Published in final edited form as:

Cancer Epidemiol Biomarkers Prev. 2018 November ; 27(11): 1252–1260. doi:
10.1158/1055-9965.EPI-17-1087.

Utility and Generalizability of Multi-State, Population-Based Cancer Registry Data for Rural Cancer Surveillance Research in the United States

Whitney E. Zahnd¹, Wiley D. Jenkins¹, Aimee S. James², Sonya R. Izadi², David E. Steward³, Amanda J. Fogleman¹, Graham A. Colditz², and Laurent Brard^{4,5}

¹Office of Population Science and Policy, Southern Illinois University School of Medicine, Springfield, IL

²Division of Public Health Sciences, Department of Surgery, Washington University School of Medicine, and Siteman Cancer Center St. Louis, MO

³Department of Internal Medicine, Southern Illinois University School of Medicine, Springfield, IL

⁴Division of Gynecologic Oncology, Department of Obstetrics and Gynecology, Southern Illinois University School of Medicine, Springfield, IL

⁵Simmons Cancer Institute at SIU, Southern Illinois University School of Medicine, Springfield, IL

Abstract

More than 46 million Americans live in rural areas, but rural populations remain relatively understudied in cancer disparities research. However, several analyses of multi-state cancer registry data that describe the rural cancer incidence burden have been recently published. In light of this, our commentary aimed to characterize the utility and generalizability of multi-state, population-based cancer registry datasets for rural cancer surveillance research. First, we described the accessibility, geographic coverage, available variables, and strengths and weaknesses of five data sources. Second, we evaluated two of these data sources—the North American Association of Central Cancer Registries (NAACCR) public use dataset (93% population coverage) and the Surveillance Epidemiology and End Results (SEER) 18 dataset (28% population coverage)—on their characterization of rural-urban cancer incidence rates and sociodemographic representation. The five data sources varied in geographic coverage and extent of available variables. SEER 18's cancer rates sociodemographic representation differed from the more geographically representative NAACCR data. We suggest that SEER increase its geographic coverage to improve their generalizability and to take advantage of their utility to assess disparities along the cancer control continuum. We also suggest that non-SEER data sources be utilized more frequently to capitalize on their extensive geographic coverage.

Corresponding Author: Whitney E. Zahnd, PO Box 19664, 201 E. Madison Rm. 235, Springfield, IL 62794-9664, Phone: (217)545-2428, Fax: (217)545-0799, wzahnd@siu.edu.

Conflicts of Interest: The authors have no conflicts of interest to report.

Keywords

rural health; cancer surveillance; SEER; cancer disparities

Introduction

The National Cancer Institute characterizes cancer health disparities as adverse differences in cancer burden along the cancer control continuum from primary prevention to incidence to survivorship among specific population groups, including those defined by their geographic location (e.g. rural vs. urban) (1). More than 46 million residents live in rural areas, but cancer disparities among rural populations are still relatively understudied (2). However, there is an increasing emphasis on the need to address cancer control in rural and other small populations through appropriate research methods, evidence-based public health and clinical interventions, and community engagement (3-6). Further, recent descriptive analyses of population-based, cancer registry data have illuminated the cancer incidence disparities experienced by rural populations in the U.S. – highlighted in our work in this “Rural Cancer Control” *CEBP Focus* issue and other recent publications from Center for Disease Control and Prevention (CDC) and National Cancer Institute (NCI) researchers (6-8). Each of these analyses utilized different sources of multi-state, population-based cancer registry data with varying extents of geographic coverage of the country. Our paper utilized the North American Association of Central Cancer Registries (NAACCR) public use dataset, which covers 93% of the United States (7). CDC researchers utilized data from the National Program of Cancer Registries (NPCR) and Surveillance Epidemiology and End Results (SEER) program (97% coverage), and NCI researchers utilized data from the SEER 18 program (28% coverage) (6,8). Interestingly, the NCI analysis of SEER 18 data indicated that the all cancers combined incidence rate was higher in rural populations than in urban, while the CDC's analysis of NPCR/SEER data and our analysis of NAACCR data found that the all cancers combined incidence rate was higher in urban populations.

