HHS Public Access

Author manuscript

Spat Spatiotemporal Epidemiol. Author manuscript; available in PMC 2024 June 01.

Published in final edited form as:

Spat Spatiotemporal Epidemiol. 2023 June ; 45: 100564. doi:10.1016/j.sste.2023.100564.

Feasibility of visualizing cancer incidence data at sub-county level: Findings from 21 National Program of Cancer Registries☆

Taylor D. Ellington[a,b], **Angela K. Werner**[c,*], **S. Jane Henley**[b], **Lisa E. Paddock**[d], **Pamela K. Agovino**[e]

[a]Oak Ridge Institute for Science and Education, Oak Ridge, TN, United States

[b]Division of Cancer Prevention and Control, National Center for Chronic Disease Prevention Health Promotion, Centers for Disease Control and Prevention, Atlanta, GA, United States

[c]Division of Environmental Health Science and Practice, National Center for Environmental Health, Centers for Disease Control and Prevention, Atlanta, GA, United States

[d]New Jersey State Cancer Registry, Rutgers Cancer Institute of New Jersey, New Brunswick, NJ, United States

[e]Cancer Surveillance Research Program, Cancer Epidemiology Services, New Jersey Department of Health, Trenton, NJ, United States

Abstract

Monitoring cancer incidence data by geography is useful for planning public health activities. However, due to anticipated confidentiality and statistical reliability issues, data on cancer incidence and mortality are more often displayed at a national, state, or county level, rather than at more local levels. To address this gap in displaying cancer data at the local level, the CDC's National Environmental Public Health Tracking Program and 21 National Program of Cancer Registries worked together on a pilot project to examine the feasibility of displaying sub-county-level incidence of selected cancer types diagnosed during 2007–2016. The results from this project are important steps for building sub-county cancer displays into data visualizations and using the data in a way that provides meaningful insights. The availability of sub-county cancer data may allow researchers to better examine cancer data at a local level which may help guide public health decisions regarding community-based interventions and screening services.

☆Participants of Phase One of the Sub-County Cancer Pilot Project (Centers for Disease Control and Prevention 2020)

*Corresponding author at: 4770 Buford Highway NE, MS F60, Atlanta, GA 30341. awerner@cdc.gov (A.K. Werner).

Declaration of Competing Interest

The authors declare no potential conflicts of interest.

Disclaimer

The findings and conclusions in this report are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention.

CRediT authorship contribution statement

Taylor D. Ellington: Conceptualization, Methodology, Formal analysis, Writing – original draft, Writing – review & editing. **Angela K. Werner:** Conceptualization, Methodology, Formal analysis, Writing – review & editing. **S. Jane Henley:** Conceptualization, Methodology, Formal analysis, Writing – review & editing. **Lisa E. Paddock:** Formal analysis, Writing – review & editing. **Pamela K. Agovino:** Formal analysis, Writing – review & editing.

Keywords

Sub county; Environmental health; Cancer surveillance; Tracking; Cancer registries

1. Introduction

Data on cancer incidence are commonly displayed at the national, state, or county level, but not at more local levels due to anticipated confidentiality and statistical reliability issues. (Centers for Disease Control and Prevention 2020) In recent years, there has been a push to develop timely, locally-relevant health information systems. (DeSalvo et al., 2016) Despite new developments in technology, public health data remain spatially unresolved, lagging, or all together unavailable. (Werner and Strosnider, 2020) Innovation is uniquely challenging in public health, as problems tend to be complex, dynamic, and context specific. (Berney et al., 2015) Some states have used environmental public health tracking to enhance their public portal to benefit programs by mapping data and developing smartphone-friendly versions of the their state's Web portal. (Jordan et al., 2015; Abookire et al., 2020)

While the benefit of using finer geographic detail to analyze cancer registry data has been discussed previously, more precise geocodes (e.g., to the census tract level) in cancer data have only been routinely used more recently. (Rushton et al., 2006; Howe, 1986) Often geographic data for cancer cases are aggregated to the county or state level before release in public-facing databases. (Yu et al., 2017) Finer resolution data can allow for health issues affecting specific populations to be examined and tailored interventions to be implemented. (Centers for Disease Control and Prevention 2013) The use of sub-county data can also increase public awareness of place-based factors and understanding how these relate to health. (Werner et al., 2018)

To address this gap in displaying cancer data at the local level, two signature programs funded by Centers for Disease Control and Prevention (CDC) worked together by testing aggregation schemes for common and rare cancer types. The National Environmental Public Health Tracking Program (Tracking Program) was established in 2002 to track exposures and health outcomes including selected cancers associated with environmental hazards and to bridge existing data gaps. (McGeehin et al., 2004) State and local Tracking Programs develop or enhance standards as part of the Nationally Consistent Data and Measures (NCDMs) for data calculation, dissemination, and display. A pilot project in 2014 suggested that NCDMs, as well as standardized geographies, be established for generating, analyzing, and sharing sub-county data to allow for comparison of different health outcomes over space and time. (Werner et al., 2018) Collaboration between the Tracking Program and jurisdictions who work with the Tracking Program is key to advancing efforts and making sub-county data available on a larger scale. (McGeehin et al., 2004)

