

9. Lu X, Chittaganpitch M, Olsen SI, Mackay IM, Sloots TP, Fry AM, et al. Real-time PCR assays for detection of bocavirus in human specimens. *J Clin Microbiol.* 2006;44:3231–5. DOI: 10.1128/JCM.00889-06

Address for correspondence: Dean D. Erdman, Centers for Disease Control and Prevention, 1600 Clifton Rd NE, Mailstop G04, Atlanta, GA 30333, USA; email: dde1@cdc.gov

## Assessment of Reporting Bias for *Clostridium difficile* Hospitalizations, United States

**To the Editor:** Burckhardt et al. (1) recently reported on *Clostridium difficile*-associated disease (CDAD) in Saxony, Germany. In contrast to the observation by Wilcox and Fawley in the United Kingdom (2), the report from Germany argued against a reporting bias for gastroenteritides as a cause of the observed increase in the

incidence of CDAD diagnoses from 2002 through 2006. To explore this issue further, I examined the potential influence of such reporting bias on the observed increase in the incidence of hospitalizations of patients with CDAD in the United States from 2000 through 2005.

In the 2000–2005 data from the National Inpatient Sample data from the Agency for Healthcare Research and Quality (3,4), I identified hospitalizations for gastrointestinal infections caused by *C. difficile*, *Salmonella*, rotavirus, and other unspecified infectious agents, using the corresponding diagnosis codes from the International Classification of Diseases, 9th Revision, Clinical Modification. I obtained censal and intercensal data on the numbers of the U.S. population from 2000 through 2005 from the U.S. Census Bureau (5). Based on these records, I calculated hospitalization incidence for each of the infectious causes.

Annual incidence of CDAD increased from 49.2 to 101.6 per 100,000 population within the period examined. Within the same time frame, the incidence of CDAD as the principal diagnosis also more than doubled, increasing from 11.6 to 25.8 hospitalizations

per 100,000. Although the incidence of hospitalizations for *Salmonella* infections per 100,000 population remained stable, rotavirus infection showed a slight increase (from 10.8 to 14.5) as did other infectious gastroenteritides (from 38.9 to 49.9/100,000) (Figure). Thus, although a slight increase in the incidence was exhibited, a reporting bias for gastroenteric infections with organisms other than *C. difficile* does not appear to account fully for the observed doubling of the overall incidence of hospitalizations with CDAD in the United States from 2000 through 2005.

**Marya D. Zilberberg\*†**

\*University of Massachusetts, Amherst, Massachusetts, USA; and †EviMed Research Group, LLC, Goshen, Massachusetts, USA

DOI: 10.3201/eid1408.080446

### References

- Burckhardt F, Friedrich A, Beier D, Eckmanns T. *Clostridium difficile* surveillance trends, Saxony, Germany. *Emerg Infect Dis.* 2008;14:691–2.
- Wilcox M, Fawley W. Viral gastroenteritis increases the reports of *Clostridium difficile* infection. *J Hosp Infect.* 2007;66:395–6. DOI: 10.1016/j.jhin.2007.05.010
- Agency for Healthcare Research and Quality. Nationwide Inpatient Sample (NIS). Healthcare Cost and Utilization Project [cited 2008 Apr 1]. Available from <http://www.hcup-us.ahrq.gov/nisoverview.jsp>
- Agency for Healthcare Research and Quality. HCUPnet. Healthcare Cost and Utilization Project (HCUP). 2000–2004 [cited 2008 Apr 1]. Available from <http://hcupnet.ahrq.gov>
- US Census Bureau [cited 2008 Apr 1]. Available from <http://www.census.gov>

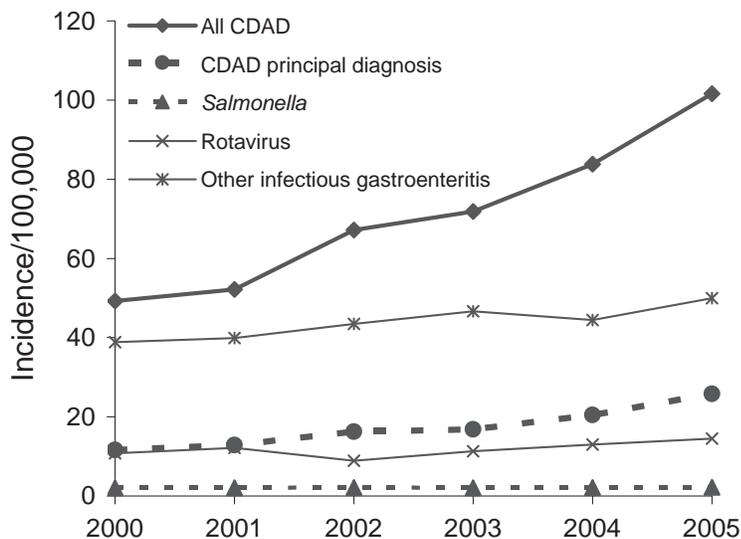


Figure. Annual incidence per 100,000 population of all hospitalizations for *Clostridium difficile*-associated disease (CDAD) compared with hospitalizations for a primary diagnosis of CDAD and with gastroenteritides caused by *Salmonella*, rotavirus, and other unspecified infectious agents, United States, 2000–2005.