Cancer surveillance data are useful for describing differences in incidence rates and trends among populations (e.g. rural and urban), but such data can also be used to assess cancer mortality, survival, risk factors, and outcomes and to help guide interventions and public policy (9). The strength of cancer registry systems is that they collect population-based data on individual cancer cases using consistent data standards which enhances the validity and generalizability of studies utilizing such data (10). Each central cancer registry is required to collect data on demographics, cancer identification, staging, treatment, follow-up, and death (11). However, each central cancer registry differs on what variables they may collect beyond required data, what processes researchers must utilize to obtain data, and what specific data elements may be available to outside researchers (12-13). While researchers may be interested in studying populations within a single state, NAACCR, NCI, and CDC all have avenues by which researchers can obtain multi-state data for cancer surveillance research. Multi-state data may be particularly meaningful for evaluating rural or other small populations nationally, assessing policy-relevant geographic contexts (e.g. the Delta Regional Authority or Appalachian Regional Commission), or exploring other important

multi-state categorizations (e.g. Medicaid expansion vs. non-Medicaid expansion states) (14-16).

In light of the recent studies utilizing multi-state cancer registry data to describe the rural cancer burden, this commentary aimed to evaluate the utility and generalizability of multi-state, population-based cancer registry datasets for rural cancer surveillance purposes.

Methods

To achieve our goal, our approach was two-fold. First, we described available multi-state, population-based cancer datasets and their strengths and limitations for rural cancer surveillance research. Second, we explored two readily available population-based cancer registry datasets to assess their generalizability to rural populations.

Description of Multi-State, Population-Based Cancer Registry Data

Information on all incident cases of cancers in the U.S. are collected through either the CDC-funded NPCR or NCI-funded SEER registry programs (11). Together NPCR and SEER provide data on 100% of cancer cases diagnosed in the United States. There are mechanisms to request access to multi-state data like the NPCR and SEER through the CDC and NCI, respectively. There are also mechanisms to access multi-state, population-based cancer registry data through NAACCR, an umbrella organization for central cancer registries. NAACCR develops and promotes cancer registry data standards, provides training and certification, and publishes and promotes the use of cancer registry data for cancer surveillance research and cancer control (17). We evaluated the accessibility, geographic coverage and scale, extent of available variables, and strengths and weaknesses of five datasets available through NAACCR, CDC, or NCI, which are the only conduits to obtain multi-state cancer registry data. These datasets included: the NAACCR public use dataset, the NAACCR Cancer in North America (CiNA) Deluxe File, SEER, NPCR public use dataset, and the United States Cancer Statistics (USCS) restricted access dataset (18-22).

Generalizability of SEER and NAACCR Analyses to Rural Populations

The NAACCR public use dataset, which we analyzed in our published study in this *Focus* issue covers 93% of the U.S. population, while SEER 18 covers 28% of the U.S. population across 18 different registries (7,18,20). There are three datasets available with the SEER program: SEER 9, SEER 13, and SEER 18. Each of these datasets includes an increasing number of registries and different years of data based upon the inclusion of each registry in the SEER program. To assess the generalizability of each dataset to rural populations, and to assess rural-urban differences in cancer incidence, we calculated age-adjusted cancer incidence rates and corresponding 95% confidence intervals for rural and urban populations in SEER 18 and the NAACCR public use dataset for cancers diagnosed between 2009 and 2013. Rate ratios with p-values less than 0.05 indicated statistically significant rural-urban differences in rates. Corresponding age-adjusted incidence rates and confidence intervals are reported from our study utilizing NAACCR's public use dataset. All analyses were done using SEER*Stat 8.3.4. We utilized these two datasets because, of the five datasets we reviewed, the NAACCR public use dataset and SEER 18 dataset are the only ones that are

publicly available, include a rural-urban variable, and are able to be analyzed remotely under a data use agreement without additional approvals from the data-managing entity. To categorize rural and urban, we used the United States Department of Agriculture's Rural Urban Continuum Codes (RUCCs), which classify counties by population size and proximity to a metropolitan area (2). We used a dichotomized RUCC designation (rural RUCC = 4-9; urban RUCC = 1-3), as this dichotomized designation was the only available rural-urban measure in the NAACCR dataset. We also performed this analysis stratified by race/ethnicity for non-Hispanic whites, non-Hispanic blacks, and Hispanics to evaluate the generalizability to the growing racial/ethnic diversity in rural populations (23). To further assess the generalizability of SEER 18 and the NAACCR public use datasets, we used 2009-2013 American Community Survey population overall estimates and stratified by race, ethnicity, poverty, and median household income to characterize rural populations covered by counties represented in SEER 18 and the NAACCR public use dataset and the United States population as a whole (24).