The National Program of Cancer Registries (NPCR) collects data on cancer occurrence, the type of initial treatment, and outcomes for 46 states, DC and 3 territorial cancer registries representing 97% of the United States population. NPCR measures progress in preventing and treating cancer in the United States. (National Program of Cancer Registries 2020) CDC

displays cancer data at the national, state, county, and Congressional District levels. (Centers for Disease Control and Prevention 2020)

In this project, the Tracking Program and 21 NPCR registries examined the feasibility of displaying cancer incidence at sub-county levels. Using data for selected cancer types diagnosed during 2007–2016, stability and confidentiality issues were examined to identify which spatiotemporal aggregation minimized suppression and instability while allowing for display of finer spatial resolution data at each sub-county spatial resolution. The participants tested different spatial and temporal aggregations for displaying lung cancer, breast cancer (females only), prostate cancer, colorectal cancer, liver cancer, melanoma of the skin, and non-Hodgkin lymphoma. The aim of this project was to suggest spatiotemporal aggregations for sub-county cancer data display and dissemination.

2. Methods

2.1. Background

To increase the availability and accessibility of sub-county data, the Tracking Program created standardized, population-based sub-county geographies using census tracts as the foundation, ensuring that the maximum number of sub-county geographies could be displayed with minimal suppression and instability. (Werner and Strosnider, 2020) Briefly, in 2014, using the Geographic Aggregation Tool and user-specified criteria (such as contiguous boundaries) to aggregate census tracts into geographies, the Tracking Program tested various population thresholds to determine optimal aggregations across health outcomes, over time, and between places. In the full context of all outcomes monitored in the Tracking Program, the optimal aggregation for common outcomes was determined to have a minimum population of 5000 persons and for rarer outcomes, a minimum population of 20,000 persons. (Werner et al., 2018) Standardized spatially aggregated geographies were created for all states and are used in this project (Werner et al., 2018) The Tracking Program also determined that when data are sparse, even with spatial aggregation, temporal aggregation across a set number of years could be used.

For this pilot project, a subset of 21 NPCR recipients volunteered to participate to address the gap in displaying cancer data at the local level. Twelve of these states also had a local Tracking Program that participated in the project. To promote active participation, states were grouped into three regional workgroups. Each regional workgroup examined a different set of cancer sites. Region 1 consisted of NPCR cancer registries and Tracking Programs from Florida, Georgia, New Jersey, North Carolina, Puerto Rico, Rhode Island, South Carolina, and Virginia. Participants in this region examined colorectal cancer and non-Hodgkin lymphoma. Region 2 consisted of Louisiana, Michigan, Minnesota, Missouri, Nebraska, and Wisconsin. Participants in this region examined female breast cancer and melanoma. Region 3 consisted of Arizona, California, Idaho, North Dakota, Texas, Utah, and Washington. Participants in this region examined prostate cancer and liver and intrahepatic bile duct cancer. All regions examined lung cancer and lung cancer stratified by sex.

2.2. Population and cancer incidence data

Population data were acquired from the U.S. Census Bureau 2010 Decennial Census for population by age and sex, categorized by census tract. (U.S. Census Bureau 2017; U.S. Census Bureau 2011) Data for invasive cancers diagnosed during 2007–2016 by cancer type, age, sex, and census tract were provided by cancer registries or Tracking Programs.

2.3. Statistical calculations

SAS 9.4 was used for statistical calculations. (SAS Institute Inc., 2013) For the pilot, crude rates were calculated as the number of cases divided by the population. For confidentiality and reliability, counts and rates were suppressed when there were <16 cases, the standard suppression criteria for cancer data, or <100 persons in the denominator. The pilot participants also tested suppression for <10 cases. For this project, any geographic units with zero cases were flagged as unstable (with an undefined relative standard error) and factored into the percent unstable calculation; although it is possible that there actually were no cases reported, it is also possible that a case was not correctly geo-coded to that geographic unit or that a case occurred but was not reported.

95% confidence intervals (CI) for rates, relative standard error, percent of geographies categorized as suppressed, and percent of geographies categorized as statistically unstable (i.e., relative standard error ≥30%) were calculated for each geographic unit in census tracts and in each aggregation level (i.e., minimum aggregated population of 5000 persons and minimum aggregated population of 20,000 persons).

