Critical Review: Spatial Analysis in Epidemiology

By

Motunrayo Aderemi

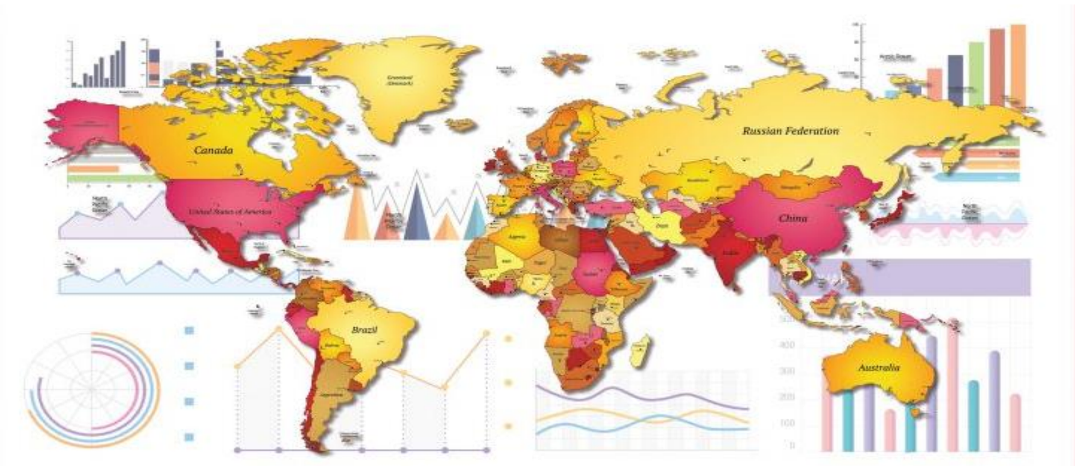


TABLE OF CONTENT

TABLE OF CONTENTS

Abstract

Awareness of the Influence of place on health on health is a critical component of epidemiological research and a range of techniques are in use to examine geographically distributed health-related data. Place has been consistently acknowledged as a significant distribution factor and possible control of diseases and other health-related factors. Spatial data analysis, remote sensing and GIS are important tools in the fight against disease which are currently underutilised. With the increased utilization of such tools in epidemiological studies, there is a possibility for better comprehension of disease mechanisms.

Epidemiology is a widely recognised in the field of health-science, which frequently employs spatial data. The purpose of this research is to discuss the impact of spatial data analysis and the techniques which includes GIS which can be used in epidemiology. Future directions for the use spatial analysis could also be proposed.

1.0 INTRODUCTION

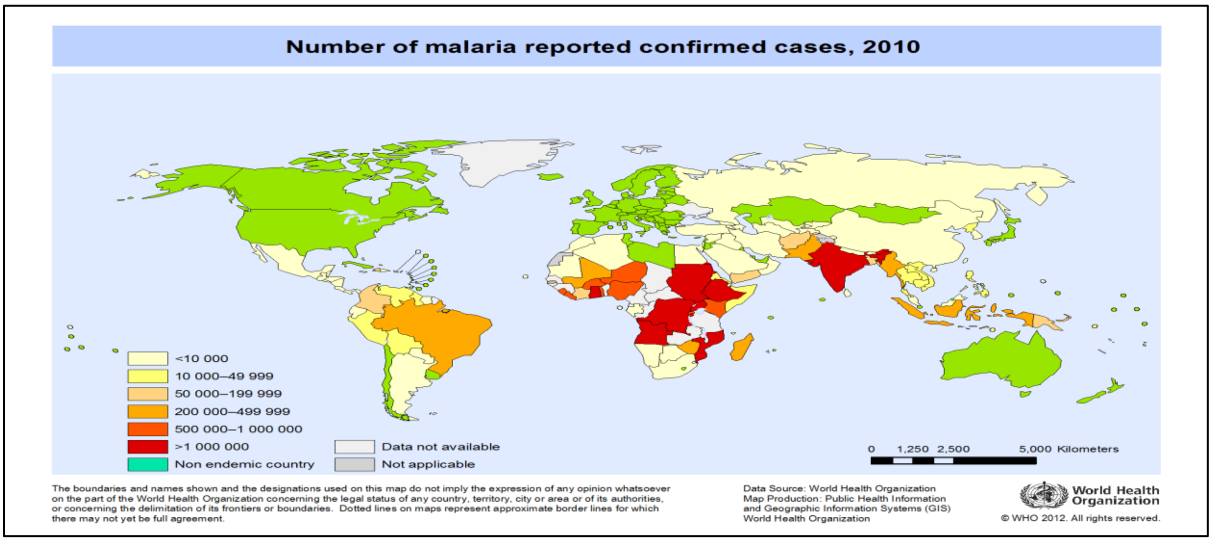
Epidemiology is the field that investigates the health and disease trends among populations. It is considered the cornerstone of public health, which aims to prevent diseases and promote the health of communities. (National Academy Press, 1998) Epidemiology is not only concerned with individual patients but also attempts to understand the health and disease patterns in populations to suggest ways to improve them. Basic medical sciences explain the causes of a particular disease and may suggest different treatments, such as pharmacological or dietary interventions, which can influence patient care.

