Ann Surg Oncol (2023) 30:6987–6989
https://doi.org/10.1245/s10434-023-14140-9

Annals of
SURGICAL ONCOLOGY
OFFICIAL JOURNAL OF THE SOCIETY OF SURGICAL ONCOLOGY

ASO RESEARCH LETTER

Evaluating Geographic Health Disparities in Cancer Care: Example of the Modifiable Areal Unit Problem

Constance P. Fontanet, MPH[1], Heather Carlos, MS[1], Julie E. Weiss, MS[2], Maria C. Gil Diaz, BS[1], Xun Shi, PhD[3], Tracy Onega, PhD[4], and Andrew P. Loehrer, MD, MPH[1,2,5]

[1]Geisel School of Medicine at Dartmouth, Hanover, NH; [2]Dartmouth Cancer Center, Lebanon, NH; [3]Department of Geography, Dartmouth College, Hanover, NH; [4]University of Utah, Huntsman Cancer Institute, Salt Lake City, UT; [5]Department of Surgery, Dartmouth-Hitchcock Medical Center, Lebanon, NH

As geographic information systems become more accessible to researchers, the role of geography in health research is becoming increasingly salient[1] and can lead to misapplication of important geographic concepts. One such example is the modifiable areal unit problem (MAUP), a type of ecologic fallacy that can lead to different results depending on the areal unit chosen for analysis.[2] This may have a disproportionate impact on rural areas versus urban areas, with the former suffering from highly variable rates for the same area due to lower population density. To truly understand how geography affects health outcomes, we must understand how the units of analysis we select influence our results. In this analysis, we demonstrated the differences in results stemming from our use of different areal units to evaluate disparities in late-stage presentation of patients with breast cancer.

METHODS

We identified patients with incident breast cancer in the Indiana State Cancer Registry from 2010 to 2015. The geospatial heterogeneity of late-stage breast cancer was analyzed at three different geographic levels: county, census tract, and block group. Counties are administrative entities that vary widely in both area and population size. Census tracts are statistical subdivisions of a county with a target population of 4000 inhabitants.[3] Block groups are subdivisions of a census tract containing 600 to 3000 people. The Global Moran's *I* statistic was used to investigate overall clustering of location.[4] We illustrated the potential impact of using different areal units with maps of rates for late-stage breast cancer at each level across Indiana. Given the de-identified nature of the Indiana State Cancer Registry dataset, the study was exempt from institutional review board (IRB) review. Areas with case counts fewer than six were "suppressed" and not shown on the map due to privacy concerns.

RESULTS

Our sample included 30,604 patients with breast cancer residing in 4814 block groups, 1511 census tracts, and 92 counties. We observed similar proportions of late-stage presentation at the levels of block group (15.2 %), census tract (15.3 %), and county (14.5 %). At the block group level, low case counts led to highly variable rates and suppression of data presentation (Fig. 1). At the county level, we were unable to appreciate local variation in late-stage presentation rates. For example, maps of areas with low rates at the county level (e.g., Indianapolis) obscured the high rates of late-stage presentation visible within the county at the census tract level. Our analysis showed decreasing variance and spatial autocorrelation with increasing size of area and loss of statistical significance when the level of county was evaluated (Table 1).

DISCUSSION

Understanding how rates of late-stage breast cancer vary geographically is essential to formulating and targeting interventions. We found that using block group-level

First Received: 16 May 2023
Accepted: 30 July 2023
Published online: 1 September 2023