2.4. Calculating expected spatial and temporal aggregations

States were provided a crosswalk file to aggregate case counts from census tracts to the standardized spatially aggregated geographies. States were also provided a shapefile for developing maps. Each state analyzed their own data and provided summary tables and maps to be shared within their workgroup. First, states calculated the observed median census tract-level case counts across combinations of the 3 spatial aggregations (census tract, 5000-person minimum geography, and 20,000-person minimum geography) and selected temporal aggregations (annually and 3-, 5-, 7-, and 10-year periods). Zero population census tracts (not zero case census tracts) were not included in the calculations to avoid inaccurate or misleading rate estimates in instances where a case was geo-coded to a census tract with zero population. Methods for calculating spatial and temporal aggregations were previously analyzed by Werner and Strosnider (2020). (Werner and Strosnider, 2020)

Next, states determined the spatiotemporal aggregation that best minimized suppression and instability while allowing for display of finer spatial resolution data based on the median case count. States compared their median census tract-level case counts to the median counts suggested for the standardized geographies previously established as a baseline when looking at annual sub-county data (Table 1). (Werner and Strosnider, 2020) Census tract was suggested for greater than 17 cases (very common outcome), the minimum aggregated population of 5000 persons was suggested for 7.3 to 16.9 cases (common outcome), and the minimum aggregated population of 20,000 persons was suggested for 1.9 to 7.2 cases (rare outcome). (Werner et al., 2018)

Temporal aggregation was suggested for less than 1.9 cases. If temporal aggregation was needed, states examined a 3-year aggregated temporal period first to obtain more than 1.9 cases. If that was not sufficient, then they moved to a 5-year aggregated temporal period or more, with a maximum of a 10-year aggregated temporal period.

Each workgroup reviewed cancer outcomes to see if the percentage of geographies that were suppressed or unstable were acceptable for each of the spatiotemporal aggregation schemes. Participants agreed that acceptable schemes resulted in less than 30% of geographies suppressed and less than 30% of geographies with statistically unstable rates. Then, each workgroup came to a consensus about which standard spatiotemporal aggregations to display for their set of cancer types. States provided descriptive statistics tables and maps using these spatiotemporal aggregations. When there were discrepancies in recommendations between workgroups, the less conservative recommendation was chosen. As a result, states with more conservative recommendations were willing to display data at this level, understanding that more than 30% of their data may be suppressed. Finally, participants discussed NCDM recommendations to allow for multiple display options, including the overall spatiotemporal recommendations for each cancer type.

3. Results

Average median, minimum, and maximum case counts by cancer type for each aggregation level were compiled for all regions (Table 2). The median annual case count was not large enough to display any cancer type at the census tract level without any temporal aggregation. States decided on the below suggested spatiotemporal aggregations by discussing observed output during regional meetings. The best outcome for all states were selected and agreed upon during the final meetings. When deciding on the final display recommendations, states considered percent of geographies that were suppressed, percent of geographies with unstable rates, as well as confidentiality (suppressing case counts less than 16).

Suggested spatiotemporal aggregations by cancer type are displayed in Table 3. The average annual number of cases and incidence rate for lung cancer (overall and stratified by sex) at the 5000 person aggregation for a 5-year period and at the 20,000 person aggregation for a 5-year period were found feasible to display. Because of regional differences in lung cancer incidence, the recommendations for spatiotemporal aggregations varied by workgroup.

States determined that it was feasible to display the average annual number of cases and incidence rate for female breast cancer at the census tract level for a 10-year period, at the 5000 person aggregation for a 5-year period, and at the 20,000 person aggregation for a 3-year period (Table 3). The average annual number of cases and incidence rate for non-Hodgkin lymphoma was found feasible to display at the 20,000 person aggregation for a 5-year period. The average annual number of cases and incidence rate for colorectal cancer was found feasible to display at the 5000 person aggregation for a 5-year period and at the 20,000 person aggregation for a 3-year period. The average annual number of cases and incidence rate for melanoma of the skin was found feasible to display at the 5000 person aggregation for a 5-year period and at the 20,000 person aggregation for a 3-year period. The average annual number of cases and incidence rate for prostate cancer was feasible to

display at the census tract level for a 10-year period, at the 5000 person aggregation for a 5-year period, and at the 20,000 person aggregation for a 3-year period. It was not feasible to display liver and intrahepatic bile duct cancer at census tract or any other aggregation level, even using a 10-year period because of the small number of cases.

An example of maps produced by each pilot state are shown in Figures. Fig. 1 displays a Minnesota state map for lung cancer incidence rates stratified by sex at the 20,000 person aggregation for a 3-year period (2014–2016). Fig. 2 displays a North Carolina map for colorectal cancer incidence rates, using 5000 person aggregation over a 5-year period (2012–2016). Fig. 3 displays Missouri maps for breast cancer incidence rates, using census tract aggregation over a 10-year period (2007–2016).