Address for correspondence: Marya Zilberberg, EviMed Research Group, LLC, PO Box 303, Goshen, MA 01032, USA; email: [marya@evimedgroup.org](mailto:marya@evimedgroup.org)

## Improving Methods for Reporting Spatial Epidemiologic Data

**To the Editor:** A recent perspective in this journal (1) pointed out problems with the present, county-referenced system for reporting spatial epidemiologic data. Problems identified included coarse spatial resolution of county-referenced data and differences across the United States in size of counties, making data for the western part of the country coarser in resolution than data for the eastern part. Eisen and Eisen correctly pointed out that these problems complicate spatial analyses of epidemiologic data (1). However, the solutions that they propose, referencing epidemiologic data to ZIP codes or census tracts, partially solve only the first problem.

The problem of regional differences in spatial resolution of county-referenced data is, unfortunately, reflected in counties, ZIP codes, and census tracts, as shown in plots of nearest-neighbor distances among unit centroids as a function of longitude (Figure). Because all 3 regionalizations are based on human populations, the much greater population density in the eastern United States creates finer scale dispersion in the east. Thus, a shift to ZIP codes or census tracts does nothing to resolve

the problem of regional differences in spatial resolution.

The problem of coarse spatial resolution is only partially addressed by the ZIP code or census tract solution. ZIP codes and census tracts cover fixed areas and can misrepresent the spatial precision of epidemiologic records. A traveling salesperson who covers the state of Wyoming each week would be represented identically as his or her next-door neighbor who is housebound, although spatial precision differs considerably between the 2 persons. Precision of the housebound neighbor could be better represented than county, ZIP code, or census tract. ZIP codes and census tracts change periodically, and ZIP codes do not have defined spatial extents per se (2). Thus, a better and more flexible solution is needed.

The biodiversity world has already addressed this challenge. The point-radius method for georeferencing locality descriptions (3) estimates a best guess for the exposure site (e.g., residence, workplace) but describes uncertainty in that georeference is a radius that expresses spatial uncertainty in the record (i.e., compare our traveling salesperson with his or her housebound neighbor) and in translation into geographic coordinates (including uncertainty in the locality descriptor, spatial footprint of the locality described, imprecision in the locality identified, and any other

sources of imprecision). Point-radius georeferences are easily recorded and reported, are consistent and reproducible, and are more precise and considerably more stable than ZIP codes or census tracts.

As an example of how the point-radius method would be applied, the locality for our traveling salesperson would be assigned to his or her house, but the error radius would be 360 km (based on corner-to-corner distance across Wyoming). The housebound neighbor might have a similar set of coordinates (next door), but the error radius might be 0.1 km (breadth of the house plus the imprecision of the global positioning system unit). When a researcher uses these data, he or she might wish to analyze occurrence of this disease with a spatial precision of 1 km; e.g., applying a filter to exclude those data records too imprecise for this study, he or she would exclude the data record for the salesperson (because the salesperson may have contracted the disease in another sector of the state) but include that for the housebound neighbor. Alternatively, the researcher may include variable degrees of precision in the analysis according each to record a precision or certainty corresponding to its error radius, as in recent spatial analyses of Marburg virus transmission risk (4) and climate change effects on plague and tularemia transmission (5).

How specifically would this meth-

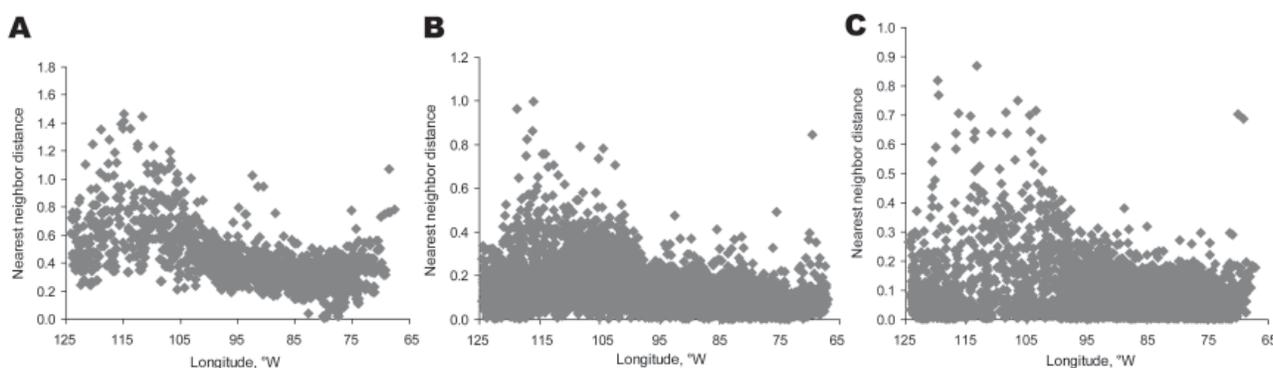


Figure. Longitudinal patterns in nearest-neighbor distances for A) counties, B) ZIP codes, and C) census tracts across the lower 48 United States, showing trends toward greater spacing among districts in the western United States compared with the eastern United States in all 3 regionalizations.

