for surveillance and control strategies. Studies using GIS have been done mainly to study clustering of patients with various types of cancers.8–13 A few studies involving infectious diseases such as cholera, malaria, leprosy, tuberculosis and giardiasis have been reported in recent years.14–22 Efforts are under way to develop GIS and remote sensing based models for various parasitic diseases including schistosomiasis, which are now being expanded to include widespread infectious diseases.23 Only a few studies have reported the application of GIS to study the clustering of cysticercosis or taeniasis. Lescano et al.24 used GIS to study swine cysticercosis risk gradient surrounding human tapeworm carriers in Peru. They also studied human cysticercosis seroprevalence gradient and cysticercosis related seizures surrounding tapeworm carriers using GIS techniques.25 The objectives of this study were: to spatially map the burden of AE, NCC, taeniasis and seroprevalence of cysticercal antibodies and positivity to circulating cysticercal antigens in Kaniyambadi block; and to detect and map spatial clusters of AE, NCC, taeniasis, seroprevalence of cysticercal antibodies and positivity rates for circulating antigens.

2. Methods

Kaniyambadi block is located in Vellore district of Tamil Nadu State in India, spanning a geographic area of 184 km² with a population of 106,000 (current updated census, 2006, of the Community Health Department of Christian Medical College (CMC), Vellore). There are 82 settlements distributed across 33 revenue villages (three or four settlements grouped for administrative and revenue purposes) in the block. The study area is bordered by Vellore town on the northeastern side, Anaicut block on the western side and Arani block on the southern part of the study area and situated between 12°40’N to 12°55’N latitudes and 79°0’E to 79°15’E longitudes (Figure 1). The adult literacy in this block is 79% and the primary occupation is agriculture. The population under study was selected from 21 randomly selected settlements from the block and is representative of the state of Tamil Nadu.
We used data from the following sources for our study:

Master Health Census collected with geo-locations sourced from Community Health Department of CMC, Vellore.
GIS data pertaining to the block, its villages, and people from the existing Geodatabase of the Department of Community Health, CMC, Vellore.

2.1. Study population

The study population covered 38,105 people between the ages of 2 and 60 yrs who had been screened for AE from 16 randomly selected rural clusters. Rural clusters in the study are village panchayats (settlements of people with local self-governance). Figure 2 shows the spatial location of all the village settlements and the study clusters in Kaniyambadi block.

2.2. Epidemiological survey

The survey methods and screening tools used for determining the AE rates have been published previously. Briefly, patients with AE were identified using a validated questionnaire administered, in a door-to-door survey, by health workers. Clinicians evaluated those identified in a hospital setting and those finally determined to have AE were advised to undergo a contrast enhanced CT scan after informed consent was obtained. All patients with AE undergoing CT scan were also tested for the presence of cysticercal antibodies and antigens in their blood using enzyme-linked immunoelectrotransfer blot assay (EITB) and ELISA techniques, respectively.

Detection of cysticercal antibodies in serum was by an EITB using lentil lectin specific *T. solium* glycoproteins as antigens, standardized in our laboratory. A sample was considered positive for cysticercal antibodies by the criteria of Tsang et al., i.e. reaction to one or more *T. solium* glycoproteins of molecular weights 50, 38–42, 24, 21, 18, 14 and 13 kDa.

Circulating *T. solium* metacestode antigens were assayed in all sera by an ELISA using monoclonal antibodies to excretory/secretory products of *T. saginata* metacestodes established by Brandt et al. and modified by Dorny et al. The ELISA is 94% sensitive for cysticercosis with no cross reactivity with sera of other parasitic infections. A sample was considered positive for cyst antigens above the Mean+3 SD absorbance of 6 negative control sera assayed with the study samples on each microtitre plate.

Each case in the block was linked with the master Census database using a unique identification number and the geographic coordinates (latitude, longitude) extracted from this database. Twenty households were randomly selected from each study cluster and 729 early morning stool samples were collected from the study population aged between 2 and 60 years and tested for the presence of coproantigens. The same households provided 960 serum samples which were tested for cysticercal antibodies and antigens to determine the prevalence rates. While the presence of cysticercal antibodies in the serum indicates exposure to the larval antigens but not necessarily infection, the presence of cysticercal antigens in the serum indicates the presence of a viable cyst in the patient and active infection which may or may not cause symptoms. Cluster analyses were performed using SatScan software.
Figure 3. The study methodology.

