SPATIAL EPIDEMIOLOGY: GEOSTATISTICAL TOOL FOR
DISEASE MAPPING

1Dr. Dipayan Datta, BDS*, 2Dr. S. G. Ramesh Kumar, MDS, 3Dr. M. B. Aswath Narayanan, BSc, MDS, 4Dr. A. Leena Selvamary, MDS and 5Dr. A. Sujatha, MDS

1Postgraduate, Department of Public Health Dentistry, Tamil Nadu Government Dental College and Hospital, Chennai- 600 003.

2Assistant Professor, Department of Public Health Dentistry, Tamil Nadu Government Dental College and Hospital, Chennai- 600 003.

3Professor and Head, Department of Public Health Dentistry, Tamil Nadu Government Dental College and Hospital, Chennai- 600 003.

4,5Assistant Professor, Department of Public Health Dentistry, Tamil Nadu Government Dental College and Hospital, Chennai- 600 003.

ABSTRACT
Spatial Epidemiology is the description and analysis of geographic variations in disease with respect to demographic, environmental, behavioural, socioeconomic, genetic and infectious risk factors. Currently the use of Geographic Information System (GIS), Global Positioning System (GPS) and Remote Sensing is increased in public health research in most of the developed countries. Advances in GIS and statistical methodology together with the availability of high-resolution, geographically referenced health databases present unprecedented new opportunities to investigate the environmental, social and behavioural factors underlying geographic variations in disease rates at small-area scale. In the future, developments in exposure modelling and mapping, enhanced study designs, and new methods of surveillance of large health databases promise to improve our ability to understand the complex relationships of environment to health.

KEYWORDS: Spatial Epidemiology; Geographic Information System; Disease mapping; Public health research.
INTRODUCTION
The term “spatial” in geography means pertaining to space on the earth’s surface. Hence, spatial variation is the variation across the global landscape usually associated with population. In health related surveys, demographic and socio-economic variables vary from one place to another. These spatial variations are potent enough to influence the risk of different diseases in different regions. The observed differences in the risks are likely to be confounded by these variables. Therefore, adjusting these confounders is important in the evaluation of the spatial variation in mortality and disease rates. Comparison between places, together with comparisons between times and between individuals, is a useful means of formulating and testing aetiological hypotheses.\[1\]

Spatial analysis can be described as a collection of techniques to analyze the geographical events where the results are dependent upon the spatial arrangement of the events. The spatial analysis includes both statistical and non-statistical methods, and starts with the application of exploratory techniques to find a good description of the data. The most important feature of spatial statistical analysis is that the places, where the events took place, are presented in the analysis in an explicit manner.\[2\]

The application of spatial analysis in the context of epidemiological surveillance and research has increased exponentially.\[3\] Currently the use of Geographic Information System (GIS), Global Positioning System (GPS) and Remote Sensing is increased in public health research in most of the developed countries. Besides, geographers have also started collaborating with public health researchers, helping this new discipline of health geography to grow using the geographical concepts and technologies to study different health-related situations.\[4\]

The central point of interest in most epidemiological studies is the identification of factors increasing or decreasing the individual risk of disease as observed in populations. In spatial studies, the central objective is focused to explore spatial variations in the risk of disease in order to identify locations (and more importantly, individuals) associated with higher risk of the disease.\[5\] Therefore, Spatial Epidemiology is the description and analysis of geographic variations in disease with respect to demographic, environmental, behavioural, socioeconomic, genetic, and infectious risk factors.\[6\]
BRIEF HISTORY

Literature in medical geography can be found in several ancient civilizations, including China, Greece and India. The earliest documented work was perhaps done by Hippocrates in the 5th Century BC, who described the relationship between health and the environment.\[7\]

The first disease map was created by German physician Leonhard Ludwig Finke (1792).\[8\] John Snow (1854) described the disease cholera with the help of the map of London and showed its association with contaminated drinking water.\[9\]

Spatial epidemiology underwent a transition in the late 19th and early 20th centuries from a largely descriptive science to an analytical science with the introduction of improved mechanical equipment and the wider availability of base maps on which statistical information are being plotted.\[10\]

Now-a-days, medical geographers assess the spatial distribution and accessibility of health services using GIS. The first true operational GIS was developed by Dr. Roger Tomlinson (1962). Environmental Systems Research Institute (ESRI) released ARC/Info and ARC View software in 1981 and 1992 respectively. By the end of 20th Century, the development of ARC View enabled viewing GIS data through internet.\[11\]

