Investigating Geographic Distribution of Colorectal Cancer Cases: An Example from Penang State, Malaysia

Narimah Samat1, Aishah Knight Abd Shattar2, Yasmin Sulaiman3, Azizah Ab Manan3 & Chan Ngai Weng1

1 School of Humanities, Universiti Sains Malaysia, Penang, Malaysia
2 Advance Medical and Dental Institute, Universiti Sains Malaysia, Penang, Malaysia
3 Penang Cancer Registry, State Health Department, Komtar, Penang, Malaysia

Correspondence: Narimah Samat, School of Humanities, Universiti Sains Malaysia, 11800 Penang, Malaysia. Tel: 60-12-501-8834. E-mail: narimah@usm.my

Received: July 9, 2013   Accepted: September 11, 2013   Online Published: November 29, 2013
doi:10.5539/ass.v9n17p38          URL: http://dx.doi.org/10.5539/ass.v9n17p38

Abstract

The map has widely been used to depict and disseminate information on the spatial distribution of various diseases, including cancer cases. Increasingly, population-based cancer data need to incorporate spatial information that enables various spatial and proximity analyses to be conducted whereby results can be graphically presented. Yet, disease maps as a communication form remain largely unexamined. This is probably due to the confidentiality of the disease cases and the cost of incorporating the spatial component in the database. In the state of Penang, Malaysia, although Penang Cancer Registry (PCR) collects and collates data of all cancer cases diagnosed in the state as well as cancer cases diagnosed elsewhere whose home address is given as a Penang address, geographical location is not included. Mapping of cases using information from PCR will give a fairly complete picture of spatial distribution of cancer cases from Penang State and clustering of cases can be readily evaluated. This study demonstrates the application of spatial analysis methods and GIS in mapping and understanding the spatial distribution of Colorectal cancer cases in Penang State. The cases were mapped to identify spatial clustering of cancer cases and measure distance from existing health facilities. This study finds that spatial information is pertinent to be included as part of the database kept by Cancer Registry since this information can effectively be used for communications with and education of the public, as well as for planning health care delivery.

Keywords: colorectal cancer, point-pattern analysis, geographic information system(GIS), disease mapping, penang state, Malaysia

1. Introduction

Cancer is a major health problem in many countries. It is a leading cause of death worldwide, accounted for 7.4 million deaths (around 13% of all deaths) in 2004. Furthermore, the deaths from cancer worldwide are estimated to reach 12 million deaths in 2030 where lung, stomach, liver, colorectal, esophagus and prostate are common types of cancer among men, while breast, lung, stomach, colorectal and cervical cancers are more frequent among women (WHO, 2009). In Iran, Babaei et al. (2009) studied cancer incidence and mortality in Ardabil and found that this area has one of the highest rate of gastric cardia cancer in the world with more than 4300 new cases during the 3 years of the study. Jafari et al. (2010) studied cancer mortality by evaluation of asbestos fibers concentrations in an asbestos-cement products factory in Iran and found that asbestos mills have the highest mortality rate, with an expected 1198 deaths per 100,000 workers after one year exposure and 14,665 deaths per 100,000 workers after 20 years of occupational exposure. In Malaysia, for example, the incidence of cancer is approximately 30-40 thousand cases per year (Ministry of Health, 2000). In 2006, for example, there was 3525 female breast cancer cases registered with the National Cancer Registry, which accounted for 16.5% of all cancer cases registered (Omar et al., 2006). Colorectal cancer, colon cancer or large bowel cancer includes cancerous growths in the colon, rectum and appendix accounted for 14.2% of male cancers making it the commonest cancer among men. This cancer accounted for 10.1% of female cancers making it the third most common cancer among women (Lim, Yahya & Lim, 2003; Lim & Yahya, 2004). The statistics from National Cancer Registry shows an alarming rate of cancer cases in Malaysia. It is, however, viewed in tabular form which limits its ability to be analyzed spatially such as in assessing clusters of cases to look for possible aetiological factors or planning...
of health services for screening or treatment.