Results

Description of Multi-State, Population-Based Cancer Registry Data

Table 1 summarizes the accessibility, geographic scale and coverage, variable availability, strengths, and weaknesses of several multi-state, population-based cancer registry sources. With the exception of the USCS restricted access dataset, all datasets are remotely accessible through SEER*Stat software. However, only NAACCR CiNA Deluxe and SEER data are accessible through the case listing session for extraction of individual level data to be analyzed in other software programs. Both the NAACCR CiNA Deluxe datasets and USCS restricted use datasets are only available for specific research projects after appropriate approvals from their respective entities. The geographic coverage varied from 28% of the US population in SEER 18 to a potential of 100% coverage in the USCS restricted access dataset. While all data reported to central cancer registries are required to be geocoded to the census tract, data available through these mechanisms are only available most precisely at the county level. Such geographic specification is necessary for rural-urban characterization of cancer data (e.g. utilizing USDA rural-urban metrics), but each central cancer registry may restrict/limit the release of geographic information. The NAACCR public use dataset includes a dichotomous rural-urban variable, but identification of specific county is not available. The NPCR public use dataset does not include any sub-state geographic indicators, severely limiting the characterization of rural cancer burden. All datasets have data on tumor characteristics, but only the NAACCR CiNA Deluxe and SEER datasets have any data available on treatment or outcomes.

Generalizability of SEER and NAACCR Analyses to Rural Populations

Table 2 displays the rural-urban age-adjusted cancer incidence rates utilizing 2009-2013 data from both the NAACCR public use dataset and SEER 18. The direction and magnitude of rural-urban all cancer incidence differences varied between the two data sources. The NAACCR dataset identified a slightly higher incidence rate in urban compared to rural populations for all cancers combined (448.7 vs. 446.4 per 100,000, respectively). SEER 18, on the other hand, identified a higher incidence rate in rural compared to urban populations

for all cancers combined (461.8 vs. 449.1 per 100,000, respectively). Among 20 of the 23 individual cancers that we evaluated, both NAACCR and SEER 18 data showed either no rural-urban differences in rates or differences that were consistent in direction. For pancreas cancer, NAACCR identified a higher rate in urban populations, while there was no rural-urban difference in SEER 18. For melanoma, NAACCR identified a slightly higher rate in urban populations while SEER 18 analysis indicated a slightly higher rate in rural populations. For myeloma, NAACCR data indicated a higher rate in urban populations while SEER 18 data indicated no rural-urban difference in rates.

Table 3 displays the race-ethnicity specific age-adjusted cancer incidence rates for rural and urban populations utilizing 2009-2013 data from both the NAACCR public use dataset and SEER 18. Both the NAACCR and SEER 18 analyses indicated elevated overall rates in urban areas among non-Hispanic whites, non-Hispanics blacks, and Hispanics. For non-Hispanic whites, both NAACCR and SEER 18 indicated the rural-urban differences across all cancer types. In general, the age-adjusted incidence rates in the analysis of SEER 18 data were higher than the analysis of NAACCR data for both rural and urban non-Hispanic whites. For non-Hispanic blacks, the NAACCR analysis identified two small rural-urban differences (Hodgkin lymphoma and myeloma) that were not identified in the SEER 18 analysis. Rural-urban differences in kidney and urinary bladder cancers were identified amongst non-Hispanic blacks in the SEER 18 analysis.

Table 4 displays the sociodemographic representation of SEER 18 and the NAACCR public use datasets compared to the nation as a whole. Rural counties represent 10.6% and 14.6% of the SEER 18 and NAACCR populations, respectively, while nationally, 14.8% of the population is rural. Rural counties represented in SEER have higher proportions of black, Asian, and Hispanic populations compared to the NAACCR population and the national populations (e.g. SEER 18 Hispanic populations were 9.1% compared to 7.8% in NAACCR and 7.7% nationwide). SEER 18 rural counties had a slightly higher percentage of the population living in poverty (19.7%) compared to the counties represented in NAACCR (18.3%) and the US rural population (18.1%). The regional distribution of rural populations in SEER 18 are notably different from the US population overall, as the Northeast and Midwest are considerably underrepresented while the South and West are overrepresented.

Discussion

This commentary provided an overview of cancer surveillance datasets that can be used for rural cancer surveillance research and described the generalizability of two datasets—the NAACCR public use dataset and SEER 18—with varying geographic coverage (93% and 28% representation of the U.S. population, respectively). Datasets also varied by opportunity to perform analyses beyond rates and trends and potential to explore cancer outcomes beyond incidence/tumor characteristics. Age-adjusted incidence rates were largely similar among both NAACCR and SEER 18 analyses for individual cancer types, but they differed in all cancer combined analyses. The NAACCR analysis indicated a slightly elevated rate in urban populations, while the SEER 18 analysis indicated a higher rate among rural populations. However, when the SEER 18 analysis was stratified by race/ethnicity higher

rates in urban populations were identified across all groups. Further, the geography of SEER 18 registries indicated underrepresentation of rural populations.

