[DA-035] GIS&T and Public Health

Abstract

Contemporary environmental problems, global climate change, globalization, and urbanization have imposed severe impacts on human health. Meanwhile, disparity became a major concern in healthcare policy making and resource allocation. Within this context, GIS have been rapidly expanding and deepening their applications in the domain of public health. GIS applications in public health can be classified into three broad categories: 1) spatial/spatiotemporal modeling of specific diseases, including chronic diseases and communicable diseases, as well as their associations with environmental risks; 2) spatial/spatiotemporal modeling of environmental exposures from physical, behavioral, and/or socioeconomic environments; and 3) studies on healthcare services, including assessment of geographic access to healthcare facilities, investigation of disparity in the access, and optimization of resource allocation. The boundaries between these divisions are not clear-cut. Meanwhile, applications in public health have also been pushing the frontiers of GIS research on spatiotemporal modeling, high-performance computing, uncertainty, big data of human mobility, and geospatial privacy.

Keywords: chronic diseases, communicable diseases, environmental exposure, environmental health, epidemic modeling, healthcare, public health, spatial epidemiology

Author & citation

Shi, X. (2019). GIS&T and Public Health. The Geographic Information Science & Technology Body of Knowledge (4th Quarter 2019 Edition), John P. Wilson (ed.).
DOI: [10.22224/gistbok/2019.4.13](https://doi.org/10.22224/gistbok/2019.4.13).

This entry was first published on December 11, 2019. No earlier editions exist.

Explanation

1. Definitions

GIS: Within the context of this entry, GIS refers to broadly defined geographic information sciences, methodologies, technologies, and software systems.

Spatial epidemiology: Within the context of this entry, this term refers to population-level and geographic-level studies of human health that take a geographic perspective and employ GIS.

Cluster: Within the context of this entry, a cluster refers to a relatively small geographic area that has abnormally (statistically significant) high intensities (measured by prevalence, incidence, or other measurements) of the disease under concern.

Spatial association: Within the context of this entry, spatial association refers to spatial match of two spatial distributions (e.g., a disease and an environmental factor).

"Big" data: Within the context of this entry, this terms particularly refers to individual-level data of a mass population.

Geospatial privacy: Privacy that is related to or can be revealed by geographic location information.

2. Overview

Contemporary environmental problems, global climate change, globalization, and urbanization have imposed severe impacts on human health. Meanwhile, disparity became a major concern in healthcare policy making and resource allocation. All these have brought unprecedented attention to the connection between human health and factors at the regional and even the global levels, and have created an enormous demand for health-related applications of GIS. On the one hand, GIS have been continuously and rapidly expanding their influence and acceptance in this area, and on the other hand, to meet and also be motivated by the expectation, as well as the challenge, from this particular area, GIS have also experienced tremendous developments (Richardson et al. 2013, Kwan 2016). Figure 1 illustrates this dynamic with the yearly number of scientific publications.

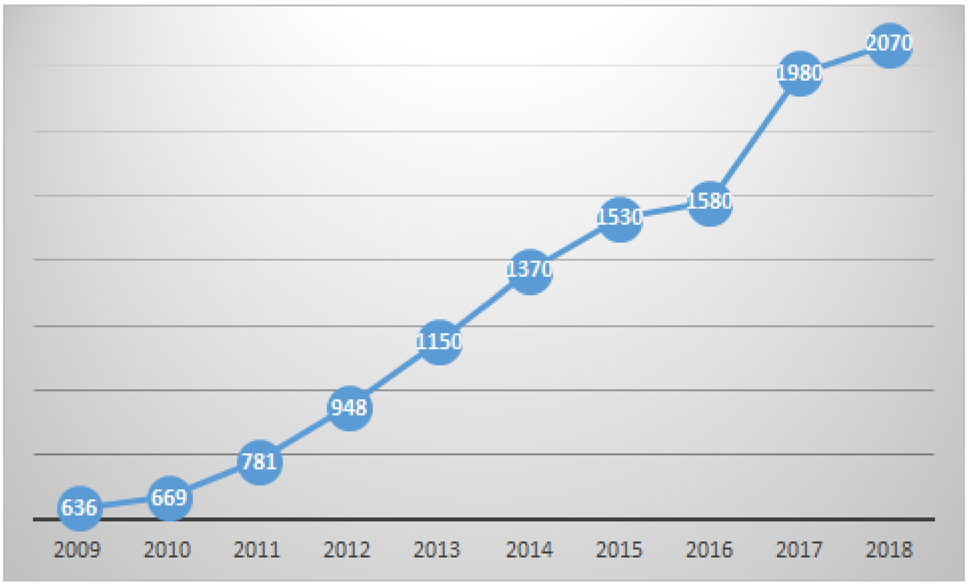

Figure 1. Yearly numbers of publications returned by a search with Google Scholar using "cancers and geospatial" as key words, for the 10-year period 2009-2018. Source: author.