According to National Cancer Society (2006), 80% of cancer cases are curable if detected early. However, most cancer cases in Malaysia were detected at stage III, which makes it difficult to achieve a cure. This is probably due to low awareness among the population regarding cancer screening or low accessibility of the people to screening facility. Since cancer has been a major cause of death in many countries, it is important to improve population health by evaluating the spatial differences in the distribution of cancer cases, mapping possible cluster of cases, identifying possible factor that might be the cause of cancer, and addressing strategies to improve health facilities (Rushton, 2007; Abdulkader, 2007). Geographical Information Systems (GIS) have widely been used in many developed countries to examine the spatial pattern of disease and analyze the accessibility to primary and secondary health services (Rosero-Bixby, 2004; Abdulkader, 2007). The analytical capabilities of GIS can possibly be used to examine cancers incidence and geographic region in order to evaluate geographic variation and pattern of specific cancers (Moore & Carpenter, 1999; Rushton, 2007). In addition, this technology can be utilized to find association between cancer incidence and health infrastructures and identify places where cancer surveillance and control programs are needed and evaluate the accessibility of health screenings and treatments (Higgs et al., 2005; Ghetian et al., 2008).

Various studies have been undertaken that used GIS technology and spatial data analysis in evaluating health data (Levine et al., 2009; Rushton et al., 2006). Spatial analysis is a useful tool to explore health data, map and identify patterns, generate new hypothesis, and provide evidences about existing hypothesis (Boscoe et al., 2004). In the western developed nations, data from cancer registry can be fully used to find association between occurrence of cancer incidents with geographical and demographic factors (Higgs et al., 2005; More & Carpenter, 1999). Such analysis could easily be performed where cluster could be identified and correlated the cases with geographical or demographic factors (Abdulkader, 2007).

GIS applications have been used to improve researches and services of healthcare. The integration of these applications with health data could bring many benefits, such that GIS analytical functions could be used to combine data from various sources and format, visualise and analyse spatial pattern of incident of diseases, develop and model a risk map, and generate report which identify areas for improvement of health services and diseases prevention with all kinds of support decision making (Richards et al., 2008, 1999). Long Island Geographic Information Systems (LI-GIS) is one of the applications which was developed by National Cancer Institute for Breast Cancer Studies in Long Island, New York, USA. The application able to hold more than 80-in depth data set with high quality of 1990’s that related with Breast cancer epidemiology, Topography, Demography, Environmental data, and incidents data. NCI made the application accessible for researchers and it can be extended to be applied on any other disease (National Cancer Institute, 2010). Another small module was developed by the World Health Organization (WHO) called Access Mod which works on ArcView 3.3, was used to analyse the data. The tool is open source to be used by developing and poor countries to calculate the physical accessibility to health care facilities, supported with full information of installation and usage (Black et al., 2004).

In Malaysia, on the other hand, health data is not easily accessible due to confidentiality of the data (Samat et al., 2010). Therefore, it curtailed the ability to understand spatial relationship between occurrence of cancer and demographic characteristics. Therefore, only basic analysis could be performed in order to visualize the clustering of cases in order to find association between cases and population density. Such analysis would reveal spatial characteristics of cases and potentially be used in mobilizing effort towards creating public awareness regarding the requirement for screening and early detection strategy. This study, therefore, aims to map the incidence of colorectal cancer cases in Penang State. It is used to evaluate spatial distribution of the cases and identify cluster of cases and measure the distance from population centres’ to existing hospitals. Such information is hoped to reduce the mortality for colorectal cancer by directing the effort to early detection at areas with significantly high number of cases.