4. Discussion

The CDC Tracking Program and NPCR worked together on this pilot project to gain insight into displaying cancer data at the sub-county level by testing aggregation schemes for common and rare cancer types. We found suggested spatiotemporal aggregation combinations varied by each cancer type examined. The results from this project are important steps for building sub-county cancer displays into data visualization portals and using the data in a way that provides meaningful analyses leading to public health actions. Easy-to-use public portals widen the potential audience for the results of a spatial analysis of health data to go beyond scientists to include the public, policymakers, the media, and a host of others. (Bell et al., 2006)

Stability, reliability, and confidentiality were identified as issues when determining the best spatial and temporal aggregations to use. In 2017, the national incidence of cancer types examined in this project varied, with female breast (125.1 per 100,000 female persons), prostate (106.5 per 100,000 male persons), lung (55.2 per 100,000 persons), and colorectal (36.8 per 100,000 persons) cancers being more common than melanoma of the skin (22.7 per 100,000 persons), non-Hodgkin lymphoma (18.5 per 100,000 persons), and liver cancer (8.3 per 100,000 persons). (U.S. Cancer Statistics Working Group 2020) For rarer cancer types, obtaining sufficient numerators was difficult, due to having fewer cases even when aggregating, when using small areas. [11] In determining which suggested spatial temporal aggregation combinations worked best for each cancer type, it was important to balance data stability and confidentiality issues while keeping geographic areas at the highest possible resolution and time periods as recent as possible. (Werner et al., 2018)

In this project, census tracts were the smallest area units used and we found that this spatial resolution required greater temporal aggregation. It is suggested that analyses use smallest area units with available data, and aggregation to larger units should be avoided unless there is appropriate rationale. (Pfeiffer et al., 2008) To allow for data users to have flexibility (i.e., finer spatial resolution with more temporal aggregation or coarser spatial resolution with less temporal aggregation), various sub-county aggregations were tested and included in the final recommendations.

Previous studies have examined the geographic variation in cancer incidence, mortality, treatment, and survival by mapping cancer data using geocodes and Geographic Information Systems (GIS). (Rushton et al.,; Sahar et al., 2019) One study found combining neighborhood measures and liver cancer rates can narrow down census tracts to target for liver cancer prevention more than standard approaches that use just demographic data (race/ethnicity and age) from the U.S. Census. (Lynch et al., 2020) Complete geocoded data on census tracts from all registries is desirable for national-level research related to spatial analyses. (Lynch et al., 2020) For diagnosis years 2001–2005 combined, valid information on census tract was submitted for 59% of the records in the NPCR dataset. (Singh et al., 2009) Further analysis of residential history may be helpful to access how residential history may impact the ability to get an accurate geocoded addresses.

The Tracking Program and NPCR is working on a second phase of this pilot project to further refine the work presented in this pilot project. This will include further discussion on counties not aggregated in this pilot project due to small populations, creating a 50,000 person minimum aggregation level to test rarer outcomes, including age-adjusted rates in analyses, creating appropriate cancer communication and messaging for the Tracking Network, including additional cancer types in analyses, building capacity for GIS and refining Tracking Program geocoding standards, and mapping additional interventions such as exposures and health outcomes and stratification by sex.

Findings of this study are subject to at least three limitations. One, age-adjusted rates were not tested in this phase of the pilot project. Second, a limited number of states (21) participated in phase one of the pilot project. Although we believe states that participated in the pilot project provided a good representation and diversity of NPCR registries, they are only a sample of the population. Additionally, sub-county areas are standardized within the Tracking Network but may not work for all purposes. The standardized sub-county areas are useful for providing nationally consistent data and measures to allow us to compare over time, space, and datasets, but some states may not be able to access or display sub-county data due to legislative reasons or have state-specific needs and may want or need to incorporate their own sub-county geographies. For example, areas with a minimum of 50,000 persons have been developed for 14 states for the purpose of disseminating cancer surveillance data. (Tatalovich et al., 2022)

The methods and results of this pilot project demonstrate how spatial and temporal aggregation can be used for cancer data at the sub-county level, which can be used for surveillance purposes beyond the Tracking Program in the United States to achieve stable estimates whilst considering confidentiality issues. (McGeehin et al., 2004) The availability of sub-county cancer data through the Tracking Network may allow public health researchers, policymakers, and the public to better examine cancer data at a local level, which may help guide decisions regarding community-based interventions and screening services. Preliminary data can be explored on CDC's National Environmental Public Health Network at (https://ephtracking.cdc.gov/DataExplorer/). Under the select data section, individuals can select the cancer type they would like to review and the corresponding measures (age-adjusted rates and number of cases at the national and state level.

Acknowledgements

We gratefully acknowledge the contributions of the state and regional cancer registry staff, National Environmental Public Health Tracking staff, and health department personnel for their work in collecting and analyzing the data used in this study. We thank D. Aaron Vinson for his invaluable work on contributing to the statistical code used by the registry and Tracking Program staff.

