

TITLE: Novel Mapping Methods to Describe Utilization of Free Breast Cancer Screening from a State Program

Kelly D. Hughes, PhD¹

David Haynes, PhD²

Anne M. Joseph MD, MPH³

¹ Corresponding Author

Minnesota Department of Health, Sage Program

85 7th Place E

St. Paul, MN 55101

651-201-5619

kelly.d.hughes@state.mn.us

² University of Minnesota

Institute for Health Informatics, Suite 8-100

516 Delaware Street SE

Minneapolis, MN 55455

612-625-7412

dahaynes@umn.edu

^c University of Minnesota

Department of Medicine, Division of General Internal Medicine

420 Delaware St SE

MMC 194

Minneapolis, MN 55455

612-625-3789

amjoseph@umn.edu

ABSTRACT WORD COUNT: 100

TEXT WORD COUNT: 2,954

PAGES: 18

TABLES: 0

FIGURES: 2

ABSTRACT

The National Breast and Cervical Cancer Early Detection Program (NBCCEDP) of Minnesota, “Sage”, provides breast cancer screening to uninsured women. We introduce a novel mapping technique, spatially adaptive filters (SAFs), to estimate utilization of Sage screening in Minnesota. Sage screenings ($N = 74,712$) were geocoded. The eligible population was modeled with the RTI synthetic population dataset. Between 2011 and 2015, 36,979 women a year were Sage eligible. Utilization was highly variable across Minnesota ($M = 37.2\%$, range $0\% - 131\%$, $SD = 18.7\%$). This replicable approach modeled utilization rates to the neighborhood-level, allowing Sage to prioritize locations and engage communities.

KEYWORDS: Geographic Information Systems; Women’s Health; Cancer Screening; Breast Cancer; Health Programs

FIGURES NOTES: Color should be used for figures in print.

INTRODUCTION

Cancer is the second leading cause of mortality among women in the United States, and 30% of new cancer diagnoses in 2020 among women will be breast cancers.^{1,2} Breast cancer mortality can be reduced via early detection through screening, yet disparities in breast cancer screening persist. Nationally, low-income women ($< \$15,000/\text{year}$) ages 50 to 64 are less likely to be screened than women with higher income ($> \$50,000/\text{year}$), 71.9% compared to 83.0%.³ Women without health insurance are screened at a lower rate (54.7%) than women with insurance (80.5%).³ Racial disparities in screening also exist, as Asian and American Indian women are screened at lower rates than other groups; women of color and American Indians face disproportionate barriers to screening, such as having lower rates of access to health care.^{4,5} Recognizing that access to early breast cancer detection is a health priority among low-income and uninsured women, the National Breast and Cervical Cancer Early Detection Program (NBCCEDP) was established in 1991 to provide screening services for such populations.

NBCCEDPs are present in every US state and territory, and in some tribal areas.⁶ To evaluate NBCCEDP delivery of screening services, previous research sought to quantify the eligible population and uptake of services, often at the national level. Among US states, Tangka and colleagues estimated that 4 million women 40 to 64 years of age were eligible for NBCCEDP breast screening services, and 13.2% received screening in 2002-2003.⁶ Howard and colleagues updated this work using 2011-2012 data and estimated a larger eligible population of 5 million women 40 to 64, with an average national screening utilization of 10.6%.⁷ A final study considering years 2006 to 2010 placed the NBCCEDP national breast screening utilization rate at 19.9%.⁸ Of these papers, only one reported state-level screening rates, and although screening rates for all states were reported, these were not labeled and so the specific screening rate for any

state was not identifiable.⁷ To better inform program effectiveness, NBCCEDPs would benefit from spatially generated state and local estimates of the eligible population and service utilization.

Variation in utilization of services within NBCCEDP jurisdictions is currently unknown. Local variation in uptake of health services is critical to addressing disparities, as health and disease are linked to demographic characteristics of a local population and resources in their immediate environment.^{9,10} That is, local mapping reveals the story of health and disease.¹¹ Eligible population uptake of NBCCEDP screening services may vary considerably. Howard and colleagues found that NBCCEDP breast screening utilization ranged from 3.2% to 52.8% among US states in 2011 to 2012 (however, specific states were not labeled).⁷ Similar variation describing the uptake of services for NBCCEDPs might be found below the state level, but to date there are no peer-reviewed accounts of this information. The 500 Cities Project, which creates small area estimates for select US cities from national datasets such as BRFSS, shows a range of 60% to 83.5% for breast cancer screening among women in large US cities, and similar ranges of variation at the neighborhood level.¹² Screening utilization varies within the states, as cancer screening does not reach all of the intended population.^{13,14}