Public health policies rely on epidemiological insight of the health and disease of the population, which are shaped by the methods and concepts used to understand a particular problem. Epidemiology has evolved significantly over time due to various factors such as advancements in hygiene practices, the bacteriological revolution, the increased use of statistical survey methods, the emergence of molecular biology, and the environmental movement.(Rothman, 1986)

The three fundamental components of investigation into outbreaks of diseases and epidemiology are person, place, and time. Even though disease mapping has been done for over a century. Epidemiological research has mainly concentrated on the individual and temporal aspects, with minimal attention given to the impact of spatial or geographic a relationship between processes of disease and spatial variables.

Maps play a critical role in identifying unusual spatial patterns of health outcomes and their links to environmental hazards, making spatial analysis techniques and methods increasingly valuable. Traditional mapping methods had several limitations, including the challenge of managing vast amounts of data. The advent of digital computers in the 1960s revolutionized the creation of quantitative thematic maps, paving the way for the development of GIS (Geographical Information Systems). (Mcdonnell, 1998).

This essay aims to explore the impact of spatial data analysis and the techniques which includes GIS and remote sensing which can be used in epidemiology. Different research, articles, studies, and platform, where spatial analysis and the techniques were used in epidemiology will be critically reviewed.



Above is a map showing the cases of Malaria all over the world in 2010

2.0 METHODOLOGY

The aim of this research is to investigate the impact of spatial data analysis and techniques such as GIS and disease mapping in epidemiology. The references in this essay were discovered through a systematical search procedure that included the search terms "Spatial analysis," "Spatial analysis in epidemiology," "epidemiology," and "Spatial data science in epidemiology." On Scorpus, PubMed and Google Scholar. Related articles written between 2019 and 2022 were downloaded without regard for the authors' country or region. The title, abstracts, and methods of the papers were used to determine their relevance. Relevant reports and blogs provided additional contributions.

The study is based on research carried out by various authors and incorporates applicable reports. In various parts of the essay, the impact of spatial data analysis and techniques such as GIS that can be used in epidemiology will be discussed.

3.0 LITERATURE REVIEW

3.1 Epidemiology

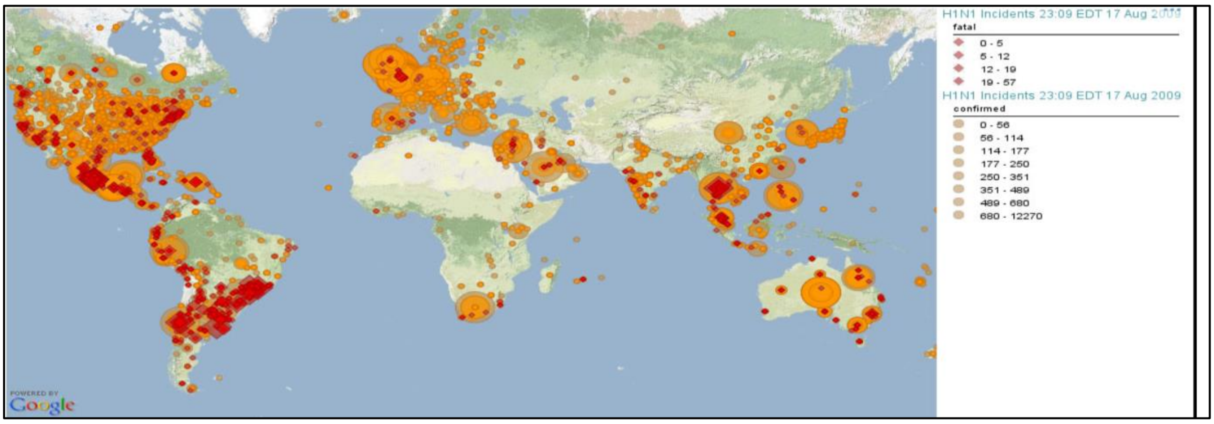
Epidemiology refers to the investigation of how disease spreads among individuals and the factors that contribute to or affect this spread. The key question that epidemiology seeks to answer is why some individuals acquire a disease while others do not. The basic premise of epidemiology is the fact that illnesses, diseases, and poor health aren't found at random in populations of people. Instead, each person has unique characteristics that can either increase their susceptibility to or protect them from various illnesses. These characteristics may stem from genetic factors or exposure to specific environmental risks. Typically, disease develop as a cause of a combination of genetic and environmental factors. (Thacker, 1988)

