Annals of Epidemiology 27 (2017) 1–9

Contents lists available at ScienceDirect

Annals of Epidemiology

journal homepage: www.annalsofepidemiology.org

Introduction

Advances in spatial epidemiology and geographic information systems

Russell S. Kirby PhD, MS [a,*], Eric Delmelle PhD [b], Jan M. Eberth PhD, MSPH [c]

[a] Department of Community and Family Health, University of South Florida, Tampa
[b] Department of Geography and Earth Sciences, University of North Carolina-Charlotte, Charlotte
[c] Department of Epidemiology and Biostatistics, University of South Carolina, Columbia

ARTICLE INFO

Article history:
Received 19 November 2016
Accepted 4 December 2016
Available online 8 December 2016

Keywords:
Spatial epidemiology
Geographic information systems (GIS)
Geocoding
Geographic mapping
Health services accessibility
Spatial regression
Spatial analysis
Spatiotemporal analysis

ABSTRACT

The field of spatial epidemiology has evolved rapidly in the past 2 decades. This study serves as a brief introduction to spatial epidemiology and the use of geographic information systems in applied research in epidemiology. We highlight technical developments and highlight opportunities to apply spatial analytic methods in epidemiologic research, focusing on methodologies involving geocoding, distance estimation, residential mobility, record linkage and data integration, spatial and spatio-temporal clustering, small area estimation, and Bayesian applications to disease mapping. The articles included in this issue incorporate many of these methods into their study designs and analytical frameworks. It is our hope that these studies will spur further development and utilization of spatial analysis and geographic information systems in epidemiologic research.

Introduction

Defining spatial epidemiology

All students of epidemiology learn that descriptive epidemiology focuses on the triad of person, place, and time [1]. Although epidemiologic research focusing on place or location historically received considerably less attention, modern epidemiology increasingly incorporates the spatial perspective into research designs and models. Spatial factors have also become prominent features in etiologic research, especially concerning host-vector-agent interactions, but also in guiding social and environmental epidemiologic investigations. Spatial methods are also progressively incorporated into health services research focused on specific diseases, health conditions, or risk factors.

The field of spatial epidemiology has evolved over the years. Elliott et al. [2] identified four types of spatial analyses in epidemiology: (1) disease mapping, (2) geographical correlation studies, (3) risk assessment in relation to point or line sources, and (4) cluster detection and disease clustering. Only a few years earlier, English [3] had defined geographical epidemiology as "the description of spatial patterns of disease incidence and mortality." More recently, in a widely used spatial statistics text, Lawson [4] states that spatial epidemiology "concerns the analysis of the spatial/geographical distribution of the incidence of disease." Here, we argue that spatial epidemiology encompasses research that incorporates the spatial perspective into the design and analysis of the distribution, determinants, and outcomes of all aspects of health and well-being across the continuum from prevention to treatment.

Spatial epidemiology is not synonymous with health and/or medical geography. Spatial epidemiology refers to inquiries that use epidemiologic study designs that involve spatial data or spatially derived information about study subjects, health facilities, or sources of exposure. Health or medical geography, a subdiscipline of human geography, encompasses research applying geographic analytical methods to health, disease, or health care issues. Its distinguishing feature is the primary focus on spatial patterns and context, whereas spatial epidemiology is inherently focused on populations [5]. Although many studies, especially those involving interprofessional research teams, combine the methods of both disciplines in creative and innovative ways, far more often spatial epidemiologic research does not fully incorporate the geographic perspective, and vice versa [6].

The authors have no financial or other material conflicts of interests which might be considered influential to the writing of this article.

* Corresponding author. Department of Community and Family Health, College of Public Health, University of South Florida, 13201 Bruce B. Downs Blvd, MDC56, Tampa, FL 33612. Tel.: +1-813-396-2347; fax: +1-813-905-9998.
E-mail address: rkirby@health.usf.edu (R.S. Kirby).

http://dx.doi.org/10.1016/j.annepidem.2016.12.001
1047-2797/

In this brief overview, we argue for an expanded role for spatial epidemiology within our discipline and demonstrate the importance of a broader scope for spatial perspectives in the study of epidemiology through the series of articles that follow in this symposium on “Geographic Information Systems and Spatial Methods in Epidemiology.”

The concepts of place and neighborhood

Health outcomes are influenced by an interplay of different factors, such as individual attributes, the physical and social environments an individual interacts with, cultural norms, and both the provision and utilization of health services [7]. It is widely recognized that the place where an individual lives or works should be considered as a potential disease determinant [8]. For example, women living in rural areas may have to travel longer distances to reach mammography facilities, potentially leading to a decrease of mammographic breast screening or a delay in their diagnosis [9–11]. Children living in a pedestrian friendly environment where parks and playgrounds are readily accessible are more likely to engage in physical activity, reducing the odds of obesity [12,13]. Along the same vein, higher risk factors for obesity are generally observed in food desert areas, which are characterized by poor-quality food environments and a lack of supply of supermarkets with fresh food [14,15].