New geospatial techniques are needed to generate the high-resolution yet high-accuracy data needed to evaluate NBCCEDP effectiveness and disparities in program use. We demonstrate the use of spatially adaptive filters (SAFs) and a synthetic population dataset developed by the Research Triangular Institute (RTI) to calculate a breast cancer screening eligible population, and variation in utilization rates at the neighborhood level for the Minnesota NBCCEDP, “Sage”. For Minnesota, the implementation of these methods will improve understanding of the uptake of programmatic services and allow for targeting of specific locations where Sage can deploy limited resources. The challenge is to depict variations in care, including disparities, at the local level to allow healthcare systems and programs to take meaningful action. In a wider scope, this project demonstrates the utility of accurate sub-state estimates of target populations, by documenting obscured local variation in utilization of services.

MATERIALS AND METHODS

Study Area Context

Sage is the NBCCEDP of Minnesota, whose primary activities are program recruitment, client connection to resources through patient navigation, creating a network of resources that provide access to cancer-related healthcare, and paying for breast and cervical screening and diagnostic services. The entire state of Minnesota is Sage’s service catchment area; Sage partners with over 400 clinics across the state, and serves women residing in every county in the state of Minnesota. Previous work estimated a Sage eligible population of 35,000.⁷ Sage screens approximately 15,000 uninsured and underinsured low-income women annually. Minnesota has one of the highest rates of health insurance access in the nation, yet the state experiences persistent disparities in breast cancer screening, incidence, and mortality that reflect national averages, as

well as similar race and ethnicity disparities in income and health care access.^{3,15} This project performs secondary analysis of observational data for the purpose of program evaluation, and IRB review was not required. Sage screening activity from 7/1/2010 to 6/30/2015 (the fiscal year starts July 1st and ends June 30th) was analyzed. The analysis was completed between the years 2017 and 2019.

Numerator

Sage eligibility and utilization rates were estimated based upon Sage's breast cancer screening eligibility guidelines between 2010 to 2015. Sage endorsed annual screening for average risk women, starting at age 40. To qualify for services a woman must have had an income below 250% of the federal poverty level and be uninsured or underinsured, defined as any out-of-pocket costs for screening. Note, that each state NBCCEDP can set its own guidelines for eligibility, patterned after nationally recommended guidelines.¹⁶ For this analysis, Sage screening cases were defined as all instances of women screened for breast cancer through Sage between the dates of analysis, which was 74,712 instances of screening. Sage NBCCEDP federal funding is supplemented by state and non-profit funds and grants that contribute 20% to 25% of Sage's total budget; screening supported by all funding streams was included in the analysis. Additional tests, such as diagnostic mammograms that occurred after the initial screening, were not included as screenings in the analysis. Sage clients that were non-Minnesota residents were removed.

We used the Minnesota Department of Health geocoder to determine geographic coordinates from client addresses. We defined geocoded clients as those with an address match score greater than 60; the lower bound of 60 was acceptable as we retroactively geocoded the data which affects match score quality.¹⁷ The address match score has a range 0 to 100, and is a result of the accuracy of the given address with the reference dataset of roads.¹⁸ In the case of multiple address matches we accepted the address with the highest score. Most screening instances (73%) had a geocode score of 80 or greater.

Some Sage clients did not have sufficient address information for geocoding, yet all clients did supply a zip code. When full address information is missing, it is common practice to geocode individuals to a zip code using zip code tabulated areas (ZCTA).¹⁹ We developed an algorithm in PostgreSQL v 10 that equally distributed all clients within a given ZCTA, if that client did not have a full address; this assigned point within the ZCTA served as the geocoded location of the screened client. PostgreSQL is open source software.

Denominator

The Research Triangular Institute (RTI) 2010 U.S. Synthesized Population dataset was used to calculate our eligible population.²⁰ The synthetic dataset is a statistical model based on American Community Survey Public Use Microdata. It provides geographic coordinates and household characteristics (i.e., household size and income) for every synthetic household. The household

record can be linked to a secondary statistically modeled person dataset. By linking the individual to the household, we created a geographic location for every statistically modeled woman in Minnesota.