3.2 Spatial Data in Epidemiology

The dataset that contains health-related information linked to geographic locations is commonly known as a georeferenced dataset. The utilization of GIS technology enables users to obtain additional information from various sources. (Jerret *et al.*,2003) There are two primary types of spatial data.: Area data and Point data. Every piece of health data (which may include environmental exposure, population, morbidity, and mortality data) can be precisely connected to a specific geographic location, such as a residential address or an area defined by a province, ward, postcode, local authority, or country. GIS is recognized for its ability to execute four essential operations on spatial data: input, storage, analysis, and output. (Jerret *et al.*,2003)

3.3 GIS for Spatial analysis in Epidemiology

GIS technology is a comprehensive computer hardware and software toolset that integrates various functionalities, including capturing, editing, storing, organizing, analysing, and displaying spatially referenced data. Its application in generating maps and data can help identify risk factors and enable spatial analysis techniques useful in healthcare research and epidemiology (Westlake, 1995). The use of GIS has significantly enhanced the ability of public health professionals and epidemiologists to work with spatial information. The technology offers numerous benefits, including the capability of dealing with recurring tasks and comparing data from multiple sources and regions, manipulate data rapidly, manage large datasets, enforce standardization, and ask hypothetical questions. Remote sensing data such as digital satellite imagery can also be used for environmental analysis. For optimal effectiveness in healthcare and epidemiology research, GISs should be able to execute or link with spatial analytical and modelling programs. They can be used to plan the provision of primary health care and access based on geography, neighbourhood, distance, or the makeup of the population (ArcView, 1998). GIS technology has been beneficial in studying infectious and chronic diseases, as well as healthcare-related issues. While some spatial data statistical analyses can be performed within GIS, more advanced approach might necessitate the analysis of data in a statistical software package before exporting to the GIS for the display of map. (Twigg, 1990). GIS can also be applied to study vector-borne diseases like Lyme disease. In infectious disease epidemiology, GIS is frequently combined with statistical modelling methods such as logistic regression. (Lawson,1999)