So far, GIS applications in public health can be classified into three broad categories: 1) modeling of diseases, 2) modeling of environmental exposures, and 3) studies on healthcare services. Each can be further divided according to the specific issues addressed. Noteworthy, the boundaries between these divisions are not clear-cut.

3. Modeling of Diseases

Spatial/spatiotemporal modeling of diseases can be further classified as studies of chronic diseases and studies of communicable diseases. The purposes and methodologies of these two kinds of studies are different.

3.1 Chronic Diseases

This can typically be a spatial environmental health study (refer to the entry GIS&T and Epidemiology), usually addressing either or both of two major issues: 1) characterizing spatial distribution of a disease and 2) detecting spatial association between a disease and certain environmental factors. Its goal is to suggest hypotheses about etiological disease-environment associations, so as to motivate and/or facilitate further epidemiological and biomedical investigations. Noteworthy, such a study by itself cannot conclude about causal disease-environment relationships (Openshaw 1996).

Chronic diseases that have been mostly tackled by this kind of study include many types of cancers (e.g., leukemia, lung cancer, breast cancer, colorectal cancer, liver cancer, and skin cancer), birth problems (e.g., birth defects, preterm birth, and low birth weight), neurological diseases (e.g., Parkinson's disease, Alzheimers' disease, and amyotrophic lateral sclerosis), obesity, and mental disorders. There are also behavioral studies adopting similar approaches, covering issues like smoking, drug abuse, and suicide.

Charactering spatial distribution of a disease, i.e., disease mapping, has been often conducted as process of cluster detection, which might be the most conspicuous application of GIS in epidemiology. An early milestone of this application is Openshaw's Geographical Analysis Machine, an essentially kernel density estimation (KDE) process for detecting areas with abnormally high intensity of point events (Openshaw et al. 1988). In 1990's, the International Agency for Research of Cancer (IARC) launched a "contest" to test different cluster detection methods (Alexander and Boyle, 1996). Currently, the spatial, temporal, or space-time scan statistics method developed by Kulldorff and his colleagues (Kulldorff and Nagarwalla 1995, Kulldorff 1997) is arguably the most widely used cluster detection method (Figure 2). In ArcGIS, the built-in tools for "mapping clusters" include Anselin Local Moran I and Getis-Ord Gi*. ArcHealth, developed as an Extension of ArcGIS, contains a sophisticated KDE-based cluster detection tool (Shi 2009).