2. Materials and Methods

The study was undertaken in Penang State, located in the north western part of Peninsular Malaysia between 5°8’ and 5°35’ latitude and 100°8’ and 100°32’ longitude is a small area that covers approximately 103,938 hectares, consisting of Penang Island and Seberang Perai on the mainland (Figure 1). Penang state is one of the urbanized states in Malaysia where urban population is approximately 87% of 1,313,449 people in 2000. The state population increased to 1,546,800 in 2008 (Statistics Department, Malaysia, 2000; 2008). Because of its highly urbanized characteristic and dense population, health facilities in Penang State are quite good. In 2004, the number of government hospitals and private hospital are 6 and 11 respectively. The main reason for choosing
Penang State for the study area is because of data availability. PCR keeps records of cancer cases, whose home addresses are in Penang State since 1994. This complete dataset is useful in mapping and understanding the distribution of colorectal cancer cases in Penang State. In addition, the Geography Section of Universiti Sains Malaysia has a collection of database for Penang State, which could be used in the analysis. Database building can be costly and time consuming (Longley et al., 2005). Therefore, existing digital data is used wherever possible to reduce the cost of database development. Finally, researchers are quite familiar with the study area, which makes it much easier to locate addresses obtained from PCR.

Physical accessibility of patients to health facilities such as hospital is very important in order to ensure treatment can be undertaken regularly (Schuurman et al., 2009). At present, addresses of the patients are recorded but this information has not been used in planning or allocating new health facility. This study, therefore, attempted to map the distribution of colorectal cancer cases in Penang State, evaluated the distribution of cases in relation to population of the sub-districts, and accessed the accessibility of patients to the nearest hospitals. The study began by getting ethical approval from the Research Ethics Committee (Human) of Universiti Sains Malaysia which is in compliance with ICH GCP guidelines (USMKK/PPP/JEPM[204.3 (8)]). This study was also registered with the National Medical Research Register (NMRR-08-1336-2734).

Then, the study acquired addresses of patients, generalized to maintain the confidentiality of the patients from Penang Cancer Registry (PCR). It was recorded from 1994 to 2003 in EXCEL format. Global Positioning System (GPS) was used to acquire geographic coordinate of the addresses. In identifying the locations, various approach were used such as Google Earth software which was used prior to undertaking the fieldwork, asking the postmen in the study area and getting help from the Department of Town and Country Planning, Penang State. In addition to the location of cancer cases, the geographic coordinates of all health facilities namely private hospitals, public hospitals, public health clinics, private health clinics and state health offices were recorded. After that the data were loaded into ArcView 3.2 and converted into ArcGIS 9.3 format. Secondary data such as roads, land use, slope and population were also gathered from other agencies in Penang State. The study mapped the cases in each sub-district or mukim, identified cluster of cases or hotspot using Getis-Ord GI* statistics and evaluated accessibility of each mukim to the existing health facilities. Maps were produced to communicate the information regarding the distribution of cases such that it allows practitioners and policy makers to better view the distribution of cases.
3. Results

Results from this study showed that between 1994 and 2003, there were 1543 colorectal cancer cases in Penang State. However, only 1266 (82.04%) cases were collected using GPS and mapped into ArcGIS 9.3 software. About 17.86% cases could not be identified due to incomplete addresses. The spatial distribution of cases is shown in Figure 2(a) below. This figure illustrates that colorectal cancer cases concentrated around Georgetown and the surrounding areas and a few clusters can be detected throughout the map. Small concentration of cases can be seen in Batu Feringghi, Tanjung Tokong and Sungai Dua in the Timur Laut District. Surprisingly, not many colorectal cancer cases can be seen in the Barat Daya area. Only a small cluster can be seen in Teluk Kumbar and Batu Maung area. In the Seberang Perai region, the major concentration cases can be seen in Butterworth and Bukit Mertajam area. Small cluster can also be seen in Kepala Batas and Telok Air Tawar in the North of Seberang Perai and a few cases scattered in the South of Seberang Perai.