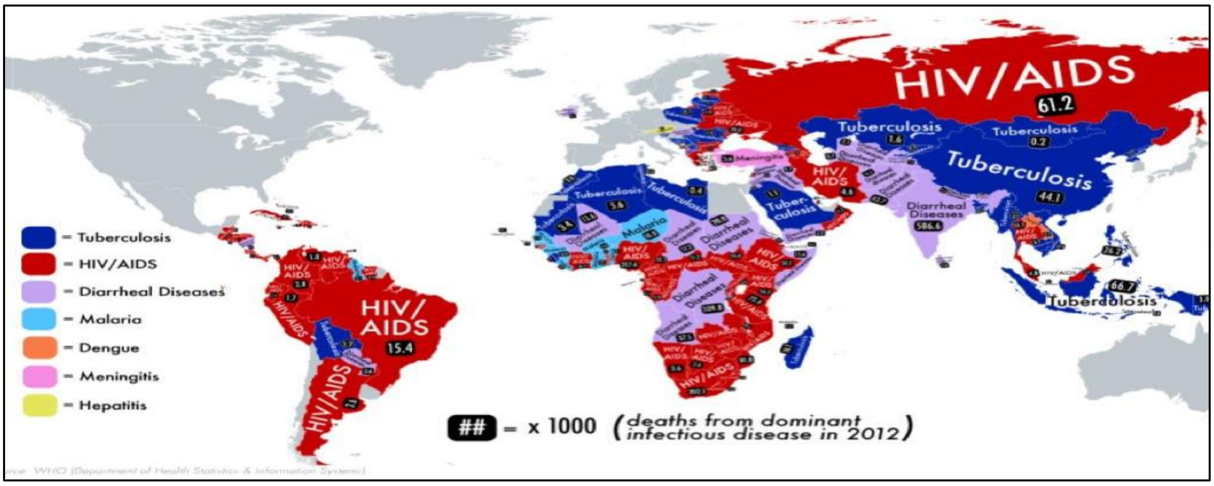


Above is a GIS Epidemiology mapping for Swine Flu

3.4 Other Spatial Analytical methods for Epidemiology

3.4.1 Disease Mapping

Visualisation of data is the initial stage in uncovering the intricate structure of data. Apart from capturing the viewer's attention, it can also lead to unexpected discoveries. While graphs and other visual aids are essential for analysts, spatial analysts often employ maps to visualise spatial data. (Everitt, 2001) Disease mapping is a subdivision of geographical epidemiology that seeks to create precise maps of disease mortality and morbidity. The disease mapping process is a visual representation of where the disease is occurring. This can be done for a variety of reasons, such as determining the location of disease outbreaks or tracking the spread of a disease. Disease mapping entails mapping case locations, incidence rates by area, and standardised rates. Pickle et al. (1997) There are several methods for creating a map of disease occurrence. One of these methods is to use data from health care facilities such as hospitals or clinics. This data can be used to generate a map that shows where cases of a disease have been reported. Data from surveys can also be used to map disease occurrence. This information can be used to create a map that shows where people with a specific disease live. This can be useful in understanding the spread of a disease or identifying associated risks for a disease.(Smallman, 1992).



Above is a map which shows the deadliest infectious diseases in the world

3.4.2 Spatial Clustering Analysis

In epidemiology, spatial cluster methods are the most used tool for evaluating non-random spatial patterns. Several statistical approaches are being implemented to identify the significance of disease clusters and not just a result of chance (Kulldorff, 2006), with global clustering tests evaluating clusters without specifying their precise locations, local clustering tests identifying small-scale clusters, and focused clustering evaluating clustering around a fixed-point source. These tests each have their own significant explanations and techniques for detection are distinct entities, and the existence of one is not always indicative of the presence of the other. (Huang, 2008)

Tools for detecting clusters are available for both point-pattern and areal data. For point-pattern data, clusters are identified by statistical tools such as the variance to mean ratio, nearest neighbour analysis, and quadrant analysis. These techniques evaluate whether events are systematically organized or randomly distributed. Areal data groups data into regions to identify clusters on a larger scale.(Lawson,1999)

Various tools are being developed to capture data for epidemiology, including HealthMapper, developed by the WHO for surveillance and disease mapping(WHO,1999); SIGEpi, created by the Pan-American Health Organization to enhance epidemiological analytical resources;(Martinez et al.2001) and GeoDa, developed at the Spatial Analysis Lab at the University of Illinois.

4.0 Related Works

Pleydell et al. (2004) conducted a study on Echinococcus multilocularis, a tapeworm, and its distribution in Europe. The authors highlight the significance of mapping infection risk in wildlife to monitor changes in human transmission risk over time and space. The study employed geostatistical interpolation of laboratory-analysed faecal matter to accomplish this task. Additionally, Echinococcus multilocularis transmission in high-risk populations in China was also investigated by Pleydell et al. (2004). Landsat remotely sensed data was used to create a map of vegetation cover, and landscape metrics were obtained and linked to the prevalence of infections for sample villages at various time and space intervals.. The findings revealed that spatially linked habitat form and function play a crucial role in determining disease transmission under specific environmental and social conditions. Graham et al. (2004) further expands on this theme of epidemiological data analysis with regards to landscape and disease transmission ecology.

Mohzen et al. (2007) in their study, highlighted the factors that must be taken into account when designing a map, such as choosing the suitable administrative unit, data classification method, and color scheme or hatching pattern. However, Westlake A (1995) cautioned that while disease mapping can be an effective tool for communication, it can also be deceptive, especially when rates are mapped in small areas.

In their study, Openshaw et al. (1987) utilized a quadrat-based approach in a geographic analysis machine (GAM) to identify and visually present the clustering of childhood leukemia cases. On the other hand, Harries (1997) employed a cluster analysis method along with map overlay to categorize neighborhoods in Baltimore, Maryland according to their levels of violence and social stressors, instead of solely detecting clusters.

