Populations, megapopulations, and the areal unit problem

Tom Koch

University of British Columbia, Dept. of Geography (Medical), 1984 West Mall, Vancouver, BC, Canada

A B S T R A C T

Models of epidemic disease and programs for their management require accurate population data as a critical component of most studies. But the traditional definitions of urban places assumed discrete borders and localized populations. The vast increase in urban travel at all scales has raised the problem of how we define those urban populations. This paper reviews the issues as an areal unit problem within the context of the evolving idea of "megaregions" and their defintion.

1. Introduction

Models of dynamic infectious disease events assume a bacterium or virus will spread within a susceptible community of persons to other, easily identified at-risk populations through a network of interpersonal contacts or commercial exchanges linking infected and at-risk populations (Altman, 1995). Studies of disease expansion, either simulations (Orbann et al., 2017) or analyses of reported epidemics, therefore have required three distinct but related data sets. The first two describe clearly defined, stable populations, one infected and the other susceptible (at risk). These must be of sufficient size for each to support first the transfer to and then the propagation of bacterial or viral entities within their jurisdictions (Balcan and Vespignani, 2012). A third dataset describes one or another measure of connectivity permitting disease transfer between those distinct population centers.

A critical question often overlooked in both disease modeling and event analysis is how best to define areal units that accurately describe those populations. It has long been understood that population models employing different areal units will return different results when individual data are aggregated to local, urban, state, or national scales of address (Duncan et al., 1961). In defining a unit for study three things are critical. First, the constancy of this or that jurisdictional boundary encompassing a population and, second, the quality of reportage on that population. Third, and of equal importance, is the degree to which those units reflect a stable population embedded in a network promoting transfer between different but similarly defined units.

Here the problem of variable areal units and their connections is reviewed in relation to infectious disease and epidemic events. Particular attention is paid to the current literature on "megapopulations" as an emerging areal population unit of potential importance in disease studies.

2. The Urban

Cities have long served as a principal areal unit in disease studies, both as sites of infection and as loci of disease transfer. It was for this reason that in the fourteenth century quarantine programs first banned travelers from infected to epidemic-free port cities in an attempt to protect at-risk populations (Rosen 1993, 43–45). In the late eighteenth century local health boards, constituted in part to address yellow fever outbreaks in the U.S., were charged with collecting primary data on local disease incidence, advising citizens on how best to avoid contagion, and on enacting measures to manage where they could not prevent an outbreak in their cities (Koch, 2017, 33–38).

Similarly the assumption has always been that human travel–international, national, and local–is a principal vector for disease transmission between population centers. It was a strong motif in Holbein the Younger's famous sixteenth century *Dance Macabre's* commentary on plague (Holbein, 1538/, 1971). In the first modern global pandemic, cholera incidence was mapped in the nineteenth century, city by city along existing sea and land routes (Brigham, 1832; Koch, 2017, 260–171). Contemporary studies have focused less on the nature of urban places as disease catchments and more on the networks that connect them (Brockmann and Hufnagel, 2006); Balcan et al. (2009)). Principal attention has centered on international airline passenger flights (Colizza, Barrat, Barthelemy and Vespignani, 2006) and international cargo ship carriage (Kaluza et al., 2010) as disease vectors (Tatem et al., 2012; Teran-Romer et al., 2017). Airline travel has been implicated in the national and international transfer of, among other diseases, influenza (Kahn et al., 2009), dengue (Lana et al., 2017), West Nile Virus (Koch and Denike, 2007) and Zika virus (Massad et al., 2016).

In these studies a pathogen or its vector was reportedly transported from an infected city (Beijing, Mexico City, Sao Palo, Brazil) to other susceptible national urban centers (London, Los Angeles, New York

City) with a density sufficient to sustain emigrant microbial populations (Veldtman and McGeoch, 2004). What has been unclear in these studies has been the appropriate definition of the designated urban areas themselves. Even less attention has been given to the process of secondary transfers from principal urban centers (London, Los Angeles, New York) to secondary or tertiary locations (for example, Bristol, UK; San Diego, CA; Harford, CT).

3. Urban populations

The expansion of travel volume at all scales and by various modes has led some to question whether traditionally defined urban centers serve today as appropriate areal units. It is no longer clear that, "at any given point in time each individual can be assigned uniquely to a single local context within which they experience the sufficiently close, sustained interaction necessary for disease transmission" (Watts et al., 2005: 1158). If individual assignments are unclear then so, too, is the utility of the urban designations based upon them.

Increased regional travel capacity permitted by high speed road networks has resulted in the construction of a number of areal catchments that include but extend beyond traditional urban boundaries (Robert et al., 2016). The United States Census Bureau, for example, has created a series of commercially based, interlocking areal units describing the regionalization of previously discrete urban populations. At present there are 388 principal "Metropolitan Statistical Areas" (MSAs), each including at least one city with a population greater than 50,000 persons (OMB Bulletin No. 15-01, 2015). There are, in addition, 929 "Core Based Statistical Areas" (CBSAs) centered on cities or towns with populations between 10,000 and 50,000 persons. There are as well 166 combined statistical areas (CSAS) that may include multiple MSAs whose constituent urban centers are presumably joined along shared commercial exchange pathways (OMB, 2010).