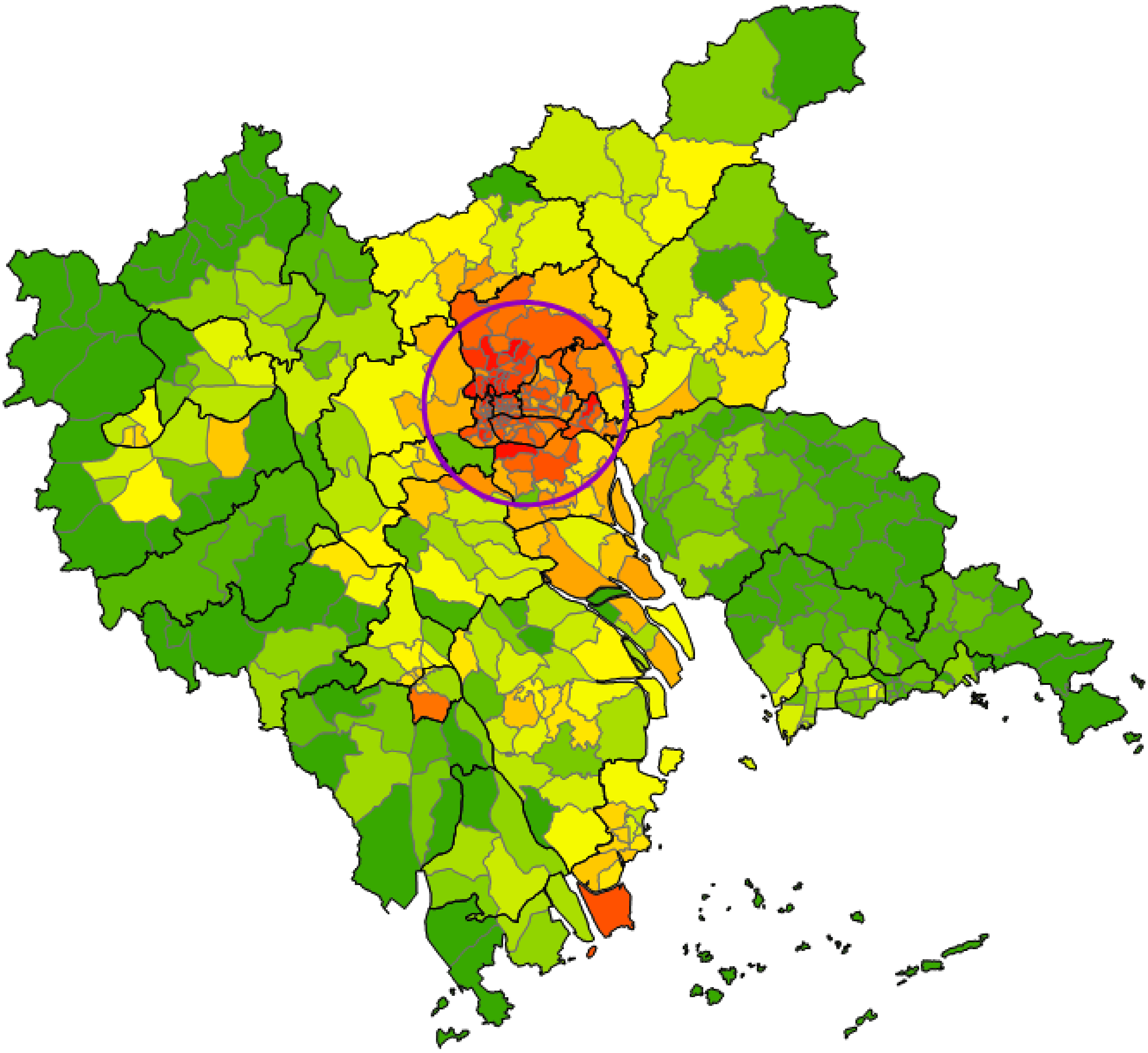

Figure 2. A cluster (delineated by the circle) of a disease detected by SaTScan. Due to the confidentiality concern, the specific incidence value represented by the color of the areal unit is not presented here. Source: author.

For detecting disease-environment associations, a representative method is the geographically weighted regression (GWR) (Brunsdon et al. 1996), which conducts localized analysis that is able to map the spatial variation in the association. Another spatial analytical method for quantifying the disease-environment association, called the sandwich method, has also become popular (Wang et al. 2013).

3.2 Communicable Diseases

Contemporary environmental changes and human dynamics have brought about emergence, re-emergence, and intensification of certain communicable diseases, exemplified by HIV/AIDS, flu, dengue fever, malaria, West Nile Virus, and Ebola (just naming a few from a long list). A local epidemic or a small number of imported cases can result in regional or even global outbreaks. For communicable diseases, a GIS-supported study aims to reveal, understand, and model spatiotemporal pattern of the epidemic. In the most classic story about “disease and map,” Dr. John Snow was dealing with a communicable disease (the 1854 London cholera outbreak), but the contemporary epidemics of communicable diseases, featuring sudden, rapid, extensive, and jumping spread processes, present much more challenging tasks to today’s spatial epidemiologists.

Conventional epidemic models for communicable diseases focus on the temporal dimension. The generic SIR (susceptible-infectious-recovered) model is to estimate the total number of people in each compartment (S, I, or R) of the entire study area at a certain time point, without considering the spatial variation in the transmission process. As the importance of spatial variation became increasingly recognized, spatialization of epidemic models have been increasingly attempted and appreciated (Emch et al. 2012, Riley et al. 2015). Conjugated with the spatialization is the individualization, i.e., the target of the model shifts from the population in an area as a whole to individuals (Bian 2013). These two trends need GIS.

