Applied Geography 71 (2016) 123–132

Contents lists available at ScienceDirect

Applied Geography

journal homepage: www.elsevier.com/locate/apgeog

An approach for estimating vaccination coverage for communities using school-level data and population mobility information

Paul L. Delamater [a,*], Timothy F. Leslie [a], Y. Tony Yang [b], Kathryn H. Jacobsen [c]

[a] Department of Geography and Geoinformation Science, George Mason University, Fairfax, VA, United States
[b] Department of Health Administration and Policy, George Mason University, Fairfax, VA, United States
[c] Department of Global and Community Health, George Mason University, Fairfax, VA, United States

ARTICLE INFO

Article history:
Received 9 November 2015
Received in revised form
29 March 2016
Accepted 14 April 2016
Available online 5 May 2016

Keywords:
Vaccination
Immunization
Mobility
Interpolation
Pertussis

ABSTRACT

Childhood vaccination data are made available at a school level in some U.S. states. These data can be geocoded and may be considered as having a high spatial resolution. However, a school only represents the destination location for the set of students who actually reside and interact within some larger areal region, creating a spatial mismatch. Public school districts are often used to represent these regions, but fail to account for private schools and school of choice programs. We offer a new approach for estimating childhood vaccination coverage rates at a community level by integrating school level data with population commuting information. The resulting mobility-adjusted vaccine coverage estimates resolve the spatial mismatch problem and are more aligned with the geographic scale at which public health policies are implemented. We illustrate the utility of our approach using a case study on diphtheria, tetanus, and pertussis (DTP) vaccination coverage for kindergarten students in California. The modeled community-level DTP coverage estimates yield a statewide coverage of 92.37%, which is highly similar to the 92.44% coverage rate calculated from the original school-level data.

1. Introduction

The refusal or delay of childhood vaccinations has been identified in both the popular press and scientific literature as an increasing public health concern across the United States (Dempsey et al., 2011; Hayes, 2015; Hughes, 2006; Omer et al., 2006; Park, 2008). One common approach to encouraging or mandating childhood vaccination is through school entry vaccine requirements. Although there is strong scientific evidence for the safety and effectiveness of immunizations, concerns about the risks associated with vaccination (and the increasing number of required and recommended childhood vaccines) have gained momentum in various popular and social media outlets. In the 2014–2015 school year, nearly 2% of all U.S. kindergarteners entered school with a nonmedical exemption (NME) from vaccination requirements (Seither et al., 2015).

Each U.S. state has autonomy to establish vaccination requirements for children enrolled in their school system (Salmon et al., 2005), and the level of restrictiveness varies from state to state (Yang & Debold, 2014). As a result, coverage among incoming kindergarteners for common vaccines such as diphtheria, tetanus, and pertussis (DTP); measles, mumps, and rubella (MMR); and varicella fluctuates substantially throughout the U.S. (Seither et al., 2015). Some states, such as Arkansas and Colorado, have DTP and MMR coverage rates that are far below the thresholds required to ensure herd immunity (Seither et al., 2015), the percentage of the population that must be vaccinated in order to prevent outbreaks in the general population, which includes those with health problems that contraindicate vaccination.

Vaccine-related behavior also fluctuates considerably at a community or regional scale *within* states (Atwell et al., 2013; Birnbaum, Jacobs, Ralston-King, & Ernst, 2013; Lieu, Ray, Klein, Chung, & Kulldorff, 2015; Omer et al., 2008; Yang, Delamater, Leslie, & Mello, 2016). Examining vaccine coverage rates at the state level, using data such as those provided by the National Immunization Survey[1] can be helpful, but obscures significant local variation in coverage rates within states and the associated risks of infectious

* Corresponding author. Dept of Geography & GeoInformation Science, 4400 University Drive, MS6C3, Fairfax, VA 22030, United States.
E-mail addresses: pdelamat@gmu.edu (P.L. Delamater), tleslie@gmu.edu (T.F. Leslie), ytyang@gmu.edu (Y.T. Yang), kjacobse@gmu.edu (K.H. Jacobsen).

[1] Higher spatial resolution NIS data (by zip code) are available as a restricted dataset from the U.S. Centers for Disease Control and Prevention (CDC) Research Data Center. Access to this information requires payment of management fees. However, the relatively low number of children surveyed per state for the NIS (e.g., 23,248 children were surveyed in the entire U.S. in 2013) does not provide sufficient coverage for detailed spatial analysis of within state variation.

http://dx.doi.org/10.1016/j.apgeog.2016.04.008
0143-6228/

disease outbreaks. The ability to identify highly vaccinated and undervaccinated neighborhoods, school districts, and other types of local communities or regions is important for policy makers, public health officials, and medical providers who aim to make informed, evidence-based decisions and recommendations for appropriate interventions and care strategies related to vaccine-preventable diseases (Lieu et al., 2015).