Residents located in proximity of major traffic corridors are exposed to particulate matter and diesel exhausts, causing a variety of respiratory and cardiovascular diseases [16]. Individuals residing in high-crime neighborhoods may indirectly develop stress-related behaviors, such as anxiety and higher blood pressure [17,18]. These instances illustrate the breadth of pathways through which both physical and social environments, as well as provision of health services, give rise to health disparities. Documenting the role of the geographic environment where individuals live and interact (often called “activity spaces”) will improve our understanding of health outcomes. This has deep policy implications for local health interventions and resource allocation decisions, ultimately leading to a reduction of health disparities.

A neighborhood is typically defined as the geographic area relevant to the specific health outcome being studied [8]. However, neighborhoods can be delineated by the extent of the individual’s spatial interaction or administratively delimited at the scale at which policies are implemented (county, state). Secondary data sets—such as the American Community Survey published by the U.S. Census Bureau—help to overcome the absence of socioeconomic data in most U.S. medical records [19]. The scale at which the primary or secondary data sets are available often dictates the spatial granularity at which the analysis is conducted [20]. Spatial analyses with individual-level data from public health databases are also often limited by privacy policies required by data managers.

Patient information in the United States may be aggregated at the zip code level to comply with the Health Insurance Portability and Accountability Act [21]. As such, uniformity is assumed within the unit of analysis, but sharp contrasts may occur among adjacent units. Furthermore, the aggregation of neighborhoods does raise the issue of the modifiable areal unit problem; using different boundaries may lead to significantly different analytical results [22]. As a rule of thumb, analyses should be conducted at different scales to test the robustness of the spatial relationships and the effect of different artificial boundaries.

What defines a neighborhood and the concept of scale (aggregated, disaggregated) will influence the choice of methodologies and ultimately impact the results. Spatially based regression, contextual, and multilevel modeling are some of the key methods developed to incorporate neighborhood effects [8,23]. These approaches allow researchers to estimate the impact of neighborhood effects after controlling for individual characteristics.

Role of geographic information systems in spatial epidemiology

Defined by Cromley and McLafferty [24] as “computer-based systems for the integration and analysis of geographic data,” geographic information systems (GIS) can describe, analyze, and predict patterns using feature (cartographic) and attribute data. GIS have been used in many epidemiologic applications, including disease mapping, rate smoothing, cluster or hot spot analysis, and spatial modeling. In its simplest form, GIS is often used to create spatially explicit variables such as availability and accessibility scores (e.g., food access), built environment measures (e.g., land use), environmental exposures (e.g., air pollutant concentrations), and demographic indicators (e.g., percent of persons in poverty). Measuring and describing the extent of spatial relationships is also a key function of GIS, which can be as simple as calculating the distance between two points or as complex as quantifying spatial dependencies in analytic models or identifying locally varying predictors. As described by Thornton et al. [25], GIS offers opportunities to integrate data across multiple databases and spatial scales for display, management, and analysis. In Table 1, we identify several ways to apply GIS tools to human immunodeficiency virus (HIV) research.

In descriptive epidemiology, thematic or color-shaded maps produced with GIS are useful for identifying areas at high risk for epidemics, highlighting population health disparities, examining resource needs, and ultimately, formulating hypotheses that lead to generation of explanatory models. Analytically, GIS tools can be used to explore spatial or spatiotemporal clustering, investigate locally varying relationships, and explicitly model or adjust for spatial dependencies in one’s data.

Just as research has shown that interventions are most effective when implemented at multiple levels [33], we can gain more insight into individual-level outcomes by considering the context in which people live and work [8]. In the context of health care, it is equally important to consider the location and characteristics of where people seek their care (i.e., provider- and system-level factors). Contextual information can be considered in both descriptive and analytic ways. For example, one could explore the distribution of green space or recreational facilities within a region as well as develop a green-space index to predict the probability of meeting physical activity guidelines or individual weight status. When both

Table 1
Functions of GIS and related epidemiologic applications in HIV research

GIS functions	Application to HIV research
Store and measure spatial relationships	Distance between homes of newly diagnosed HIV cases and the closest Ryan White Clinic
Display spatial relationships	Bivariate map showing the relationship between county HIV prevalence and poverty rate
Analyze attribute and feature data simultaneously	Count of the number of HIV cases within 30 miles of each Ryan White Clinic
Manage data from multiple sources	Create a geodatabase with county HIV incidence and prevalence rates, Ryan White Clinic locations, community-based organizations providing HIV services, and county demographic indicators
Identify spatial patterns	Conduct a hot spot analysis of newly diagnosed HIV cases in U.S. counties
Explain spatial patterns	Determine where the prevalence of men who have sex with men is associated with the incidence rate for HIV in U.S. counties using GWR

A variety of studies have used GIS approaches in the context of HIV patient and service provider data [26–32].