The individual-based mobile model (Bian et al. 2012) is a representative of these trends. This model creates simulated individuals in a study area, each with certain attributes and behaviors, assigned according to the characteristics of the population in the area. All individuals’ daily travel trajectories are simulated based on their assumed daily life routines, mainly about how much time they spend at home, workplace, shopping places, and others. An individual’s infection status may change during her/his daily activities through contacts with others. The model can prospectively estimate numbers of people in the S, I, and R compartments. Alternatively, Li et al. (2019) explored a retrospective approach to modeling a historical epidemic based on actual individual level data. The approach aims to construct epidemic trees (forest) that represent transmissions between individual patients, and in turn the overall spatiotemporal pattern of the epidemic (Figure 3). The linkage between two patients can be modeled with the increasingly available “big” data of human mobility.

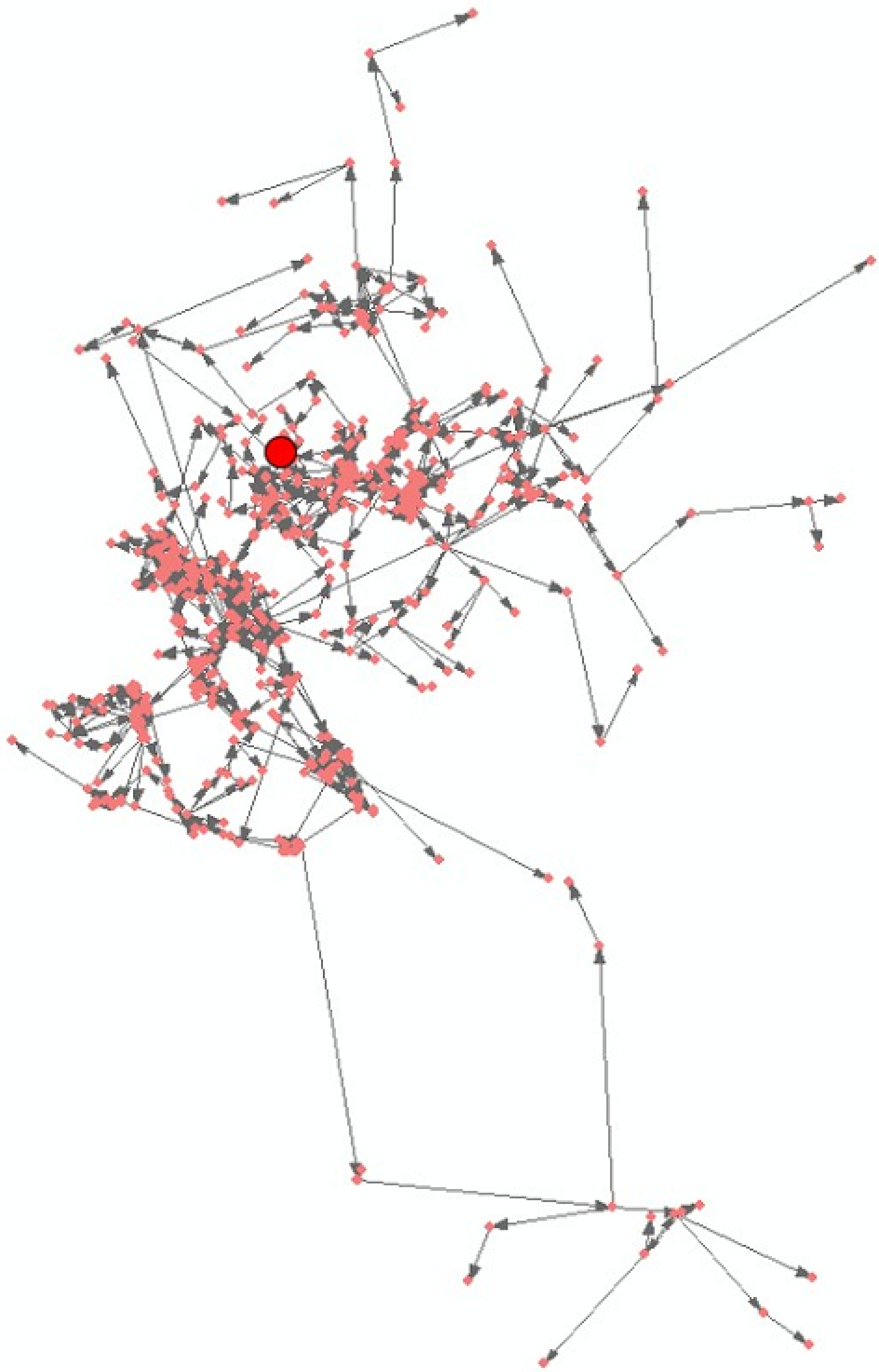

Figure 3. An epidemic of a communicable disease starting from a primary case, modeled by the Epidemic Forest method. Source: author.