We defined the potential eligible client population using synthetic population characteristics of age and income, reflecting Sage eligibility criteria (equation 1). Additionally, age-appropriate American Indian or Alaska Native women screened through Indian Health Service were eligible for Sage regardless of income (equation 2). The income threshold of < 250% federal poverty level was calculated through household income and size¹⁸. For women who were not American Indian or Alaska Native, we applied estimated rates of uninsured and underinsured to the entire population. From published public data, the underinsured rate in Minnesota was estimated to be 10%, and the uninsured rate for Minnesota women ages 40 to 64 with income ≤ 250 FPL was estimated at 12.6%.^{21,22} Together, this gives a proportion of 0.226 women who were uninsured or underinsured, which we applied to all women in equation 1.

Equation 1, income and insurance status based eligibility criteria

$$\text{Eligible Person} \times 0.226 = (\text{Income} - 30350 - (10800 \times \text{Householdsize}) \leq 0) \text{ AND sex} \\ = \text{Women AND Age} \geq 40$$

Equation 2, Indian Health Services eligibility criteria

$$\text{Eligible Person} = (\text{sex} = \text{Women AND Age} \\ \geq 40 \text{ AND race IN (American Indian, Alaska Native)})$$

These two equations are used together to define all individuals whose characteristics meet the criteria for eligibility in the RTI dataset. This population of women was multiplied by five to generate a denominator for five years' of Sage screening data, as women were considered eligible for screening every year.

Geospatial Model

Sage breast cancer screening utilization rates were generated using the numerator and denominator populations in conjunction with spatially adaptive filters (SAFs), an advanced geospatial technique that creates stable rates by controlling for changes in population density.²³ The SAFs approach was selected for several reasons. SAFs do not rely on pre-defined administrative boundaries (i.e., zip codes, county), and are robust to issues that arise from

working with small numbers, such as unstable estimates, suppression of data, and the Modifiable Areal Unit Problem.²⁴

When applying SAFs, a dense grid is first “laid” over the area to be mapped (i.e., all of Minnesota); in our model we selected a grid of 9,522 points that were 5,000 meters apart. Next, centered at each grid point, the circle is expanded until it captures 500 individuals in the denominator (i.e., the Sage eligible population). When the circle’s area has been determined, a point utilization rate at the centered grid point is calculated; the numerator consists of all the Sage screened geocoded individuals within that same circle. The method for computing the areas of each circle is complex and necessitated the big data platform Apache Spark, using the geospatial library GeoSpark; Apache Spark is open source software.²⁵

Utilization map

After point estimates were determined for the entire grid, an inverse distance weighted interpolation was applied across the Minnesota grid to calculate utilization rates between points, increasing the resolution of our map. The interpolation results in a smoothed surface with 873,983 cells, with a spatial resolution of 500m². The interpolation and final plotting of the map was accomplished with ArcPro v 2.4.

RESULTS

All Sage clients screened for breast cancer between fiscal years 2011 and 2015 were used in the mapping process; 90% were geocoded to physical addresses, and the remaining 10% were assigned a location within a ZCTA using our algorithm. Sage clients self-identified as white (53%), Hispanic (22%), African American (9%), American Indian (7%), and Asian (3%). Six percent of clients did not provide racial identification information or were multi-racial. Almost half of all Sage clients, 48.9%, lived within the seven-county metro area of Minneapolis and St. Paul.

Our approach estimates that 36,979 Minnesota women 40 years and older were eligible for Sage breast cancer screening services per year between fiscal years 2011 and 2015, on average. This represents 3% of the Minnesota population of women 40 and older in 2010. The average breast screening utilization for Sage services, which assumes annual screening, was 37.2%. Utilization rates varied locally. The range of point estimates of utilization, for the spatially adaptive filters was 0% to 131%, with a standard deviation of 18.7%. Interpolated areas > 100% utilization accounted for 0.8% of all areas.