The study, then, evaluated the concentration of colorectal cancer cases with population of the Mukim or sub-district. It was found that percentage number of cases compared to population is not really big. The result is shown in Figure 2(b) below. The highest number of cases can be found in Georgetown, Mukim 17 and Mukim 18 of Timur Laut District that is between 0.16 and 0.30 percent of its populations. Other areas that has quite a high number of cases is Mukim 8 of Barat Daya District, Mukim 7 and Mukim 10 of the North of Seberang Perai and Mukim 5, Mukim 7, Mukim 8, Mukim 9, Mukim 15 and Mukim 18 of the Middle of Seberang Perai where each has between 0.10 and 0.15 cases. Other areas have slightly lower number of cases as compared to its population. The study found that higher number of cases can be found in the north of Timur Laut District, this distribution, however, was probably due to the accessibility of people in this area to health care facilities which made it easy for people to undertake screening for this type of cancer.

Results showed that when the density of colorectal cancer cases within square kilometre of the state was evaluated, the Eastern and Southern areas of Georgetown have about 30 to 35 cases per km² and are among the highest in the state. The cases are gradually decreased to 25 to 29 cases/km² at the outer border of the sub-district until become to 5 to 10 cases/km². Northern area of Mukim 18 has 20 to 25 cases in one km² as highest recorded clustered. The cases decrease to 10 to 15 cases/km² and 5 to 9 cases/km² in the same area in east of Mukim 18, east and partial southern of Mukim 13 in Timur Laut in Penang Island, and Mukim 14, Mukim 8 in Daerah
Tengah of Seberang Prai. Elsewhere, the density of cases was between 0 and 4 cases/km² (refer to Figure 3). The study also mapped the cluster of the cases among the sub-districts by using Hot Spot analysis to calculates the Getis-Ord Gi* statistics (Figure 3). It allows the cluster of cases to be detected. The study found that George Town has the highest with 5.378 of z value, which means there is a high number of occurring cases surrounding that area within 2-5 km². Mukims 18 and 16 comes after with 4.468 and 3.648 of z value (Figure 4). Referring to the graph which shows the significant z value (Figure 3 (b)) that lead to the null hypothesis of geo-sources affects existence if the z values were in range of -4 to 4. But, z values results in the study leads to another process to figure out what might cause the statistically significant clustered cases in those areas. This information is useful for health practitioners or policy makers in planning for health care delivery or undertaking screening program in Penang State. At present, however, this study has no information regarding the gender, ethnics or stages of disease. Therefore further analysis on trend or distribution based on gender, ethnics or number of yearly occurrence could not be conducted.

Figure 3. Colorectal cancer cases density per square kilometers
In addition to spatial distribution of cases, the study also evaluated the accessibility of cases to health facility particularly hospital. The study also evaluated the proximity of population to existing both private and public hospitals in Penang Island. The location of the centre of the population was derived from centroid of the mukims. Although this location was probably not the actual centre of the population, it would give an overview on spatial distribution of population in the state. Euclidean distance between centroid of the mukims and hospitals was calculated using ArcGIS 9.3 software as shown in Figure 5. It was found that most part of the state was not far from existing hospital. Many mukims were located between 0.5 km and 2.6 km from the nearest hospitals. The farthest mukims was only between 9.2 km and 12.4 km away from the nearest hospitals. However, when compared the distance of the hospitals and existing colorectal cancer cases, it was found that a few cases in the north of Seberang Perai and Tasek Gelugor area was more than 6 km away from hospitals. This study, however, has not considered the location of hospitals in the neighbouring state of Kedah and Perak. In Penang Island, most of the mukims were less than 10 km from existing hospital except for Mukim 1 in the north west of the island. This area, however, is covered with forest reserve with approximately 5500 people only.