[1]We acknowledge participants of Phase One of the Sub-County Cancer Pilot Project: Arizona Cancer Registry and Tracking Program, California Cancer Registry and Tracking Program, Florida Cancer Data System and Tracking Program, Georgia Comprehensive Cancer Registry, Cancer Data Registry of Idaho, Louisiana Tumor Registry and Tracking Program, Michigan Cancer Surveillance Program and Tracking Program, Minnesota Cancer Reporting System and Tracking Program, Missouri Cancer Registry and Research Center and Tracking Program, North Carolina Central Cancer Registry, Nebraska Cancer Registry, New Jersey State Cancer Registry and Tracking Program, North Dakota Statewide Cancer Registry, Rhode Island Cancer Control Registry and Tracking Program, South Carolina Central Cancer Registry, Texas Cancer Registry, Utah Cancer Registry and Tracking Program, Virginia Cancer Registry, Washington-Seattle Puget Sound Cancer Surveillance System and Tracking Program, and Wisconsin Cancer Reporting System and Tracking Program.

Funding

This research was supported in part by an appointment (TD Ellington) to the Research Participation Program at the Centers for Disease Control and Prevention administered by the Oak Ridge Institute for Science and Education through an interagency agreement between the U.S. Department of Energy and the Centers for Disease Control and Prevention.

New Jersey State Cancer Registry is funded by the National Cancer Institute's Surveillance, Epidemiology and End Results (SEER) Program (#75N91021D00009), Centers for Disease Control and Prevention's National Program of Cancer Registries (#5NU58DP006279) with additional support from the State of New Jersey and the Rutgers Cancer Institute of New Jersey.

Data availability

Please contact corresponding author for data request. Crosswalks and other sub-county materials can be found at https://github.com/CDCgov/EPHTracking-Subcounty.

References

2010 Census Summary File 1 [United States] Prepared By the U.S. Census Bureau, 2011.

Abookire S, Plover C, Frasso R, Ku B Health design thinking: an innovative approach in public health to defining problems and finding solutions. front public health. 2020;8:459. Published 2020 Aug 28. doi:10.3389/fpubh.2020.00459. [PubMed: 32984247]

Bell BS, Hoskins RE, Pickle LW, Wartenberg D, 2006. Current practices in spatial analysis of cancer data: mapping health statistics to inform policymakers and the public. Int. J. Health Geogr. 5, 49. 10.1186/1476-072X-5-49. Published 2006 Nov 8. [PubMed: 17092353]

Berney D, Camponeschi J, Coons M, Creswell PD, Schirmer J, Walsh R, 2015. Using environmental public health tracking to identify community targets for public health actions in childhood lead poisoning in Wisconsin. J. Public Health Manag. Pract. 21 (2), S80–S84. 10.1097/PHH.0000000000000161. Mar-AprSuppl.

Centers for Disease Control and Prevention, 2013. Creating a Health Profile of Your neigh- borhood: a how-To Guide. https://www.cdc.gov/healthyplaces/toolkit/sources_of_health_data.pdf.

Centers for Disease Control and Prevention, 2020. U.S. Cancer Statistics Data Visualizations Tool Technical Notes. https://www.cdc.gov/cancer/uscs/pdf/uscs-data-visualizations-tool-technical-notes-508.pdf.

DeSalvo KB, O'Carroll PW, Koo D, Auerbach JM, Monroe JA, 2016. Public Health 3.0: time for an Upgrade. Am. J. Public Health 106 (4), 621–622. 10.2105/AJPH.2016.303063. [PubMed: 26959263]

Howe HL, 1986. Geocoding NY state cancer registry. Am. J. Public Health 76, 1459–1460.

Jordan M, DuClos C, Folsom J, Thomas R, 2015. Developing a smartphone interface for the Florida Environmental Public Health Tracking Web portal. J. Public Health Manag. Pract. 21 (2), S50–S54. 10.1097/PHH.0000000000000158. Mar-AprSuppl. [PubMed: 25621446]

Lynch SM, Wiese D, Ortiz A, et al. , 2020. Towards precision public health: geospatial analytics and sensitivity/specificity assessments to inform liver cancer prevention. SSM Popul. Health 12, 100640. 10.1016/j.ssmph.2020.100640. Published 2020 Aug 7. [PubMed: 32885020]

McGeehin MA, Qualters JR, Niskar AS, 2004. National Environmental Public Health Tracking Program: bridging the information gap. Environ. Health Perspect. 112 (14), 1409–1413. [PubMed: 15471734]

National Program of Cancer Registries, 2020. Division of Cancer Prevention and Control. CDC. https://www.cdc.gov/cancer/npcr/index.htm.

Pfeiffer DU, Robinson TP, Stevenson M, Stevens KB, Rogers DJ, Clements AC, 2008. Spatial Analysis in Epidemiology. Oxford University Press, New YorkNY.

Rushton G, Armstrong MP, Gittler J, Greene BR, Pavlik CE, West MM, Zimmerman DL, 2006. Geocoding in cancer research: a review. Am. J. Prev. Med. 30 (2), S16–S24. 10.1016/j.amepre.2005.09.011. FebSuppl. [PubMed: 16458786]

Sahar L, Foster SL, Sherman RL, et al. , 2019. GIScience and cancer: state of the art and trends for cancer surveillance and epidemiology. Cancer 125 (15), 2544–2560. 10.1002/cncr.32052. [PubMed: 31145834]

SAS Institute Inc, 2013. Base SAS® 9.4 Utilities: Reference. SAS Institute Inc, Cary, NC.