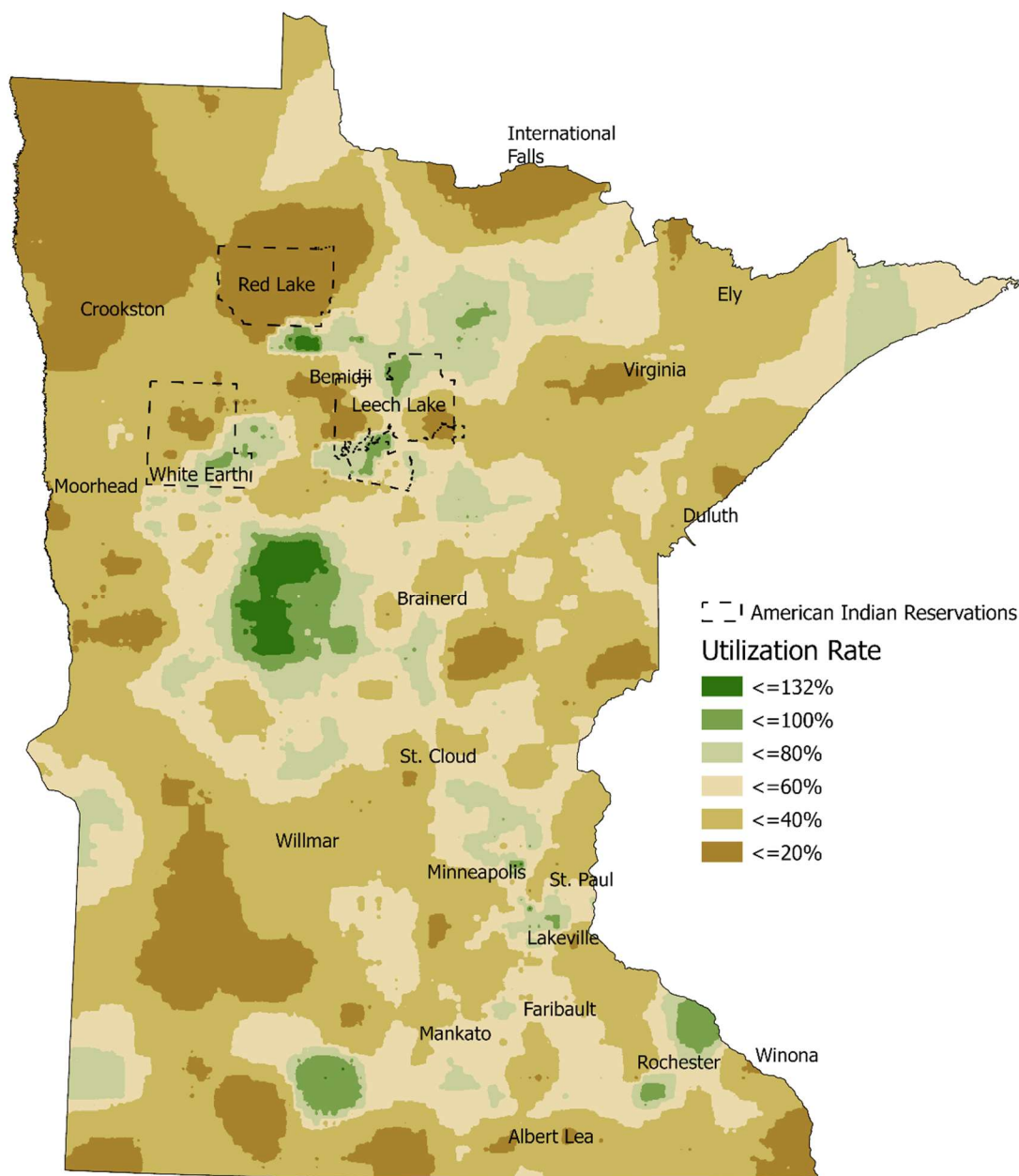


Figure 1. Sage Utilization Rates in Minnesota.

Visual inspection of the map shows clusters of high and low breast cancer screening across the state (Figure 1). Especially high utilization is found in central Minnesota west of Brainerd, in clusters surrounding Bemidji, and in pockets west of Mankato and adjacent to Rochester. Notably, the clusters of high utilization near Bemidji correspond spatially with the Red Lake, White Earth, and Leech Lake tribal areas. Rochester, MN is the home of the Mayo Clinic, a long standing Sage partner and world class medical clinic.