DATA REQUIRED IN SPATIAL EPIDEMIOLOGY

Two main types of spatial data are used in spatial epidemiology: point data and area data. Each item of health data (including population, environmental exposure, mortality and morbidity) may be connected with a point, or precise spatial position such as a home, a street address or an area, which could be defined as a spatial region by postcode, ward, local authority, province and country.\[6\]

For spatial analysis in health care surveys, data are collected from different source. But these data are often handled without taking into account the interests of the geographical epidemiologists. Therefore, it is extremely important to ensure that complete and accurate point and/or area health data should be used in spatial epidemiology.\[12-14\]

TYPES OF SPATIAL EPIDEMIOLOGIC STUDIES

Studies using spatial epidemiology can be described into 3 major fields\[6\]:

1. Disease Mapping
2. Geographic Correlation Studies
3. Clustering and Surveillance

1. Disease Mapping: Disease maps establish a visual summary of complex geographic information. They can identify subtle patterns in the data that are not given importance in tabular or other form of presentations. Disease mapping is used mainly for descriptive purposes, to generate hypotheses regarding the etiology, for surveillance to highlight those areas which are apparently at high risk, and to aid policy formation and resource allocation. They are also useful in placing specific disease clusters and results of point-source studies in proper context.\textsuperscript{[15]}

The use of mapping in the medical context has developed so rapidly during recent decades that the presentation of maps is now established as a basic tool in the analysis of public health data.\textsuperscript{[16]} There are also other important issues that need to be considered while creating a map. These include the selection of an appropriate administrative unit for mapping, the selection of an appropriate method of data classification in the map and the selection of an appropriate colour scheme or collection of hatching patterns.\textsuperscript{[17,18]}

2. Geographic Correlation Studies: The aim of the geographic correlation studies is to assess the geographic variations across different population groups in exposure to environmental variables (which may be measured in air, water, or soil), socioeconomic and demographic measures (like race, income etc.), or lifestyle factors (like smoking, diet etc.) in relation to health outcomes measured on a geographic scale. This approach often takes advantage of data that are routinely available and can be used to investigate natural experiments where the exposure has a physical basis (e.g, soil, water).\textsuperscript{[19]} For example, Investigation of a regional excess of sinonasal cancer was consistent with studies in other countries showing risks associated with working in the furniture industry\textsuperscript{[20-22]} and study of local lung cancer excess was associated with residence near or employment in the arsenic industry\textsuperscript{[23]}

3. Clustering and Surveillance: Searching for disease clustering involves an assessment of local or global accumulation of disease. There are different types of clustering, including general and specific. In general clustering, the analysis of the overall clustering tendency of the disease incidence is done in a study region. The specific clustering is designed to examine the exact location of the clusters.\textsuperscript{[24]}
Surveillance, or the systematic routine collection and analysis of health outcome data for disease prevention and control purposes can be applied to the problem of disease clusters through the use of space, time and space-time pattern detection methods.\cite{25-27}

**GEOGRAPHIC INFORMATION SYSTEM (GIS)**

The Geographic Information System (GIS) is the computer software for data capturing, thematic mapping, updating, retrieving, structured querying and analyzing the distribution and the differentiation of various phenomenon including communicable and non-communicable diseases across the world with reference to various time periods. It has also been used for studying and mapping the surrogate information relevant to the environments of the disease transmission, disease surveillance, and monitoring disease control program.\cite{28}

**GLOBAL POSITIONING SYSTEM (GPS)**

A GPS is composed of a system of at least 24 and up to 32 solar-powered satellites that orbit the earth every 12 hours and transmit radio waves at a precisely defined timed intervals.\cite{29} To assess a position in three dimensions i.e. latitude, longitude and elevation, a receiver needs specific signals from at least four satellites.\cite{30} GPS is becoming a standard technology for collecting data in geographical epidemiology and public health studies.\cite{31} Moreover, different components of a GPS receiver work efficiently under severe weather conditions such as high temperature, sandstorm, torrential rain etc. Therefore, they may have a key role in combination with GISs, especially in emergency humanitarian activities.\cite{32}