Figure 4. Spatial distribution of colorectal cancer cases in Penang state
4. Discussion

This study is an early attempt to understand the spatial arrangement of cancer cases particularly colorectal cancer with the hope to gain critical insights into the nature of the cases and help to plan for ensuring health facility is accessible to the patients. From the analysis undertaken it was found that high concentration of cases can be found in major town centres such as Georgetown, Air Itam and Bayan Baru in Penang Island and major concentration of cases in Butterworth and Bukit Mertajam. Higher concentration of colorectal cancer cases in major town centres probably due to accessibility of the population to screening facilities and the highest concentration of Chinese population in town centers of Penang State as compared to Malays and Indian. As stated by Lim and Halimah Yahaya (2004), the number of colorectal cancer cases was the highest among Chinese population as compared to Malays and Indian in 2002 and 2003.

At present, this study only mapped spatial distribution of colorectal cancer cases. It would be beneficial if the distribution of cases could be correlated with eating habits of the people. However, this study allows for the examination of the causes and effects of Colorectal cancer cases from multiple approaches, giving new outlooks into health issues (Moore & Carpenter, 1999; Black et al., 2004). GIS provides the opportunity to revisit methods of spatial analysis through the use of spatial data integration and visualization capability which might be useful for planners and health practitioners in planning and evaluating the location of health facilities in the region. It was found that GIS was very effective in depicting the spatial and temporal patterns of cancer cases particularly colorectal cancer as these patterns can now be used to plan for ensuring health facilities and accessibility to the patients.

There are some limitations in this study. Mapping of colorectal cancer cases is one of the examples used to evaluate the distribution of cases over geographic space. Although the study managed to visually and statistically analyze cluster distribution of cases and disseminate the result to those that can take action such as by planning and improving health services, not much conclusion can be made regarding the incidents. The spatial analysis of cancer data poses unique challenges because most cancers develop over a period of 20 to 30 years and are a result of multiple exposures interacting with the individual's genetic susceptibility (Pickle et al., 2005). Furthermore, in the study of colorectal cancer cases distribution, eating habits data was absent which might be useful covariates relating to cases.

To obtain useful spatial and temporal data is challenging in conducting health related research. Health data that is required for analysis are typically scattered across many sources and often collected by different groups and agencies. For example, PCR collated data on ascertain cases whose home address is Penang State. Such records collected for clinical purposes also rarely include demographic and geographic information desirable for the data analysis. Furthermore, in protecting the privacy and confidentiality of the patients, collection agencies and
medical facilities are imposing increasingly strict requirements for data release and often only identify a place (usually patient's address) to a broad administrative unit. The reportable specificity of location is often not good enough to allow the analysis to answer research questions about the spatial patterns of the disease. Methods are currently being explored that would allow use of specific individual information in the analysis but would mask identifying characteristics in the results reported only at an aggregated level.

5. Conclusion

The study mapped spatial distribution of colorectal cancer cases in and determined the cluster of cases the accessibility of cases to health facilities and calculated the density of cases in Penang State, Malaysia. It was found that major concentration of cases could be found in major town centres of Georgetown, Bayan Baru, Butterworth and Bukit Mertajam. Furthermore, most of the cases were accessible to hospitals or other health facilities in Penang state. The application of GIS in mapping the distribution of cases helps in identifying clustered of cases and calculating the accessibility of cases to health facilities. The findings from this study are useful for planners and health practitioners in understanding the pattern and visualising the hot spots of colorectal cancer cases. It is beneficial in planning as more efforts can be directed towards controlling and providing support services for patients in the hot spots. In addition, by incorporating spatial information into the health data system, it would give new insight into looking at health data and allow data to be visually analyzed.

Acknowledgements

Authors wish to thank Universiti Sains Malaysia for funding this research under Research University Grant, Penang Cancer Registry, located at Penang Health Department (Ministry of Health, Malaysia) for giving permission to used data for the study and Town and Country Planning Department, Penang for providing spatial data used in the study.

References


Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (http://creativecommons.org/licenses/by/3.0/).