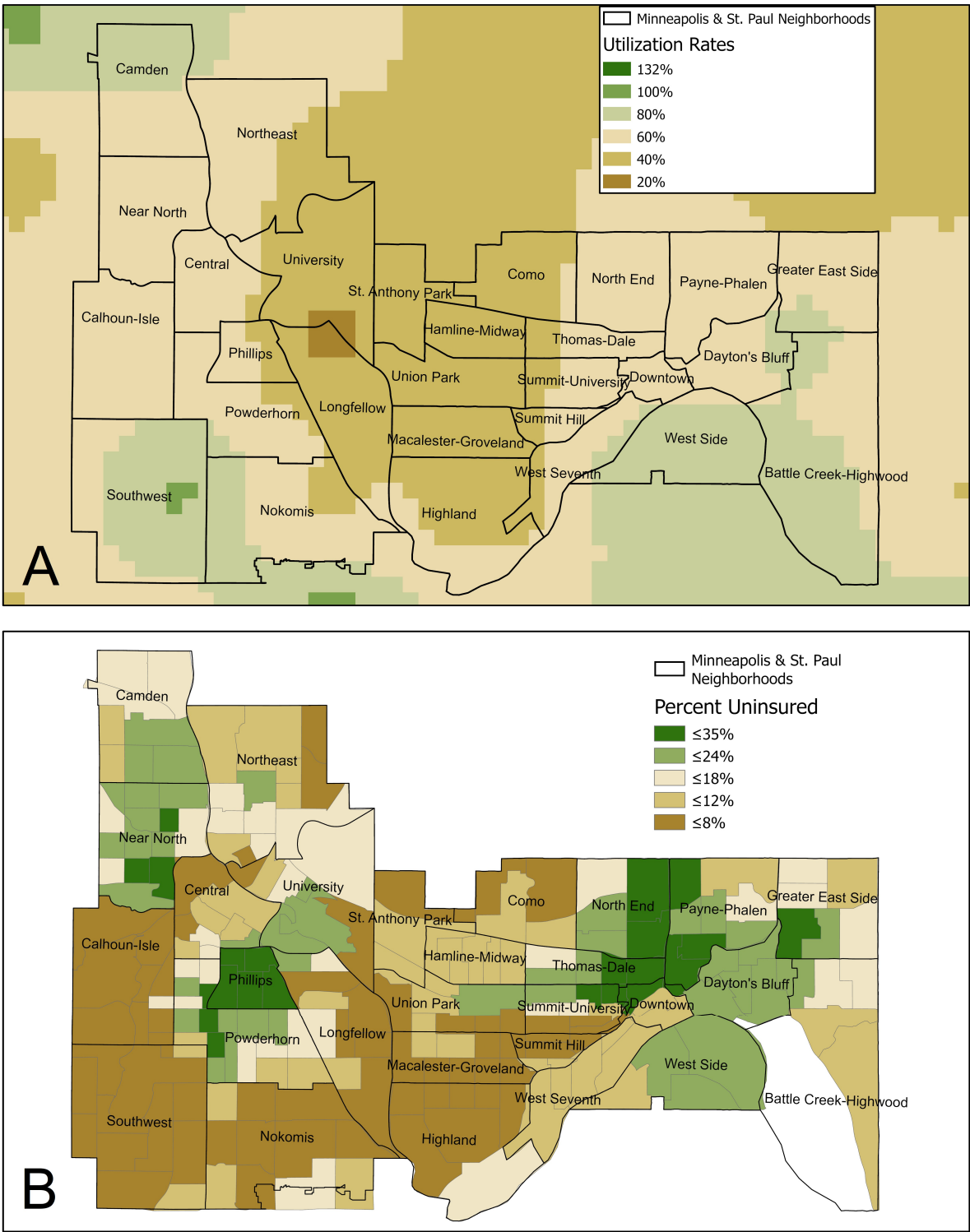


Figure 2. A: Sage utilization rates in Minnesota’s Minneapolis and St. Paul neighborhoods, B: 500 Cities dataset on rates of lack of health insurance in Minneapolis and St. Paul neighborhoods, derived from BRFSS data 2012 - 2016.¹² Sage programs can zoom into Minnesota locales and examine breast cancer screening utilization alongside relevant

demographic and health data that provide context for Sage's programmatic efforts, such as rate of lack of health insurance.

Examination of local rates in Minneapolis and St. Paul neighborhoods demonstrate the high-resolutions possible with SAFs (Figure 2A). Each pixel on the map represents 500 m². The map reveals low utilization in a corridor between the two city centers, and few places where utilization rates are particularly high. Pixels with very high utilization scores in southwest Minneapolis correspond with high income areas with a very small Sage eligible population.