Lin and Zhang (2004) centered their research on the examination of spatial clustering and devised a testing methodology to identify the existence of clustering with high or low values. Such techniques could be applied to scrutinize spatial aberrations in a "typical" disease surface by leveraging epidemiological data. In contrast, Baker (2004) incorporated a temporal element to the spatial clustering methodology and stated that "a cluster of cases that are near in both space and time indicates an infectious cause." Baker's study employed a test that modeled a point with a magnitude that increased with proximity to an infector.

Noor et al. (2004) are particularly interested in data acquisition and analysis using geographic information systems (GIS). They provide information on geographic disparities in Kenyan health-care delivery, as well as imbalances caused by social, economic ,service, and environmental factors. "In much of the developing world, our understanding of the spatial determinants of risk, access to services, and health equity is limited" the authors write in their paper.

Hyndman et al. (1997) investigated the location of mammography and abdominal aortic aneurysm screening clinics in Perth, Australia, using a geographic information system (GIS). Her team was able to investigate the effect of distance from residence on screening clinic use using small area census demographic data and geocodes for household water metres.

5.0 DISCUSSION

When considering the use of GIS and spatial analytical techniques in the field of epidemiology and research in health, there are several factors to consider. The authors of this study investigate the significance of incorporating spatial data techniques and methods in epidemiology based on their review of existing research. The articles included in this work emphasise the increasing significance of spatial data collection and analysis in epidemiology. It is recommended that future epidemiological studies make explicit use of spatial information, and surveys should routinely include spatial location measurement to facilitate a complete analysis of the space-time phenomenon of disease outbreaks. This study showcases the potential of GIS, disease mapping, and spatial cluster analysis in epidemiological analysis by showcasing a selection of available techniques for use by epidemiologists and health workers. However, user groups must be supportive of their adoption for this potential to be fully realised.

Jacquez (2000) identified several obstacles to the adoption of GIS techniques in health sciences, including a lack of spatial thinking training among epidemiologists, inadequate studies demonstrating GIS's potential, as well as a scarcity of specialised GIS software for applications in healthcare. To overcome this limitation, it is recommended that epidemiologists receive appropriate training on spatial techniques, and that dedicated GIS packages be developed for health purposes.

Several factors will shape the future of spatial analysis in epidemiology. These include advancements in computer speed and storage capacity, the creation of spatial disease transmission models based on processes, as well as the distribution of useful software to healthcare professionals. The research community aims to obtain larger and more accurately positioned datasets to improve the effectiveness and accuracy of spatial analysis.

Developing precise process-based models of disease transmission can be challenging due to most disease systems' difficulty. However, epidemiologists are striving to create spatially distributed process models that link the function of disease transmission, time and space (Gewin, 2004). These models are crucial because they will allow for predictions for future studies to be made on the basis of real-time scenarios using analytical methods for in order to provide early warning effective disease prevention. Additionally, these models will facilitate "what-if" scenario planning for uncertain climate change situations. It is essential to have a comprehensive understanding of disease transmission to create models that can accurately predict future disease outbreaks.

When considering a spatial analytical technique for epidemiological research, it is important to consider various factors such as the type of data being analyzed (continuous, categorical or dichotomous), the process being explored (cluster detection, interpolation, or diffusion), accessibility of spatial data, and the scale of analysis. Additionally, commercial software is available for the chosen technique should also be considered.

However, one of the major limitations in conducting small-area studies using routine data sources is data availability and quality. For instance, in the US and UK, data on residential history is not easily accessible, which is crucial for reconstructing exposure histories for long-term health outcomes like cancer incidence and certain types of mortality. This highlights the need for improved data quality and availability in order to conduct more comprehensive and accurate epidemiological analyses.

6.0 CONCLUSION

This study focuses primarily to the use of GIS and discusses several spatial analytical methods used in Epidemiology which include Disease mapping and spatial clustering analysis. The study highlights the significance of acquiring and analysing spatial data in epidemiology and recommends that future epidemiological research must explicitly incorporate spatial information, including spatial location measurements in surveys, to enable a comprehensive analysis of disease outbreaks as a space-time phenomenon.

However, the techniques discussed in these publications can be easily automated and adapted to be significantly useful for subsequent disease outbreaks.

REFERENCES

ArcView GIS, 3.1. Redlands, CA: Environmental Systems Research Institute, 1998.

Beck, L.R., Lobitz, B.M., and Wood, B.L. (2000) ‘Remote sensing and human health: new sensors and new opportunities’. Emerg. Infect. Dis. 6, 217–226.