Data accessible through the CDC have greater geographic coverage (96-100% of the United States) than other sources. A recent study by Kuo and colleagues utilized the USCS restricted use dataset to evaluate its generalizability to the U.S. population and recommends that researchers take greater advantage of this resource because of its more complete geographic coverage compared to SEER 18 (25). Indeed, both the USCS restricted use dataset and the NPCR public use dataset have great geographic coverage, but only have variables relevant to tumor characterization (e.g. site, stage). The restricted use dataset, which has the potential to characterize rural areas, is only available for analysis at research data centers, which may be a barrier to some researchers. Additionally, the NPCR public use dataset while remotely accessible, does not include any characterization of rural and urban populations, as there is no geographic indicator available more granular than state. Further, this dataset only allows for analysis of incidence rates and trends.

The CDC has made addressing rural disparities an increasing priority, as highlighted by a series of surveillance reports on areas such as cancer disparities and health behaviors (26). Because of this increasing priority, we offer two recommendations. First, we suggest that data on treatment and outcomes be integrated into the USCS restricted use dataset to allow for quantification of rural-urban disparities/differences at points along the cancer control continuum (e.g. treatment, survival) to be explored using a generalizable, population-based data source. Second, we suggest that the public use dataset incorporate a rural-urban variable(s), such as a USDA rural-urban measure, much like that available in the NAACCR public use dataset, to allow researchers to utilize this generalizable data source to describe differences/disparities in cancer rates and trends. This would allay potential confidentiality and privacy concerns when working with counties with small numbers/populations while providing researchers the opportunity to evaluate rural-urban differences in cancer incidence.

While data accessible from the CDC is more representative of rural populations, SEER underrepresents some rural populations in its geographic coverage. The distribution of rates by rural-urban designation in our study using NAACCR data corroborated those published by the CDC using similarly representative data, which indicated nominally higher overall rates in urban areas, while SEER indicated higher overall rates in rural areas (7,8). Further, the sociodemographic characteristics of counties represented in SEER 18 differ from the US rural population as a whole. This corroborates previous studies which indicated that geographic areas covered by SEER were more urban than non-SEER areas of the country, one of which also showed that SEER's rural representation has become less generalizable to the nation over time (25,27). A strength of SEER is that there is purposeful inclusion of populations who have been historically underrepresented, particularly racial and ethnic minority populations. Indeed, over the years the coverage of SEER has expanded to include registries like those incorporating rural Georgia, an area with a high proportion of rural black residents. However, as a whole, rural populations are not well-represented in SEER. Like racial and ethnic minorities, rural populations have been historical underrepresented and understudied in cancer research. Two priorities in NCI's Division of Cancer Control and

Population Sciences include both rural cancer control and enhancing SEER for research use (28). Further, one of the goals of the SEER program is to identify changes and differences in cancer incidence among geographically and otherwise sociodemographically defined population subgroups (29). To simultaneously address these priority areas and goals, the SEER program could improve its rural coverage by inclusion of additional registries to enrich the data for cancer disparities research purposes. This is especially important because SEER data are the most readily accessible population-based dataset to explore cancer treatment and survival. Additionally, other datasets are available that link SEER data to Medicare, Medicare-Health Outcomes Survey, and Medicare-Consumer Assessment of Healthcare Providers (30). Thus, it is a valuable resource for not only cancer surveillance research, but also for health services research that may elucidate rural-urban disparities in cost, utilization of tests and procedures, quality of life, and quality of care. For example, SEER-Medicare data have been used to explore temporal trends in cancer incidence and mortality, treatment, and costs among rural populations or in comparing rural-urban differences (31-33). Expanding SEER's geographic coverage will help researchers address rural health service research questions in addition to rural cancer surveillance research questions.

Aggregated analysis of SEER 18 data revealed elevated incidence rates in rural compared to urban areas. However, when rates were stratified by race/ethnicity, the urban rates were higher across all groups. Because of the solely descriptive nature of our commentary's analysis, this result may be attenuated with more rigorous analyses (34-35). Further, it is indicative of an imbalance of the regional distribution of rural and urban populations by racial/ethnic groupings within SEER (i.e. a conditional disparity) (36). This phenomenon was not identified with our NAACCR analysis, however, likely due to the fact that its distribution among geographic regions largely mirrors that of the nation. This phenomenon also underscores the importance of considering the interplay of race and rurality on health outcomes especially across multi-state regions and the need to expand SEER 18's geographic coverage (37).