od be implemented in public health surveillance? If data are to be captured initially on paper, the data recorder would simply record the focal point of the person's activities (usually a residence) and an approximate description of the person's movements (e.g., broadly across the state, housebound, within 20 miles). These descriptions are easily georeferenced post hoc by using recently developed software tools (e.g., Biogeomancer, [www.biogeomancer.org/](http://www.biogeomancer.org/)). A more promising solution, if initial data capture is electronic, would be adaptation of some of these software solutions to the public health challenge. A flexible-resolution map with political boundaries, named places, and roads and streets could enable immediate digitization of the central point and the error radius even during direct consultation with the patient (when feasible).

The point-radius approach is novel to most epidemiologic applications but offers considerable advantages. When fine-resolution data are available, researchers will have this more precise information and can distinguish it from coarser resolution data; when actual data are coarser, this information is also expressed. Researchers will be able to filter epidemiologic occurrence information to retain those data that are sufficiently precise for particular applications, thus offering a considerable improvement over any of the 3 polygon-based approaches (ZIP codes, census tracts, and counties). Thus, the recent publication cited (1) got the question right but the answer wrong.

#### A. Townsend Peterson\*

\*University of Kansas, Lawrence, Kansas, USA

DOI: 10.3201/eid1408.080145

#### References

1. Eisen L, Eisen RJ. Need for improved methods to collect and present spatial epidemiologic data for vectorborne diseases. *Emerg Infect Dis.* 2007;13:1816–20.
2. Grubestic TH, Matisziw TC. On the use

of ZIP codes and ZIP code tabulation areas (ZCTAs) for the spatial analysis of epidemiological data. *Int J Health Geogr.* 2006;5:58. DOI: 10.1186/1476-072X-5-58

3. Wieczorek J, Guo Q, Hijmans R. The point-radius method for georeferencing locality descriptions and calculating associated uncertainty. *International Journal of Geographical Information Science.* 2004;18:745–67. DOI: 080/13658810412331280211
4. Peterson AT, Lash RR, Carroll DS, Johnson KM. Geographic potential for outbreaks of Marburg hemorrhagic fever. *Am J Trop Med Hyg.* 2006;75:9–15.
5. Nakazawa Y, Williams R, Peterson AT, Mead P, Staples E, Gage KL. Climate change effects on plague and tularemia in the United States. *Vector Borne Zoonotic Dis.* 2007;7:529–40. DOI: 10.1089/vbz.2007.0125

Address for correspondence: A. Townsend Peterson, Natural History Museum and Biodiversity Research Center, University of Kansas, Dyche Hall, 6th Floor, Lawrence, KS 66045, USA; email: [town@ku.edu](mailto:town@ku.edu)

**In Response:** In his comment, Peterson reiterates the need for improved methods for collecting and presenting spatial epidemiologic data for vector-borne diseases (1). He agrees with us that lack of reliable data on probable pathogen exposure site is an obstacle to the development of predictive spatial risk models (2). In that article we noted, “New methods are urgently needed to determine probable pathogen exposure sites that will yield reliable results while taking into account economic and time constraints of the public health system and attending physicians.” Peterson suggests that the point-radius method is a viable solution to this problem. Unfortunately, its practical implementation for vector-borne diseases is neither reliable nor cost-effective.

With regard to practical implementation of the point-radius method in a public health setting, Peterson states, “If data are to be captured

initially on paper, the data recorder would simply record the focal point of the person's activities (usually a residence) and an approximate description of the person's movements (e.g., broadly across the state, housebound, within 20 miles)” (1). We find a number of serious problems with this approach to determining probable sites of pathogen exposure, primarily that meaningful use of the point-radius method 1) will require not only recording detailed movements during the perceived window of opportunity for pathogen exposure but also weighting of risk by activity type and, for some vector-borne diseases, time of day; and 2) will require the public health community to allocate resources to in-depth interviews conducted by specially trained personnel.

Our first concern is that Peterson's scenario does not distinguish between a car trip to the mall at noon and spending an evening on the golf course. In reality, one activity presents minimal risk for exposure to mosquitoes infected with West Nile virus, whereas the other is a potential high-risk activity. Giving equal weight to the movements represented by these activities will assuredly produce an unreliable result for probable pathogen exposure site. Other issues are patient recall and reluctance to provide information on movement patterns and specific activities. Peterson's suggestion that the data recorder would simply record the focal point of the person's activities and an approximate description of the person's movements is therefore a grossly oversimplified solution to a complex public health problem.

With regard to the second concern, the average physician likely lacks the knowledge, time, and training in vector-borne disease epidemiology and ecology needed to accurately assess when and where risk for pathogen exposure occurred. To be of use, the method will require in-depth patient interviews by specially trained personnel from local or state health