The study methodology was illustrated in Figure 3.

2.3. Statistical analysis

Data were entered in Microsoft Excel™ 2002 (Microsoft Corp. 1985–2001) and statistical analysis was done using Statistical Package for Social Sciences v. 12 (SPSS Inc., Chicago, IL, USA)33 software. Prevalence rates of AE, NCC, taeniasis and seroprevalence for cysticercal antibodies and rates of positivity to circulating cysticercal antigens were calculated for each of the clusters and at the block level.

2.4. Spatial mapping and analysis

Garmin GPS V (Garmin International Inc., Olathe, KS, USA) was used for collecting waypoints (latitude, longitude) of the people who did not have existing co-ordinates. These were then downloaded using GPS utility 4.10.4 (GPS Utility Ltd., Southampton, UK) and mapped using ArcView GIS 9.1 software. To explore the spatial distribution of AE, seroprevalence of cysticercal antibodies, positivity rates of circulating cysticercal antigens, taeniasis and to identify significant spatial clusters, if any, spatial scan statistic implemented in SaTScan 7.0.1 software program was used. The unit of analysis was considered to be a case of either AE or taeniasis or a seropositive individual.

SaTScan needs three input files: one population file, one case file and one coordinate file. The case file could be single or aggregated data. Thus, in our study the case file consisted of the number of active epileptics or patients with taeniasis in the study area, and thus the data is aggregated at the block level in our case. The population file carries the population of the block, and the coordinate file contains the coordinates of the individuals.34 SaTScan operates by passing a circle over the study region. The software moves a circular window systematically over a geographic area to detect significant spatial clusters. The radius of the window may vary from zero to
Figure 4. (a) Prevalence of Active Epilepsy in the study clusters. (b) Prevalence of Neurocysticercosis in the study clusters. (c) Prevalence of positivity rates to serum cysticercal antigens in the study clusters. (d) Prevalence of seropositivity rates to cysticercal antibodies in the study clusters. (e) Prevalence of taeniasis in the study clusters.

A user-defined upper limit. The centroids (mathematical or geographical center point of a polygon or the midpoint of a line, described as an x, y coordinate) of some individuals will lie within the circle. SaTScan tests if the number of cases in the 'within circle' region outstrips the expected number of cases, within that region. It increases the size of the circle until the user-defined threshold is reached. The circle moves from one centroid to another until all the centroids within the study region are covered. Once this testing is done, generating a large number of random permutations of the dataset tests the test statistic.
The cluster assessment is performed by comparing the number of cases within the window with the number expected, if cases are randomly distributed in space. The test of significance of the identified clusters is based on a likelihood ratio test whose $P$-value is obtained through Monte Carlo testing.\textsuperscript{34,36,37}

Identification of spatial high clusters was done under the Bernoulli probability model assumption using a maximum spatial cluster size of 5% of the total population. For statistical inference, 999 Monte Carlo replications were performed. The null hypothesis of no clustering was rejected when the
simulated $P$-value was $\leq 0.05$ for the primary clusters.

3. Results

3.1. Prevalence of AE, and NCC, taeniasis and seropositivity for cysticercal antigens and antibodies

The prevalence of AE in the surveyed population was 3.04 per 1000 (116 cases out of a total population of 38 105). The prevalence of AE in the individual clusters ranged from 0.58 to 6.59/1000 population. Out of the 116 patients with AE, a contrast enhanced CT scan was performed in 87% of the patients (101/116). In the remainder, these tests could not be done for various reasons, the most common of which was refusal (6/15). The other reasons included severe physical or mental disability, patient working or studying out of town, pregnancy, and history of severe head injury. Using criteria described by del Brutto et al. for the diagnosis of NCC, the prevalence of NCC in the study population was 1.02 per 1000 population (39 cases out of a total population of 38 105). A diagnosis of NCC was made in 39 (38.6%) of the 101 AE patients undergoing a CT scan. The most common lesion identified on the CT scan was a solitary calcification; one or two calcifications were noted in 81.3% of the scans positive for NCC. The prevalence of taeniasis in the community was 30.2 per 1000 population (22/729 samples). The population seroprevalence of cysticercus antibodies was 19.2% (184/960) and the positivity rate of cysticercal antigens was 4.8% (46/960).