Everitt, B.S.E, Dunn, G. (2001) ‘Applied multivariate data analysis’, London: Arnold.

Gewin, V. (2004) Disease control: virtual plagues get real. Nature 427, 774–775.

Gewin, V.(2004) ‘Disease control: virtual plagues get real’ . Nature 427, 774–775

Graham, A.J, Danson, F.M, Giraudoux, P., and Craig, P.S. (2004) ‘Ecological epidemiology: landscape metrics and human alveolar echinococossis’. Acta Trop., this issue.

Harries, K.(1997) ‘Social stress and trauma: synthesis and spatial analysis’. Soc Sci Med ;45:1251-64.

Huang, L., Pickle, L.W, and Das, B., (2008) ‘Evaluating spatial methods for investigating global clustering and cluster detection of cancer cases’. Stat. Med. 27:5111–42

Huang, L., Pickle, L.W., Das, B.(2008) Evaluating spatial methods for investigating global clustering and cluster detection of cancer cases. Stat. Med. ; 27:5111–5142. [PubMed: 18712778]

Hyndman, J., Holman, C.D., Jamrozik, K.(1997) ‘The effect of spatial definition on the allocation of clients to screening clinics’. Soc Sci Med ;45:331^0.

Institute of Medicine Committee for the Study of the Future of Public Health. (1988). ‘The Future of Public Health’. Washington, DC: National Academy Press.

Jacquez, G.M. (2000) ‘Spatial analysis in epidemiology: nascent science or a failure of GIS?’ J. Geogr. Syst. 2, 91–97.

Jerrett, M., Burnett,R.T, Goldberg, M.S, et al.(2003) ‘Spatial analysis for environmental health research: concepts, methods, and examples’. J Toxicol Environ Health A66:1783–810.

Kulldorff, M. (2006) ‘Tests of spatial randomness adjusted for an inhomogeneity’: a general framework. J. Am. Stat. Assoc.; 101:1289–1305

Lawson, A.B, Bohning, D, Biggeri, A, et al, (1999). ‘Disease mapping and its uses, disease mapping and risk assessment for public health’. Chichester: Wiley.

Martinez, R., Vidaurre, M., Najera, G.P., Loyola, E., Castillo-Salgado, C., and Eisner, C. (2001) SIGEpi: Sistema de Información Geográficaen Epidemiología y Salud Pública. Epidemiol Bull 22:4–5

Mattews, S.A. (1990) ‘Epidemiology using a GIS: the need for caution’ , Comput Environ Urban Syst

Mcdonnell, R.A. (1998) *Principle of Geographic Information Systems, MENAdrought View project MENA-Regional drought management system View project*. Available at: https://www.researchgate.net/publication/37419765.

Mohzen, R., Graham, D., Seleyn, S., and Louis. A.(2007). 'Geographical epidemiology, spatial analysis and geographical information systems'. a multidisciplinary glossary. 10.1136/jech.2005.043117

Openshaw, S., Charlton, M., Wymer, C., et al. (1987) 'A Mark 1 Geographical Analysis Machine for the automated analysis of point data sets'. Int J Geogr Inf Syst;l:335-58.

Pickle, L.W, Mungiole,M and Jones,G.K. (1997) 'Atlas of United States mortality',Hyattsville, MD: US Department of Health and Human Services, Centers for Disease Control and Prevention, National Center for Health Statistics.

Pleydell, D.R.J., Raoul, F., Tourneux, F., Danson, F.M., Graham, A.J., Craig, P.S., Giraudoux, P. (2004) 'Modelling the spatial distribution of Echinococcus multilocularis infection in foxes' Acta Trop., this issue.

Rothman, K.J. (1986) 'Modem Epidemiology'. Boston: Little, Brown and Co.

Smallman-Raynor, M.R, Cliff, A.D, Haggett, P. (1992) 'Atlas of AIDS'. Oxford, England: Blackwell Publishers.

Thacker S, Berkelman R.L.(1988) *Public health surveillance in the United States*, Epidemiol Rev 10:164-190.

Twigg, L. (1990) 'Health based geographical information systems: their potential examined in the light of existing data sources'. Soc Sci Med 1990,30:143-55.

Westlake, A.(1995) 'Strategies for the use of geography in epidemiologic analysis'. In: de Lepper MJC, Scholten HJ, Stern RM , eds. The added value of geographical information systems in public and environmental health. Boston, MA: Kluwer Academic Press,:135-44