4. Modeling of Environmental Exposures

GIS-based environmental modeling became an important approach to estimating environmental exposures. Its underlying assumption is that the concentration of an environmental influence at a human subject's location (e.g., residential location) can represent the subject's actual exposure to that influence. Typically, the environmental data are available only at particular locations (e.g., air quality monitoring stations), and thus spatial inference is needed to estimate concentrations at the human subjects' locations. Commonly employed inferences include interpolation and kernel density estimation (KDE). Interpolation has been widely used in environmental health studies to map contamination of soil (e.g., Carlon et al. 2001, Hooker and Nathanail 2006, Aelion et al. 2013) and sediments (e.g., Katz et al. 2013), model chemical concentrations in groundwater (e.g., Goovaerts et al. 2005, Ayotte et al. 2006), and quantify the degree of local air pollution (e.g., Yuan et al. 2015, Borge et al. 2016, Martens et al. 2017). Studies that employ KDE are often under topics of material flow, environmental fate, and pollution emission (e.g., Jerrett et al. 2005, Tonne et al. 2006, Hu et al. 2008, Gottschalk et al. 2010), and access to food outlets and recreational facilities (Kestens et al. 2010, Thornton et al. 2011).

Remote sensing and digital elevation models (DEM) have also been widely used for modeling the habitat environment of hosts and/or media in vector-borne diseases, measuring terrain attributes that may be related to certain health-related behaviors, and evaluating the walkability as a healthy factor in a neighborhood.

5. GIS-based Studies of Healthcare Services

Access to healthcare services is a major topic in the debate about healthcare system reform, which has been fiercely ongoing in many countries. Geographic access (or spatial access) is a critical aspect of the more general access. Essentially, geographic access is about the geographic distance, and can adopt various measurements, typically travel time. The studies of geographic access to healthcare services is rooted in the research of economic geography and urban planning, for which GIS became a necessary platform and tool.

The fundamental issue in this field is the assessment of geographic access. This kind of study can be about the patients of a certain disease (e.g., breast cancer) and/or facilities of a certain type (e.g., those that provide mammography imaging), and can be about a certain scale (e.g., nationwide, statewide, or a metropolitan area) and/or at a certain resolution (e.g., based on data aggregated to counties, zip codes, or raster cells). The access can simply be represented by the distance (or travel time) to the nearest facility, or can take into account of the convenience of being close to multiple facilities. The most sophisticated method so far is the two-step-floating-catchment-area (2SFCA) method (Luo and Wang 2003), which is essentially a kernel-density estimation over inhomogeneous background (Shi et al. 2012). It takes in factors including all facilities around the location (i.e., not only the nearest one), the capacity of each facility, the local population (or number of patients), and the distance decay effect.

Many studies then use the outputs from the assessment to detect the impact of access on certain health outcomes. Others associate the access with demographic or socioeconomic variables to investigate disparities in healthcare resource allocation.

Some researchers have taken a further step to optimize future healthcare resources allocation based on the assessment (Wang et al. 2015). There can be different objectives in this kind of planning. Essentially, it tries to balance the economic efficiency and social justice.

6. Studies of Methodology and Infrastructure

While GIS have kept expanding and deepening their applications in public health, the applications have also been pushing the frontiers of GIS research. This section briefly describes a few topics for which research has been particularly active.