3.2. Spatial distribution

Prevalence rates for AE varied within the study clusters and some of the clusters with high AE rates over 3.04/1000 population were Salamanatham, Sathupalayam, Adukamparai, Pennathur and Virupakshipuram as shown in Figure 4a.

The prevalence rates for NCC in the study clusters are shown in Figure 4b. Three clusters Salamanatham, Munjurpet and Mottupalayam had higher rates of both AE and NCC as compared to other clusters.

Two villages Nanjukondapuram and Kammasaudram had the highest seroprevalence rates for cysticercal antigens, measuring over 160 per 1000 population (Figure 4c) and Salamanatham village had the highest seroprevalence rate for cysticercal antibodies of 390/1000 population (Figure 4d).

The clusters with comparatively higher rates of taeniasis in the study area were Mottupalayam, Sathupalayam, Pennathur, Sathumadurai and Adukamparai as shown in Figure 4e.

3.3. Spatial clustering

Cases of AE were distributed throughout the study area. GIS analysis revealed a significant high primary spatial cluster of AE located at 12.8129N, 79.1347 E in Kaniyambadi village with a log likelihood ratio (LLR) of 10.8, relative risk (RR) of 22.4 and P value = 0.048 as shown in Figure 5. There were no significant spatial clusters noted for NCC in the study area.
Spatial analysis revealed five significant spatial clusters for seropositivity to cysticercal antibodies. The most significant one was located in Singirikoil village with a location id of 12.7641 N, 79.1236 E, LLR of 11.5, RR of 4.28 and $P$ value = 0.007. The other significant clusters were located at Nanjukondapuram, Kammavanpet, Sholavaram and Dhar-mavaram villages as shown in Figure 6.

Two ‘hot spot’ clusters were identified for positivity to circulating cysticercal antigens located at Kammasaudram village with a location id of 12.7770 N, 79.1849 E, LLR of
16.7, RR of 8.1 and \( P \) value = 0.001 and at Nanjukondapuram village with a location id of 12.7588 N, 79.0850 E, LLR of 13.0, RR of 14.7 and \( P \) value = 0.002 as shown in Figure 7.

One ‘hot spot’ cluster of taeniasis was observed in Kilvallam village with a location id of 12.7545 N, 79.1521 E, LLR of 8.35, RR of 36.67, \( P \) value of 0.02 as shown in Figure 8.

4. Discussion

4.1. Geographical information systems (GIS)

Epidemiologists have long used maps to track the spread of disease, and in the past decade, GIS technology has added a powerful new tool that helps reveal far more than simply the ‘where’ and ‘when’ of epidemics. Recent advances in GIS have allowed the application of disease mapping and spatial analysis, such as spatial clustering and cluster detection in epidemiological research. There are several important consequences of detecting clusters of diseases. Efforts can be focused on these clusters to study the influence of potential causative factors of a disease or the modes of transmission of infectious diseases. The application of GIS techniques in the field of health and disease mapping has been on a gradual rise, however, its use with respect to neglected diseases such as cysticercosis has been minimal.

4.2. Utility of this study

The results of our study demonstrate the application of GIS techniques in mapping the indices of \( T. solium \) infection. It could be seen that even in a small geographic study area the prevalence of the various indices vary considerably between the individual villages. Some of the villages with higher rates for seroprevalence of cysticercal antibodies, which indicate exposure either active or due to past infection also demonstrated a higher positivity rates for cysticercal antigens, indicating that the exposure seems to be from a continuous source and could pose a hazard to the people living in these clusters. Even though NCC contributed to 38.6% of AE in this region, there were no significant clusters of NCC while significant spatial clustering of AE was observed indicating that in this region, other causes of AE could be important.

Villages with high prevalence rates for taeniasis as estimated by the coproantigen assay were geographically distributed over most of the study area and two villages also had higher rates of AE but not of NCC. An interesting observation from our study is that the clusters of positivity for circulating cysticercal antigens were different from that of taeniasis. Possible reasons for this anomaly are that taenia carriers or carriers living in the vicinity of subjects positive for cysticercal antigens have been missed in the sampling process or that the individuals who are positive for circulating cysticercal antigens acquired their infection from outside their area of residence, possibly while traveling.