CONCLUSIONS

Main findings

Our work adds to the literature through introduction of powerful geospatial methods that quantify the number of individuals eligible for Sage breast cancer screening services and utilization of those services at the local level. We were capable of generating these estimates despite challenges associated with small numbers and a need for local, high-resolution yet reliable estimates. These methods provided the first published state-level and local estimates for NBCCEDP breast screening rates.

Our estimate of statewide utilization of breast cancer screening services (37.2%) is well within the previously reported range of NBCCEDP breast screening rates at the state level (3.2% to 52.8%), yet is higher than previous national estimates of NBCCEDP breast cancer screening utilization that ranged from 10% to 20% for the same time period.⁶⁻⁸ Our estimation of Sage's breast cancer screening is likely higher than previous estimates as we included screenings funding by all revenue streams, not just federal funding, as was the case in previous publications.⁶⁻⁸ Our estimated denominator population is similar to previously published denominators; we estimated that 36,979 women 40 and older were eligible for Sage breast screening, while Howard and colleagues estimated that 35,000 women 40 to 64 were eligible for Sage.⁷ Sage programs screened more of the eligible population compared to the median NBCCEDP program, yet all estimates are relatively low and Sage could potentially screen many more women.

Local variation in utilization of Sage breast cancer screening services varied within Minnesota, 0% to 100%. Estimates of utilization at the national or state level obscured significant local variation. Local variation in utilization was revealed by SAF methods, and would have remained unknown using common mapping methods that rely on pre-defined administrative units (e.g., ZCTAs). Pre-defined administrative areas are often used in analyses because they are easily accessible. However, a major limitation is that they vary in size and include heterogeneous populations, which can obfuscate health patterns and socially-relevant boundaries within the unit.²⁶ It is important to reveal this variation to document and address both geographic and racial or ethnic disparities in utilization of Sage program resources.

While the utilization rates do range above 100%, they represent a small percentage of overall map (0.8% of pixels). These areas reflect spatial filters with 500 or more instances of screening within five years. There are two main explanations for this outcome. First, the breast cancer screening utilization rate is analogous to a standardized incidence ratio, which standardizes for population density. While ratio values generally range from zero to one (i.e., 100%), they can exceed 1 when more cases are observed than expected. Second, some of our model assumptions may underestimate the denominator. It is likely that the RTI Synthesized Dataset, which is based on the 2010 US Census, underestimates the presence of individuals that are undercounted in the census, such as Blacks/African Americans, Hispanics, and American Indians.²⁷ In particular, Minnesota is home to a large American Indian population who live on sovereign land and receive screening through American Indian services. The near 100% utilization of Sage breast cancer screening within sovereign American Indian populations can be attributed to both high utilization and some inflation of utilization via denominator underestimation. We also used a constant uninsured and underinsured rate across the entire population, which underestimates uninsured rates for certain groups. For example, in 2014 the uninsured rate was 26% among Hispanics compared to 12.6% we used in our model.²² Since we are applying a local spatial model, if the denominator is underestimated due to the RTI dataset or insurance rate assumptions, the utilization rate will be overestimated, resulting in values above >100%.

This project was enabled by a government-academic partnership. A funded post-doctoral scholar collaborated with state public health to provide mapping expertise for health outcomes. The materials for this project, except for ArcGIS, were open source and free. Code that was used to perform these analyses were shared with the state and available on Github. Therefore, replication of this project is possible, in regards to finance and resources, for other governmental institutions if pursued through collaborative partnerships.

Limitations

A limitation of this research is the accuracy of geocoding of some of the Sage participants. Self-reported addresses can lead to address inaccuracies or inconsistencies that hamper geocoding and rural addresses are known to be more difficult to geocode.²⁸ Our model could be made richer with more detailed uninsured and underinsured information, such as age, race, or income specific uninsured rates. This is an area of future work. Finally, we limited the analysis of Sage data to 2011 to 2015 to match the available denominator RTI dataset, based on the 2010 US Census. The Minnesota Sage eligible population may have changed since that time, since implementation of the Medicaid expansion under the Affordable Care Act in 2014.²⁹⁻³¹

Public Health Implications

This research demonstrates how spatial approaches, and in particular SAFs, can provide insight into the programmatic uptake of services, in this case utilization rates of breast cancer screening services of a state NBCCEDP. Service uptake can be examined to the neighborhood level to identify areas of great success and neglect. Maps of programmatic uptake can be inspected in the

context of maps of relevant contextual variables such as uninsured rates (Figure 2B), race and ethnicity, or income, providing informative clues as to correlates of service uptake and how to structure successful interventions to increase utilization. Statistical analysis of correlates is possible, but beyond the scope of this publication. Finally, local utilization data enhance the capacity of health programs to engage communities, by providing data that are community-specific and relevant.³² NBCCEDPs are present in every state and territory in the US and our model can be applied to other NBCCEDPs or similar cancer screening programs, allowing decision-makers the ability to understand screening service utilization at a level of detail previously unavailable.