**REMOTE SENSING (RS)**

Remote Sensing is collecting information regarding an object, area or phenomenon without coming to direct contact with it. A Remote Sensor can be a seismograph, a fathometer etc. Seismographs are capable of measuring the intensity of earthquake being far away from the focus of the earthquake. Similarly, fathometers can measure the depth of the ocean without coming in contact with it. In modern Remote Sensing, information about earth’s land and water surfaces is acquired by using reflected or emitted electromagnetic energy.\cite{11} Most of the studies on public health using remote sensing data have utilized data from Landsat’s Multispectral Scanner (MSS) and Thematic Mapper (TM), the National Oceanic and Atmospheric Administration (NOAA)’s Advanced Very High Resolution Radiometer (AVHRR) and France’s Système Pour l’Observation de la Terre (SPOT).\cite{33}
**SPATIAL AUTOCORRELATION**
Spatial autocorrelation analysis is one method used to discover the extent to which given observations can be regarded as spatially independent or clustered. The correlation or dependency implies that rates for geographically close areas are more highly related than those from areas that are geographically distant.\(^{[34]}\) For example, suicide rates in neighbouring areas are likely to be more similar than those in distant ones.\(^{[35]}\)

This is because neighbouring areas may have similar underlying social, economic and cultural characteristics that trigger suicidal behaviour. Detecting spatial dependency, which is accomplished by the use of spatial autocorrelation statistics, would help researchers to justify their selected regression models in an ecological analysis, or when mapping a rare disease or when mapping in small boundaries.\(^{[1,36]}\)

**PUBLIC HEALTH ACHIEVEMENTS THROUGH DISEASE MAPPING**
Technologies utilizing spatial epidemiology are facilitating the mapping, visualizing, monitoring, retrieving, analyzing and modelling the geographic data related to public health with high accuracy. Disease mapping can also be used for monitoring the ongoing disease control programs.\(^{[28]}\) Till date, GIS, GPS and RS have helped to geographically describe numerous vector borne diseases in different parts of the world like: dracunculiasis (Nigeria), filariasis (Egypt, India), leishmaniasis (South-West Asia), lyme disease (USA), Malaria (Mexico), trypanosomiasis (Africa), Cholera (Bangladesh) etc.\(^{[33]}\)

**PRESENT CHALLENGES**
Though GIS is becoming increasingly popular among the medical geographers and epidemiologists, there is also a growing concern over its application for public health research. In spite of its being a powerful tool for many descriptive techniques, like spatial statistics, location-allocation, cartography etc. GIS does not offer any innovative idea about the determinants of population health and disease. So far, GIS and associated spatial analytical techniques are still presenting only a static snapshot of disease distribution based on cross sectional data, rather than representing the dynamic progression of disease through time and space. Moreover, GISs do not contain complex statistical functions. This has limited the application field of GIS as health researchers use software like SPSS, STATA, SAS, R etc. to perform the actual statistical analyses.\(^{[7,37]}\)
FUTURE PROSPECTS
Spatial epidemiology and its technologies may offer a more efficient and cost-effective solution, at least for exposures that can be readily characterized geographically. With this approach, a nested case–control or case–cohort study can be conducted within a large-scale population-based cohort by specifying a geographic subset of the cohort with high relative exposure, on average, for direct study.\textsuperscript{[6,38]}

Space-time representation is one of the current frontiers in the evolution of GIS, and is considered by many to be an essential component of the spatial analysis of disease patterns. One notable example of the innovative application of space-time representation to disease monitoring is the Dynamic Continuous-Area Space-Time (DYCAST) system. Another current advancement is coming from the increasing ability to collect and analyze mass amounts of information, a phenomenon known as Big Data. Since GIS incorporates spatial analysis with available data, utilizing large data sets could eventually provide us with up-to-date information on medical and social trends. Big Data will range “from few dozen terabytes to multiple petabytes (thousands of terabytes)”\textsuperscript{[7,39,40]}

Advances in GIS and statistical methodology together with the availability of high-resolution, geographically referenced health databases present unprecedented new opportunities to investigate the environmental, social and behavioural factors underlying geographic variations in disease rates at small-area scale. Such studies must be guided by good questions, excellent statistical methodology and sound epidemiologic principles, including taking proper account of problems of data quality and the potential for bias and confounding.\textsuperscript{[6]}

CONCLUSION
Spatial epidemiology has grown to a potential connection between social and biomedical sciences. This is a ground breaking approach in identification and mapping medically vulnerable populations, health outcomes, risk factors and the relationships between them. Healthcare organizations can now visualize, analyze, interpret and display multifaceted geo-location data through the use of GIS technology. Spatial epidemiologic studies is going to become a common tool in the future, both because of the instant visual appeal and wide availability of the new geographic techniques, and the desire for cleaner and healthier environments. The past decade has seen a sharp rise in the number of epidemiological studies employing GIS, particularly in the areas of health disparities, resources availability, and health-related behaviours, as well as continued use in more foundational fields such as cancer
and environmental epidemiology. With ongoing improvements in the data and methodologies, these studies will play an increasingly important role in our understanding of the complex relationships between environment and health.[6,7,41]

REFERENCES