In the study area and neighbouring blocks, the prevalence of porcine cysticercosis as estimated by testing porcine sera for antibodies is 9.7% (Oommen et al., unpublished data). Over the last 5 years or so, there has been a decline in pig rearing in this region mainly due to the banning of unorganized pig rearing by the state government, and periodic state organized campaigns to kill free roaming pigs whenever there are reports of small outbreaks of
Japanese encephalitis in the district. However, free roaming pigs are still visible in most villages. The vast majority of the people in this region belong to the lower and middle socioeconomic strata and open field defecation is widespread. This combined with poor sanitary and personal hygiene provides a good setting for the transmission of cysticercosis.

4.3. Comparisons with other GIS studies on cysticercosis

Lescano et al.\textsuperscript{24,25} have recently demonstrated the utility of GIS techniques in studying the clustering effect of cysticercosis. They found that while seroprevalence rates were significantly correlated with the distance from a taenia carrier, there was no such clustering of NCC related seizures cases. Seroprevalence rates were highest among the household members of the taenia carrier and decreased in households within 1 to 50 metres of the carrier’s house. It declined further when the households were > 50 metres from the taenia carrier’s residence.\textsuperscript{25} Interestingly, no such gradient was found with respect to NCC related seizure rates. The authors have three main explanations for this discrepancy: the life expectancy of a tapeworm, the latency between infection by the parasite and onset of seizures, and possible migration in and out of the region of the carrier and patients with seizures.

Our study was focused more on the determination of clustering of the different manifestations of the parasitic infection in humans rather than on elucidating a correlation between them. However, the presence of distinct and spatially separated clusters of taeniasis, high prevalence rates of cysticercal antibodies and antigens and AE suggests that the transmission of the disease in our community might be different from that which prevails in Peru. The possible explanations for differences in transmission dynamics could include differences in defecation practices, pig rearing practices, food habits and sources of water for human consumption.

Our finding of a lack of clustering for NCC related AE cases is similar to that reported by Lescano et al.\textsuperscript{25} This reinforces the suggestion that patients with NCC related AE need not always be in the vicinity of a taenia carrier. As speculated by Lescano et al.,\textsuperscript{25} it is possible that patients with NCC related AE might have acquired the disease elsewhere or that the adult worm within the taenia carriers from whom the infection was acquired has died and hence these carriers were not diagnosed at the time of the survey. The latter hypothesis is based on the premise that the life span of the adult worm is around 3–4 years, the time period that is usually required for a patient with NCC to manifest clinically with seizures.

There are several difficulties in studying the relative clustering of \textit{T. solium} infection; while some are imposed by the life cycle of the parasite others stem from human activities such as migration. Some of these issues, such as the 3–4 year life expectancy of the adult worm, have been discussed above. The transient nature of antibody response and even the antigen presence in the serum of those exposed to the parasite eggs is yet another confounder.\textsuperscript{41}

4.4. Limitations of the study

The addition of geographical features such as the elevation of villages, the direction of flow of rivers or streams might have provided additional information. Since the study was conducted in a relatively small area, weather...
patterns are unlikely to have played any role in disease transmission. However, environmental factors such as the relative location of grounds used for defecation and the source of water supply could have possibly provided clues to disease transmission.

5. Conclusions

GIS has provided a map of the distribution of cysticercosis and taeniasis in Vellore District and enabled identification of clusters of increased exposure to the disease. However, transmission dynamics that resulted in such clustering has not been elucidated. It is possible that additional data input as mentioned above might help identify modalities of transmission of the disease in the community. Our study has nonetheless demonstrated that using existing health data, GIS and spatial scan statistic could provide public health officials with additional tools necessary for disease surveillance and to identify pockets with high transmission rates so that preventive measures could be focused on an intensive scale.

Authors' contributions: VR, JM, AO designed the study and drafted the manuscript. VP and TJ carried out all laboratory work. MVR carried out the community studies, GIS analyses and also helped in drafting the manuscript. PD, AO and JV helped in the analyses and drafting of the manuscript. All authors analyzed the data, read and approved the final manuscript. AO, VR and JM are guarantors for the paper.

Acknowledgements: The authors would like to thank Mr Vijayakumar and Mr Yesudass for their help in conducting the field studies.

Funding: The study was supported in part by a Task Force grant from the Indian Council of Medical Research, New Delhi and the Flemish Interuniversity Council, Belgium.

Conflicts of interest: The authors declare no conflict of interest.

Ethical approval: The study was approved by the Institutional Review Board of the Christian Medical College, Vellore.

References