The size and population of the resulting jurisdictions different for each areal unit. For example, The US Census Bureau reported a 2017 population for Boston, MA, of 685,094 living in a densely settled area of 148 km². As a metropolitan area including proximate towns and suburbs, however, Boston's population in 2017 was reported as 4628,910 persons across 11,700 km². As the center of a Census Bureau primary statistical area encompassing New Hampshire, Rhode Island and half of Massachusetts, the Boston-centered population swells to 8099,575 person in an area of 27,600 km². In the reportage of disease events it is typically unclear, however, precisely which "Boston" is being included.

Without careful attention to the precise definition of the areal unit in which disease incidence is described the choice of population denominators in disease models–for example, the gravity model is at best uncertain. Worse, as the resolution and scale of areal statistical units changes from local populations to larger, typically economically defined regions, the result is discontinuous with blank spaces appearing in their mapping (Fig. 1). This is especially true in less populated western regions but also occurs in otherwise densely settled regions, for example the Mid-Atlantic, Southwest, and West Coast states.

4. Megaregions

To better reflect modern realities some urban researchers have proposed the existence of new "megapopulations" centered, like the Census Bureau's CBSAs, on major metropolitan hubs linked to an orbit of sub-centers in relatively discrete, dynamic systems of semi-unique, exchange (Adams and Kapan, 2009; Nelson and Rae, 2016). "Megaregions" have been employed in studies of Ebola in West Africa (Dudas et al., 2017), and measles in the Sub-Sahara (Ferrari et al., 2008). There is, however, no single, accepted heuristic dictating their construction. Nor has there been a careful comparison of the relative utility of one or another megapopulation model in economic, epidemiological, or population studies.

Three fundamentally different approaches have been proposed.

Each is based on linkages presumed to be sufficiently strong to reflect a single economic or population area. The first employs the circulation of bank notes (Viboud et al., 2006; Brockmann and Hufnagel, 2006) as a measure of intra-urban connectivity. The data is based on the marking of and then later location of those marked bills in "wheresgeorge.com," an online public reportage system, A serious limit of this approach is that it does not include electronic transfers or credit cards that characterize much of contemporary commerce. Nor is it clear that the identification of bank notes at any location is the result of one, several, or many individual transactions. Finally, because the data relies on citizen online reportage the data is limited by participant knowledge of the program, an ability to access it online and a willingness to do so.

A second approach employs the origin-destination records of mobile phone calls as proxies for population dynamics (Gonzalez et al., 2008; Simini et al., 2012; Expert et al., 2011). As a means of describing intra-urban activity the results are limited by the percentage of the population having and using cell phones and the accuracy of systems that can calculate, report, and map the density of intra-urban calls. More importantly, perhaps, the result reflects only distanced connectivity and not actual interactions that might promote actual bacterial or viral transfer. So while useful in other study areas this approach is unlikely to be of real utility in disease studies.

4.1. Commuter data

Others have suggested using commuter travel data to describe actual interpersonal linkages between different populations and sub-populations (Adams and Kapan, 2009). The United States Census Bureau's American Community Survey (ACS) annually questions approximately 143 million citizens living at 3.5 million addresses where each worked "last week." Respondent data are available in one, three or five year collections in which, for the five-year period ending in 2010, 4156,426 origin-destination vectors were documented across 74,002 U.S. census tracts (U.S. Census Bureau, 2017). The resulting data has been reformatted for 2005–2010 by researchers to create a flexible, spatially grounded, mappable database (Nelson and Rae 2016). It includes for each origin-destination vector a measure of "flow," the number of persons travelling between any two specific census tracts, and the Euclidian distance between the centroids of each census tract. The resulting dataset can be analyzed (and projected) in most GIS programs including ArcGIS and QGIS, the program in which it was constructed, as well as a range of statistical packages.

The analysis of precise commuter data collected at the census tract level offers significant advantages for epidemiologists and medical geographers. First, to the extent that travel is a vector for disease propagation it presents a potentially applicable medium for the study of disease transfer between populations at different scales. Second, because the data is available at the level of census tracts it can be aggregated to varying scales. Third, because it includes travel at all distances it permits a study of the potential of "intervening opportunities" (Stouffer, 1940) and other forms of "impedance" (Sallah et al., 2017) to affect the transfer of disease agents across and between populated areas. On the basis of linked networks of commuter exchange it seems likely that disease may expand across these links.

That said, the dataset while extensive is not encyclopedic. It does not include tourism and other non-work travel flows in and out of major destinations (Chicago, Las Vegas, Los Angeles, Miami, New York). And because the data is limited to U.S. commutes it does not include volume across northern and southern borders (Buffalo-Toronto, Vancouver-Seattle, San Diego and Tijuana, for example) is excluded. This potentially creates a kind of Modifiable Areal Unit Problem (MAUP) in border areas (Caillaud et al., 2013). Finally, as others have noted, cargo traffic provides a separate but still potentially significant vector for the accidental transport of disease vectors (Kaluza et al., 2010). A parsing of commuter data will not, however, include that avenue of microbial transfer.