Singh S, Ajani UA, Wilson RJ, Thomas CC, 2009. Evaluating quality of census tract data in the NPCR dataset. J. Registry Manag. 36 (4), 143–146. Winter. [PubMed: 20795557]

Tatalovich Z, Stinchcomb DG, Ng D, Yu M, Lewis DR, Zhu L, Feuer EJR, 2022. Developing geographic areas for cancer reporting using automated zone design. Am. J. Epidemiol. kwac155. 10.1093/aje/kwac155. Aug 31Epub ahead of print..

U.S. Cancer Statistics Working Group, 2020. U.S. Cancer Statistics Data Visualizations Tool, Based On 2019 Submission Data (1999–2017). U.S. Department of Health and Human Services, Centers for Disease Control and Prevention and National Cancer Institute. www.cdc.gov/cancer/dataviz. released in June.

U.S. Census Bureau, 2017. Cartographic Boundary Shapefiles - census tracts, https://www.census.gov/geo/maps-data/data/cbf/cbf_tracts.html.

Werner AK, Strosnider H, Kassinger C, Shin M, 2018. Sub-county data project workgroup. lessons learned from the environmental public health tracking sub-county data pilot project. J. Public Health Manag. Pract. 24 (5), E20–E27. 10.1097/PHH.0000000000000686. Sep/Oct. [PubMed: 29227419]

Werner AK, Strosnider HM, 2020. Developing a surveillance system of sub-county data: finding suitable population thresholds for geographic aggregations. Spat. Spatiotemporal. Epidemiol. 33, 100339 10.1016/j.sste.2020.100339. JunEpub 2020 Mar 6.. [PubMed: 32370944]

Yu M, Reiter JP, Zhu L, Liu B, Cronin KA, Feuer EJR, 2017. Protecting confidentiality in cancer registry data with geographic identifiers. Am. J. Epidemiol. 186 (1), 83–91. 10.1093/aje/kwx050. [PubMed: 28453646]

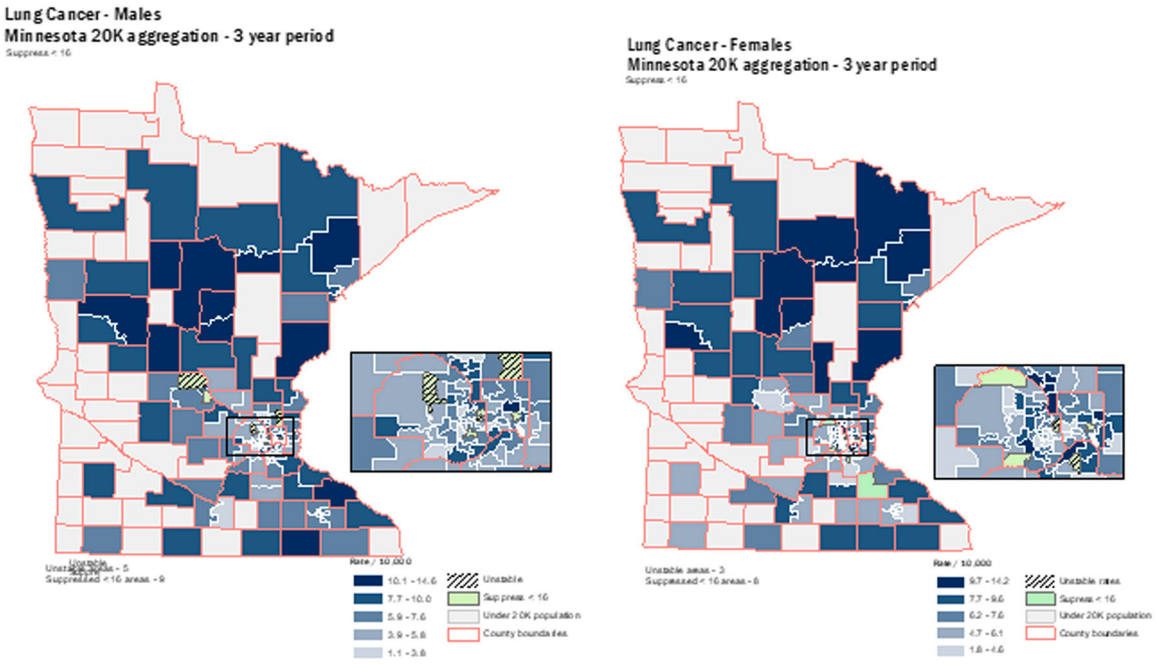

Fig. 1.

Minnesota lung cancer maps[a] by sex, using 20,000 person[b] aggregation over a 3-year period Figure displays map for Minnesota lung cancer by sex 2014–2016 rates for 2010 census tract based on a minimum population aggregation of 20,000 persons. Map includes legend that explains color codes for rate per 10,000 population and tracts with suppressed rates.