A. P. Loehrer, MD, MPH
e-mail: andrew.p.loehrer@hitchcock.org

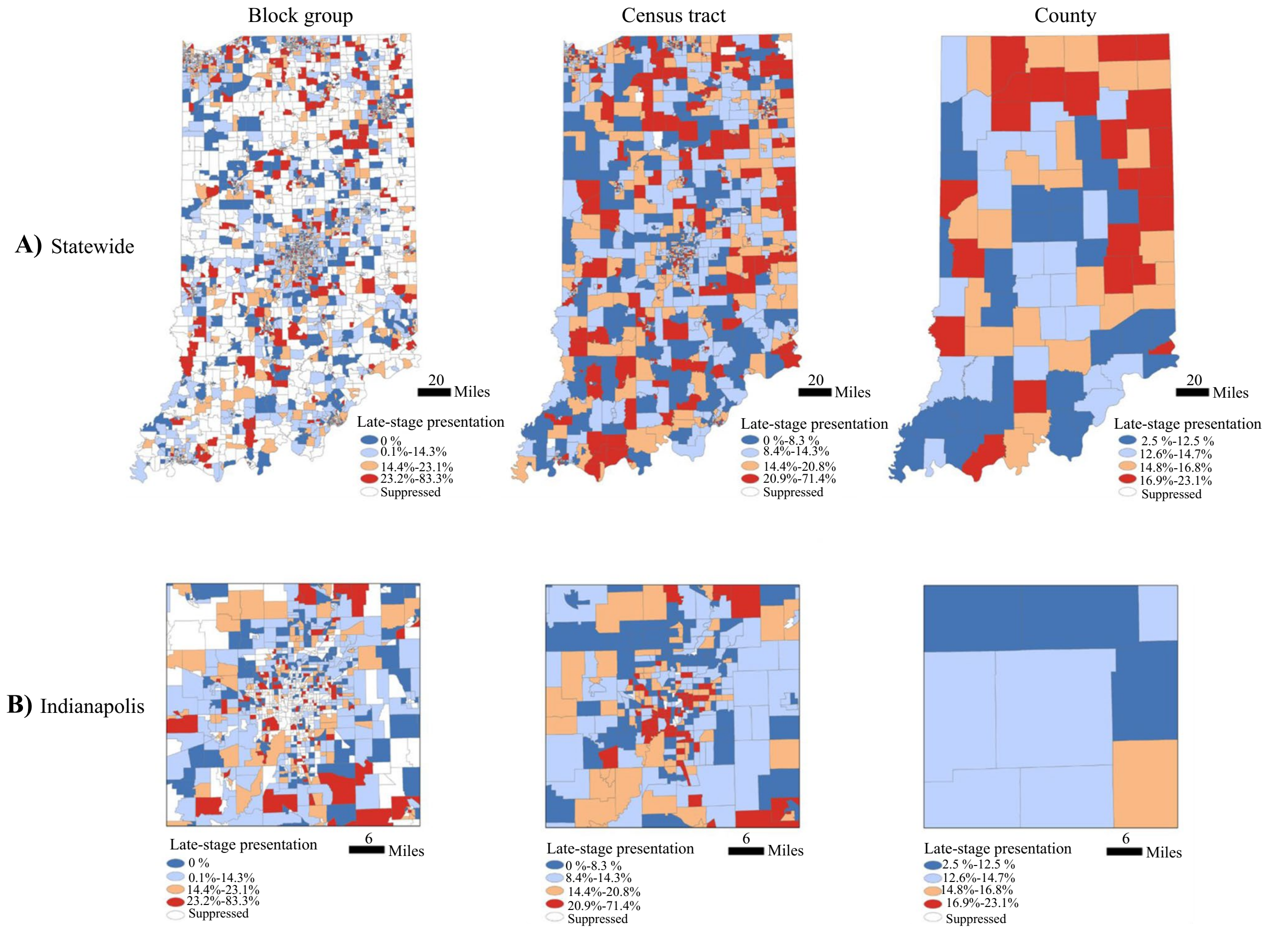

FIG. 1 Percentage of patients presenting with late-stage breast cancer in Indiana from 2010 to 2015. This figure shows rates of late-stage presentation for breast cancer at the block group level, census tract level, and county level throughout the state of Indiana (first row) and Indianapolis (2nd row). Rates for areal units with case counts <6 were suppressed

TABLE 1 Mean rates of late-stage presentation and clustering measurement by areal unit[a]

	Block group	Census tract	County
Mean rate (%)	15.2	15.3	14.3
Variance	19.1	11.0	3.8
Global Moran's Index	0.02	0.04	0.09
p Value	<0.001	<0.001	0.19

[a]In the Global Moran's Index, 0 indicates no autocorrelation, 1 indicates perfect clustering, and –1 indicates perfect dispersion

data prevented meaningful evaluation of rate changes due to small denominators, whereas using county-level data obscured potentially important within-county differences. These analyses showed the importance of empirically examining the impact of selecting different areal units rather than simply using those available in a given dataset. Health services researchers are increasingly using different area-level measures of socioeconomic factors and should consider the implications of the areal unit being applied, including fluctuating incidences of disease with small populations or masking of heterogeneity with larger areas.[5] Researchers also should explore alternative methods that minimize the MAUP. For example, the restricted and controlled Monte Carlo process disaggregates polygon-level data (e.g., block group, census tract, or county) to achieve mapping aggregate data at an approximated individual level based on pre-existing population distributions, transforming area-based data into point-based data and thus avoiding the MAUP.[6] Given the potential for inconsistent if not conflicting results, spatial analyses must thoughtfully approach the most appropriate, accurate, and relevant methods when evaluating geospatial differences in care.

ACKNOWLEDGMENTS Support was provided by the GeoSpatial Resource, a section of the Biostatistical and Bioinformatics Shared Resource at the Dartmouth Cancer Center, with NCI Cancer Center support Grant 5P30CA023108. The research reported in this publication also was supported by the National Cancer Institute of the National Institutes of Health under Award No. NCI K08 CA263546.

REFERENCES

1. Wang F. Why public health needs GIS: a methodological overview. *Ann GIS*. 2020;26:1–12 (**Epub 19 December 2019**).
2. Openshaw S. The modifiable areal unit problem. Norwich, UK: Geo Books; 1984.
3. United States Census Bureau. Glossary. Retrieved 2 March 2023 at https://www.census.gov/programs-surveys/geography/about/glossary.html.
4. Jerrett M, Gale S, Kontgis C. Spatial modeling in environmental and public health research. *Int J Environ Res Public Health*. 2010;7:1302–29.
5. Markey C, Bello O, Hanley M, Loehrer AP. The use of area-level socioeconomic indices in evaluating cancer care delivery: a scoping review. *Ann Surg Oncol*. 2023. https://doi.org/10.1245/s10434-023-13099-x.
6. Shi X, Miller S, Mwenda K, Onda A, Reese J, Onega T, et al. Mapping disease at an approximated individual level using aggregate data: a case study of mapping New Hampshire birth defects. *Int J Environ Res Public Health*. 2013;10:4161–74.