- **Spatiotemporal Analysis**. Incorporating the temporal dimension is a major topic in current GIS research, and public health is among the domains that have the greatest demand for such developments.
- **High-performance computing**. The applications of GIS in public health may require high-end computing resources, not only because they constantly handle large datasets, but also because they are adopting the computational approach to problem solving. For example, the Monte Carlo simulation have been widely used in these applications. The applications of GIS have been an important driving force to the development of GIS with the high-performance computing capacity or the cyberGIS (Wang et al. 2013).
- **Spatial Uncertainty.** The Spatial Uncertainty Program launched by the US National Institutes of Health (NIH) in 2011 have brought about a big wave of research on this topic in both epidemiology and geography, and greatly facilitated the communication between the two communities. During this wave, the concept of the uncertain geographic context problem or UGCoP (Kwan 2012) has facilitated bridging quantitative and qualitative studies in this area.
- **Big Data of Human Mobility**. Spatiotemporal dynamics of people have always been a critical factor in health issues. Particularly, the spatialization and individualization of epidemic models, as discussed earlier in this entry, have found great support from the increasingly available “big” data of human mobility, such as cellphone data, social media data, and transportation data. How to take advantage of these data and how to deal with related issues, e.g., the privacy issue, are highly active research topics.
- **Geospatial Privacy**. Patients’ privacy and data confidentiality are a prominent issue in health-related studies, particular for GIS-based studies, as the accurate and precise location information is the whole basis of such studies, but is considered identifiable and is to be concealed. This dilemma leads to critical research about the use of GIS in public health on the one hand, and technical research about getting around the problem on the other (e.g., geomasking).

7. Summary

Epidemiologists, medical geographers, and public health practitioners have known for long that human health is not only about the individual body, but should also be understood at the population level within particular geographic contexts. However, only GIS provide the possibility of eventually implementing this understanding. The applications of GIS in public health have seen a tremendous growth in the past 20 years, and is still in its rapid inclining phase. Meanwhile, the applications in public health have also brought about challenges to GIS, and have been an important driving force to the development of GIS.

References

Aelion, C.M. et al. (2013). Associations between soil lead concentrations and populations by race/ethnicity and income-to-poverty ratio in urban and rural areas. Environmental Geochemistry and Health, 35(1), 1-12.

Alexander, F.E. and Boyle, P. (1996). Methods for Investigating Localized Clustering of Disease. The simulated data-sets. IARC Scientific Publications (135), 21-27.

Ayotte, J.D. et al., (2006). Bladder cancer mortality and private well use in New England: an ecological study. Journal of Epidemiology & Community Health, 60(2), 168-172.

Bian, L. (2013). Spatial approaches to modeling dispersion of communicable diseases–a review. Transactions in GIS, 17(1), 1-17.

Bian, L., Huang, Y., Liang, M., Lim, E., Lee, G., Yang, Y., Cohen, M., and Wilson, D., (2012). Modeling individual vulnerability to communicable diseases: A framework and design. Annals of the Association of American Geographers, 102(5), 1016-1025.

Borge, R. et al., (2016). Assessment of microscale spatio-temporal variation of air pollution at an urban hotspot in Madrid (Spain) through an extensive field campaign. Atmospheric Environment, 14:432-445.

Brunsdon, C., Fotheringham, A.S., & Charlton, M.E. (1996). Geographically Weighted Regression: A Method for Exploring Spatial Nonstationarity. Geographical Analysis, 28(4), 281-298.

Carlon, C., Critto, A., Marcomini, A., and Nathanail, P. (2001). Risk based characterization of contaminated industrial site using multivariate and geostatistical tools. Environmental Pollution, 111(3), 417-427.

Emch, M., Root, E. D., Giebultowicz, S., Ali, M., Perez-Heydrich, C., & Yunus, M. (2012). Integration of Spatial and Social Network Analysis in Disease Transmission Studies. Annals of the Association of American Geographers, 102(5), 1004-1015.

Goovaerts, P., AvRuskin, G., Meliker, J., Slotnick, M., Jacquez, G., and Nriagu, J. (2005). Geostatistical modeling of the spatial variability of arsenic in groundwater of southeast Michigan. Water Resources Research, 41(7).

Gottschalk, F., Scholz, R.W. and Nowack, B. (2010). Probabilistic material flow modeling for assessing the environmental exposure to compounds: Methodology and an application to engineered nano-TiO2 particles. Environmental Modelling & Software, 25(3), 320-332.

Hooker, P. J. and Nathanail, C.P., (2006). Risk-based characterization of lead in urban soils. Chemical Geology, 226(3-4), 340-351.

Hu, Z., Liebens, J. and Rao, K.R., (2008). Linking stroke mortality with air pollution, income, and greenness in northwest Florida: an ecological geographical study. International

Journal of Health Geographics, 7(1), 20.

Jerrett, M. et al. (2005). A review and evaluation of intraurban air pollution exposure models. Journal of Exposure Science and Environmental Epidemiology, 15(2), 185.