[a]Maps were submitted by all participants. Map displayed is an example of one of the submissions.

[b]Counties that had <20,000 persons were not included and are shown as suppressed in maps above. Counties with a relative standard error of >30% were considered unstable and are hatched. Case counts <16 were suppressed.

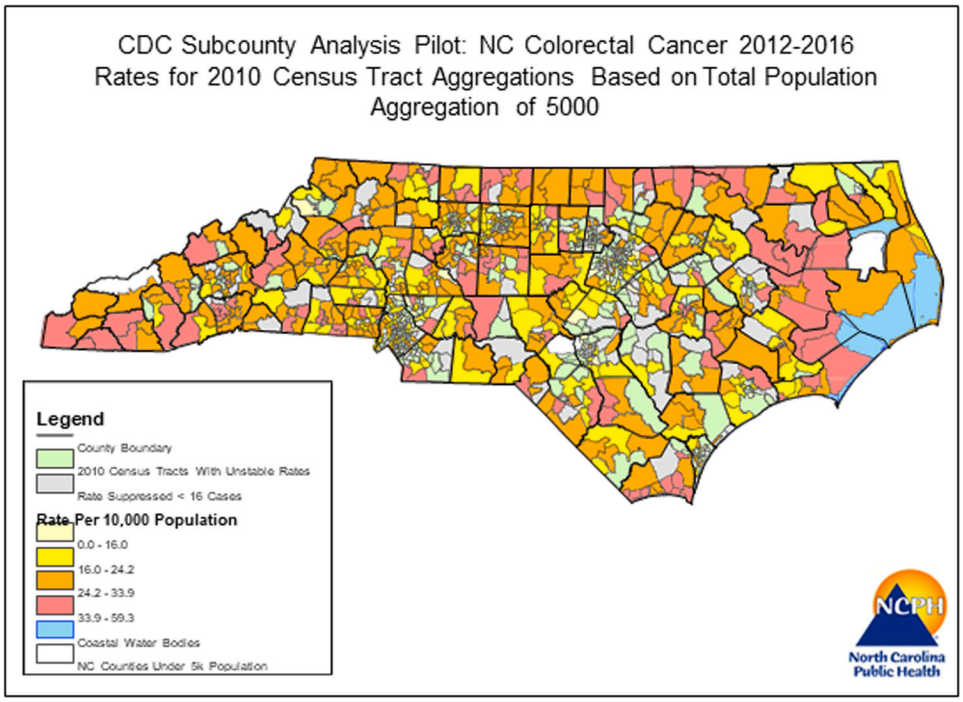

Fig. 2.

North Carolina colorectal cancer maps[a], using 5000 person[b] aggregation over a 5-year period Figure displays map for NC colorectal cancer 2012–2016 rates for 2010 census tract based on a minimum population aggregation of 5000 persons. Map includes legend that explains color codes for rate per 10,000 population and tracts with suppressed rates.

[a]Maps were submitted by all participants. Map displayed is an example of one of the submissions.

[b]Counties that had <5000 persons were not included and are shown as suppressed in maps above. Counties with <30% were considered unstable. Case counts <16 were suppressed.