Several states make vaccination coverage information publicly available for specific checkpoint grade levels, such as kindergarten and sixth grade. While some states provide data aggregated by the county (e.g., Kansas) or public school district (e.g., New Mexico and Tennessee), others allow access to school-level data (e.g., California, Oregon, and Virginia). School-level data can be considered high spatial resolution information, since schools are located at precise locations, which can be georeferenced and integrated into a Geographic Information System (GIS). Analysis of the school-level data has the potential to provide important knowledge regarding regional or local trends in vaccination coverage. Yet, the particular nature of this data creates an interesting set of problems for GIS and spatial analysis efforts. The main limitation is that schools represent discrete *point* locations in space with a set of attributes, while the students who attend the schools reside and interact within some larger *areal* region, such as a school district. The school-level point data only represent a terminus location for a set of individuals whose residences are distributed spatially throughout some region. Previous estimates of vaccination coverage rates for local communities have yet to fully account for the spatial mismatch between the school-level data and the area served by a particular school.

In this paper, we present a new data-integration method for estimating community-level vaccination coverage rates using school-level data and population mobility information. Our approach accounts for the spatial mismatch school-level point representation and the areal regions in which students reside. We incorporate aspects of both interpolation and aggregation, but tailor our approach for the particularities of U.S. school attendance geography to estimate the coverage rate for the DTP vaccine for California (CA) kindergarteners entering school in 2014. California experienced outbreaks of pertussis, also known as whooping cough, in 2010 and 2014 (California Department of Public Health, 2015), which may signal that the vaccine coverage rate had dropped below the herd immunity threshold in some communities. We estimate DTP coverage percent for census blocks, the smallest areal units of census geography, providing a highly detailed spatial representation of community-level vaccination coverage throughout the state. The resulting map and areal-level data provide geographic visualization and spatial analysis opportunities that are not available when relying solely on the school-level point data.

2. Background

The two most commonly-used approaches for estimating areal data from a point observations are: (1) interpolation, which estimates values for locations without data using a set of known values from the surrounding region and (2) aggregation of the point data to preexisting areal units.

In the interpolation approach, known values from a point dataset are used to estimate values at unsampled locations. A significant problem arises if using interpolation to estimate vaccination coverage from school-level data. Interpolation assumes that the characteristic being studied is continuously distributed throughout space and the values at the known point locations represent a set of sample points drawn from this spatially distributed phenomenon. However, this assumption does not match the reality of the school-level vaccination coverage data. In this case, a school is not a sample point drawn from a continuous surface, but instead is a destination point for the students who reside at locations distributed throughout some region.

In the aggregation approach, a set of areal units is chosen, and then the points falling within each unit are aggregated and summarized into a single value representing the entirety of that unit. For estimating vaccination coverage rates, two important problems arise in this approach. First, if large areal units such as counties are chosen as the aggregation unit, the size of the areal units has the potential to obscure the geographic detail available from the school-level data. Second, if small areal units such as U.S. Census block groups are chosen, there is a possibility that some units will not contain a school and, therefore, require an additional decision rule and processing step to assign a vaccination coverage percentage to these units.

In the U.S., the problems associated with the size of the areal units are exacerbated by the particularities of primary and secondary (grades K-12) school attendance geography, which are summarized below. Generally, schools in the U.S. fall into one of three broad categories, (1) public schools, which are publicly-funded, free to attend, and have defined geographic service areas, (2) charter (and magnet) schools, which are publicly-funded and free, but only in some cases have specifically defined service areas, and (3) private schools, which do not receive public funds, may charge tuition for attendance, and do not have defined service areas (United States Network for Education Information, 2008).

The public school system provides much of the overall education in the U.S., and attendance at public schools is tied to the residential locations' of the students (Danielsen, Harrison, & Zhao, 2014). A public school district defines the areal extent of the region or catchment in which students must attend a particular public school or set of schools. When multiple public schools are located within a single district, the district is further divided into subregions that delineate the particular service area for each school. Historically in the U.S., public schools were distributed systematically to serve a single neighborhood or community and thus minimize travel for students (Bell, 2007). More recently, the link between geography, distance, and school attendance has been weakened via increases in the availability of both public and private school choice options (Ely & Teske, 2015; Henig, 2009). Public school choice provides parents with the option of sending their children to a school that is not located in the school district or subregion in which they reside (Ely & Teske, 2015).

The use of public school district boundaries as an aggregation unit for the school-level vaccination data may appear to be a reasonable option. However, this approach generates at least three potential complications related to student attendance geography in the U.S. First, private and charter school student populations are not bound by the public school district geography. A private or charter school located within a public school district's boundaries may have numerous students who reside outside of that particular district. Second, in states offering public school choice programs, students may attend a school located outside the public school district or subregion assigned to their residence. Third, some school districts have multiple public schools within their boundaries and do not provide the detailed subregion boundary information for each school in a useable format. Supplemental Fig. 1 provides examples of the potential complications arising from public school districts as an aggregation unit for the school-level data.

3. Community-level vaccination coverage

Our approach for estimating vaccination coverage at a community-level is based on two extensions of a traditional