Katz, D. R., et al. (2013). Factors regulating the accumulation and spatial distribution of the emerging contaminant triclosan in the sediments of an urbanized estuary: Greenwich Bay, Rhode Island, USA. Science of the Total Environment, 443123-133.

Kestens, Y., Lebel, A., Daniel, M., Theriault, M., and Pampalon, R. (2010). Using experienced activity spaces to measure foodscape exposure. Health & Place, 16(6), 1094-1103.

Kulldorff, M. (1997). A spatial scan statistic. Communications in Statistics - Theory and Methods, 26(6), 1481-1496.

Kulldorff, M. and Nagarwalla, N. (1995). Spatial disease clusters: Detection and inference. Statistics in Medicine, 14(8), 799-810.

Kwan, M-P. (ed.) (2016). Geographies of Health, Disease and Well-being: Recent Advances in Theory and Method. Routledge.

Kwan, M. P. (2012). The Uncertain Geographic Context Problem. Annals of the Association of American Geographers, 102(5), 958-969.

Li, M., Shi, X., Li, X., Ma, W., He, J., and Liu, T. (2019). Epidemic Forest: A Spatiotemporal Model for Communicable Diseases. Annals of the American Association of Geographers, 109(3):1-25.

Luo, W. and Wang, F. (2003). Measures of Spatial Accessibility to Health Care in a GIS Environment: Synthesis and a Case Study in the Chicago Region. Environment and Planning B: Planning and Design, 30(6), 865-884.

Martens, D.S. et al., (2017). Prenatal air pollution and newborns' predisposition to accelerated biological aging. JAMA Pediatrics, 171(12), 1160-1167.

Openshaw, S. (1996). Using a geographical analysis machine to detect the presence of spatial clustering and the location of clusters in synthetic data. IARC Scientific Publications, 1(135), 68-87.

Openshaw, S., Charlton, M., Craft, A.W., and Birch, J.M., (1988). Investigation of leukemia clusters by use of a geographical analysis machine. The Lancet, 331(8580), 272-273.

Richardson, D.B., Volkow, N.D., Kwan, M.-P., Kaplan, R.M., Goodchild, M.F., and Croyle, R.T., (2013). Spatial turn in health research. Science, 339(6126), 1390-1392.

Riley, S., Eames, K., Isham, V., Mollison, D., and Trapman, P. (2015). Five challenges for spatial epidemic models. Epidemics, 1068-71.

Shi, X. Alford-Teaster, J., Onega, T., and Wang, D. (2012). Spatial Access and Local Demand for Major Cancer Care Facilities in the United States. Annals of the Association of

American Geographers, 102(5), 1125-1134.

Shi, X., (2009). A Geocomputational Process for Characterizing the Spatial Pattern of Lung Cancer Incidence in New Hampshire. Annals of the Association of American Geographers, 99(3), 521-533.

Thornton, L.E., Pearce, J.R. and Kavanagh, A.M., (2011). Using Geographic Information Systems (GIS) to assess the role of the built environment in influencing obesity: a glossary. International Journal of Behavioral Nutrition and Physical Activity, 8(1), 71.

Tonne, C. et al., (2006). A case–control analysis of exposure to traffic and acute myocardial infarction. Environmental Health Perspectives, 115(1), 53-57.

Wang, F., Fu, C. and Shi, X., (2015). Planning Towards Maximum Equality in Accessibility of NCI Cancer Centers in the U.S. In: Kanaroglou, P. and Delmelle, E. (eds.), Spatial Analysis in Health Geography, 261-274. London: Routledge.

Wang, J.F., Haining, R., Liu, T.J., Li, L.F., and Jiang, C.S. (2013). Sandwich Estimation for Multi-Unit Reporting on a Stratified Heterogeneous Surface. Environment and Planning A, 45(10), 2515-2534.

Wang, S., Anselin, L., Bhaduri, B., Crosby, C., Goodchild, M. F., Liu, Y., and Nyerges, T. L., (2013). CyberGIS software: a synthetic review and integration roadmap. International Journal of Geographical Information Science, 27(11), 2122-2145.

Yuan, T. et al., (2015). Assessment of the levels of urinary 1-hydroxypyrene and air polycyclic aromatic hydrocarbon in PM2. 5 for adult exposure to the petrochemical complex emissions. Environmental Research, 136219-226.