ACKNOWLEDGMENTS

We would like to thank Christina Nelson for championing this project within the Sage Program. We thank Tossy Kelly and Cheemeng Vang for IT help for geocoding Sage clients, and for maintaining and updating the relational database that housed all data. David Haynes's time was supported by NIH grant 5T32CA163184 while working on this project. This project did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

DECLARATION OF INTERESTS

The authors have no competing interests or conflicts of interest to declare.

Kelly D. Hughes is employed by the State of Minnesota and the University of Minnesota. Her effort on this project was supported by the State of Minnesota. Materials for this project were supplied by the State of Minnesota. Representatives of the Sage Program of the State of Minnesota were key stakeholders study design and data collection for the project. The State of Minnesota has approved this manuscript.

David Haynes' time was supported by NIH grant 5T32CA163184 while working on this project and while employed at the University of Minnesota. David Haynes is an unpaid contractor in relation to the State of Minnesota. David Haynes' sponsors had no input into the project.

Anne M. Joseph is employed by the University of Minnesota. She is supported in part by funding from the National Cancer Institute.

FINANCIAL DISCLOSURES: Kelly D. Hughes has no financial disclosures.

David Haynes has no financial disclosures.

Anne M. Joseph has no financial disclosures.

REFERENCES

1. Hahn RA, Chang M-H, Parrish RG, Teutsch SM, Jones WK. Trends in Mortality Among Females in the United States, 1900–2010: Progress and Challenges. *Prev Chronic Dis*. 2018;15. doi:10.5888/pcd15.170284
2. American Cancer Society. Cancer Facts & Figures 2020. Published 2018. Accessed January 5, 2020. <https://www.cancer.org/content/dam/cancer-org/research/cancer-facts-and-statistics/annual-cancer-facts-and-figures/2020/cancer-facts-and-figures-2020.pdf>
3. Centers for Disease Control and Prevention. BRFSS Web Enabled Analysis Tool, Cross Tabulation. Published 2018. Accessed January 7, 2020. <https://nccd.cdc.gov/weat/#/analysis>.
4. Artiga S, Foutz J, Cornachione E, Garfield R. Key facts on health and health care by race and ethnicity. *Kais Fam Found*. Published online 2016.
5. Kaiser Family Foundation. State Health Facts. State Health Facts. Published 2018. Accessed August 16, 2018. <https://www.kff.org/statedata/>
6. Tangka FK, Dalaker J, Chattopadhyay SK, et al. Meeting the mammography screening needs of underserved women: the performance of the National Breast and Cervical Cancer Early Detection Program in 2002–2003 (United States). *Cancer Causes Control*. 2006;17(9):1145–1154.
7. Howard DH, Tangka FK, Royalty J, et al. Breast cancer screening of underserved women in the USA: results from the National Breast and Cervical Cancer Early Detection Program, 1998–2012. *Cancer Causes Control*. 2015;26(5):657–668.
8. Subramanian S, Tangka FKL, Ekwueme DU, Trogon J, Crouse W, Royalty J. Explaining variation across grantees in breast and cervical cancer screening proportions in the NBCCEDP. *Cancer Causes Control*. 2015;26(5):689–695. doi:10.1007/s10552-015-0569-5
9. Gibson A, Asthana S, Brigham P, Moon G, Dicker J. Geographies of need and the new NHS: methodological issues in the definition and measurement of the health needs of local populations. *Health Place*. 2002;8(1):47–60. doi:10.1016/S1353-8292(01)00035-1
10. Dummer TJB. Health geography: supporting public health policy and planning. *CMAJ Can Med Assoc J*. 2008;178(9):1177–1180. doi:10.1503/cmaj.071783
11. Koch T. *Cartographies of Disease: Maps, Mapping, and Medicine*. Esri Press Redlands, CA; 2005.
12. Centers for Disease Control and Prevention, National Center for Chronic Disease Prevention and Health Promotion, Division of Population Health. 500 Cities Project Data [online]. Accessed August 18, 2020. <https://www.cdc.gov/500cities>
13. Earp JAL, Altpeter M, Mayne L, Viadro CI, O'Malley MS. The North Carolina Breast Cancer Screening Program: foundations and design of a model for reaching older, minority, rural women. *Breast Cancer Res Treat*. 1995;35(1):7–22. doi:10.1007/BF00694740
14. Coughlin SS, Leadbetter S, Richards T, Sabatino SA. Contextual analysis of breast and cervical cancer screening and factors associated with health care access among United States women, 2002. *Soc Sci Med*. 2008;66(2):260–275. doi:10.1016/j.socscimed.2007.09.009
15. American Cancer Society. *Breast Cancer Facts & Figures 2017-2018*. American Cancer Society, Inc; 2017.
16. Find a Screening Program Near You | NBCCEDP | CDC. Published January 24, 2020. Accessed September 2, 2020. <https://www.cdc.gov/cancer/nbccedp/screenings.htm>