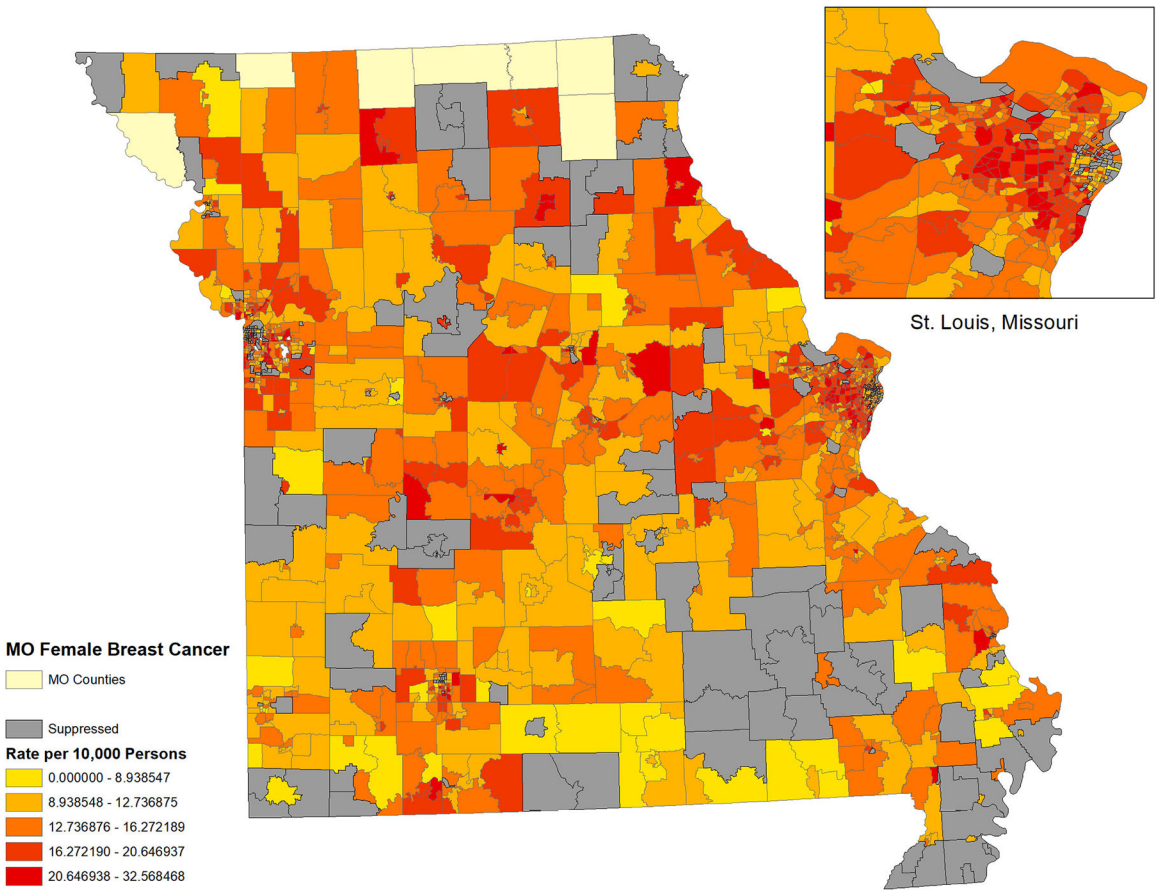

Fig. 3.

Missouri breast cancer maps[a] using census tract aggregation over a 10-year period (2007–2016)[b] Figure dispays maps for Missouri breast cancer rates using census tract aggregation over a 10-year period (2007–2016). Map includes legend that explains color codes for rate per 10,000 population and tracts with suppressed rates.

[a]Maps were submitted by all participants. Maps displayed is an example of one of the submissions.

[b] Counties with <30% were considered unstable. Case counts <16 were suppressed.

Table 1

Overview of aggregation schemes, the suggested median case count ranges, and population thresholds for a single year..

Classification	Median case count range	Spatial aggregation level
Very common outcome	≥ 17.0 cases	Census tract
Common outcome	7.3 to 16.9 cases	Aggregated minimum population 5000 persons
Rare outcome	1.9 to 7.2 cases	Aggregated minimum population 20,000 persons

Table 2

Average median, minimum, and maximum case counts by aggregation level and cancer type.

	Census tract Annual	5000 person aggregation 5-year period	20,000 person aggregation 3-year period	5-year period
Lung cancer				
Region 1	3 (0–19)	31 (1–116)	70 (16–238)	123 (29–322)
Region 2	3 (0–15)	30 (5–109)	64 (14–166)	97 (23–268)
Region 3	2 (0–19)	23 (0–119)	47 (8–164)	77 (14–268)
Colorectal cancer	2 (0–12)	19 (2–77)	44 (11–144)	
Non-Hodgkin lymphoma	1 (0–7)	10 (1–37)		36 (12–94)
Melanoma of the skin	1 (0–9)	16 (4–58)	32.5 (5–98)	
Female breast	2 (0–15)	33 (4–106)	77 (24–202)	
Prostate cancer	2(0–22)	21 (2–104)	51 (13–157)	
Liver and IHB	1 (0–5)	5(0–21)	9 (1–35)	

Table 3

Suggested spatiotemporal aggregation combinations for selected cancer types.

	Census tract	5000 person aggregation	20,000 person aggregation
Lung cancer			
Region 1	Not feasible to display	5-year period	5-year period
Region 2	10-year period	5-year period	3-year period
Region 3	Not feasible to display	7-year period	5-year period
Consensus	Not feasible to display	5-year period	5-year period
Lung cancer (female)			
Region 1	Not feasible to display	Not feasible to display	3 or 5-year period
Region 2	Not feasible to display	10-year period	3-year period
Region 3	Not feasible to display	Not feasible to display	Not feasible to display
Consensus	Not feasible to display	5-year period	5-year period
Lung Cancer (male)			
Region 1	Not feasible to display	Not feasible to display	3 or 5-year period
Region 2	Not feasible to display	10-year period	3-year period
Region 3	Not feasible to display	Not feasible to display	Not feasible to display
Consensus	Not feasible to display	5-year period	5-year period
Region 1			
Colorectal Cancer	Not feasible to display	5-Year period	3-year period
NHL	Not feasible to display	Not feasible to display	5-year period
Region 2			
Breast Cancer	10 -Year period	5-Year period	3-year period
Melanoma	Not feasible to display	5-Year period	3-year period
Region 3			
Prostate Cancer	10 -Year period	5-Year period	3-year period
Liver and IHB	Not feasible to display	Not feasible to display	Not feasible to display