17. Goldberg DW, Ballard M, Boyd JH, et al. An evaluation framework for comparing geocoding systems. *Int J Health Geogr*. 2013;12(1):50. doi:10.1186/1476-072X-12-50
18. Goldberg DW. Improving Geocoding Match Rates with Spatially-Varying Block Metrics. *Trans GIS*. 2011;15(6):829-850. doi:10.1111/j.1467-9671.2011.01295.x
19. Rushton G, Armstrong MP, Gittler J, et al. Geocoding in Cancer Research. *Am J Prev Med*. 2006;30(2):S16-S24. doi:10.1016/j.amepre.2005.09.011
20. Wheaton WD, Cajka JC, Chasteen BM, et al. Synthesized population databases: A US geospatial database for agent-based models. *Methods Rep RTI Press*. 2009;2009(10):905.
21. Schoen C, Doty MM, Robertson RH, Collins SR. Affordable Care Act Reforms Could Reduce The Number Of Underinsured US Adults By 70 Percent. *Health Aff (Millwood)*. 2011;30(9):1762-1771. doi:10.1377/hlthaff.2011.0335
22. United States Census Bureau. SAHIE. Accessed February 26, 2020. <https://www.census.gov/data-tools/demo/sahie/#/>
23. Tiwari C, Rushton G. Using spatially adaptive filters to map late stage colorectal cancer incidence in Iowa. In: *Developments in Spatial Data Handling*. Springer; 2005:665–676.
24. Fotheringham AS, Wong DWS. The Modifiable Areal Unit Problem in Multivariate Statistical Analysis. *Environ Plan A*. 1991;23(7):1025-1044. doi:10.1068/a231025
25. Yu J, Zhang Z, Sarwat M. Spatial data management in apache spark: The geospark perspective and beyond. *Geoinformatica*. 2019;23(1):37–78.
26. Galea S, ed. *Macrosocial Determinants of Population Health*. Springer; 2007.
27. O'Hare WP. Who Lives in Hard-to-Count Neighborhoods. *Int J Soc Sci Stud*. 2016;4(4):43-57. doi:10.11114/ijsss.v4i4.1441
28. Ward MH, Nuckols JR, Giglierano J, et al. Positional accuracy of two methods of geocoding. *Epidemiology*. Published online 2005:542–547.
29. Kaiser Family Foundation. Medicaid in Minnesota. Published online October 2019. Accessed February 26, 2020. <http://files.kff.org/attachment/fact-sheet-medicare-state-MN>
30. Levy AR, Bruen BK, Ku L. Health Care Reform and Women's Insurance Coverage for Breast and Cervical Cancer Screening. *Prev Chronic Dis*. 2012;9. doi:10.5888/pcd9.120069
31. Tangka F, Kenny K, Miller J, Howard DH. The eligibility and reach of the national breast and cervical cancer early detection program after implementation of the affordable care act. *Cancer Causes Control*. 2020;31(5):473-489. doi:10.1007/s10552-020-01286-0
32. Dubowitz T, Williams M, Steiner ED, et al. Using geographic information systems to match local health needs with public health services and programs. *Am J Public Health*. 2011;101(9):1664–1665.