NAACCR's CiNA Deluxe File provides high quality data on multiple cancer outcomes, including incidence, site specific factors (e.g. breast cancer subtype), treatment, and survival. The potential availability of two rural-urban measures (RUCC and RUCA) and county of residence for cancer cases make it a valuable data source for studying cancer in rural populations, although release of county-specific data requires active consent from each registry. The CiNA Deluxe file has been used to explore cancer disparities utilizing data from as many as forty central cancer registries and to explore disparities within policy-relevant, multi-state rural regions like Appalachia (38-39). Additionally, the extensive geographic coverage of the NAACCR public use dataset provides a generalizable data resource for rural cancer surveillance research. While the dichotomous nature of the rural-urban variable within the public use dataset may limit more nuanced rural characterization, it provides data with broad geographic coverage. There is a tremendous opportunity for researchers to use these data for rural cancer surveillance research.

As rural cancer disparities are increasingly studied, it is important to ensure that population-level surveillance data are generalizable to the rural population and that these data can be

utilized to assess multiple aspects of the cancer control continuum from prevention to diagnosis and to survival. We provide recommendations to entities who provide access to population-based data and to researchers who utilize these data to enhance the rural cancer surveillance research. We recommend that the CDC increase the available variables in their two, multi-state datasets to enable researchers to perform research that is generalizable to the population and that NCI should expand SEER's geographic coverage to capitalize on the availability of treatment and outcome variables and linkage to Medicare and other datasets. Further, we suggest that researchers take advantage of the USCS restricted use dataset and NAACCR's CiNA deluxe datasets for rural cancer surveillance research purposes.

Acknowledgments

Financial Support: Dr. Jenkins, Ms. Fogleman, Dr. Steward, and Dr. Brard (1P20CA192987-01A1) and Dr. James, Ms. Izadi, and Dr. Colditz (1P20CA192966-01A1) are funded by grants from the National Cancer Institute.

References

1. National Cancer Institute. Rockville, MD: National Cancer Institute; How does the National Cancer Institute (NCI) define "cancer health disparities"? Available from <<https://www.cancer.gov/about-nci/organization/crchd/cancer-health-disparities-fact-sheet#q1>>
2. United States Department of Agriculture. Washington, DC: United States Department of Agriculture; 2013. Rural Urban Continuum Codes. Available from: <http://www.ers.usda.gov/data-products/rural-urban-continuum-codes/documentation.aspx>
3. Wheeler SB, Davis MM. "Taking the Bull by the Horns": Four Principles to Align Public Health, Primary Care, and Community Efforts to Improve Rural Cancer Control. *J Rural Health*. 2017; doi: 10.1111/jrh.12263
4. Meilleur A, Subramanian SV, Plascak JJ, Fisher JL, Paskett ED, Lamont EB. Rural Residence and Cancer Outcomes in the United States: Issues and Challenges. *Cancer Epidemiol Biomarkers Prev*. 2013; 22(10):1657–1667. [PubMed: 24097195]
5. Srinivasan S, Moser RP, Willis G, Riley W, Alexander M, Berrigan D, et al. Small Is Essential: Importance of Subpopulation Research in Cancer Control. *Am J Public Health*. 2015 Jul; 105(3):S371–3. DOI: 10.2105/AJPH.2014.302267 [PubMed: 25905825]
6. Blake KD, Moss JL, Gaysynsky A, Srinivasan S, Croyle RT. Making the Case for Investment in Rural Cancer Control: An Analysis of Rural Cancer Incidence, Mortality, and Funding Trends. *Cancer Epidemiol Biomarkers Prev*. 2017; 26(7):992–7. DOI: 10.1158/1055-9965.epi-17-0092 [PubMed: 28600296]
7. Zahnd WE, James AS, Jenkins WD, Izadi SR, Fogleman AJ, Steward DE, Colditz GA, Brard L. Rural-Urban Differences in Cancer Incidence and Trends in the United States. *Cancer Epidemiol Biomarkers Prev*. 2017 Jul 27. [Epub ahead of print].
8. Henley SJ, Anderson RN, Thomas CC, Massetti GM, Peaker B, Richardson LC. Invasive Cancer Incidence, 2004–2013, and Deaths, 2006–2015, in Nonmetropolitan and Metropolitan Counties - United States. *MMWR Surveill Summ*. 2017; 66(14):1–13. DOI: 10.15585/mmwr.ss6614a1
9. Glaser SL, Clarke CA, Gomez SL, O'Malley CD, Purdie DM, West DW. Cancer surveillance research: a vital subdiscipline of cancer epidemiology. *Cancer Causes Control*. 2005; 16:1009–1019. [PubMed: 16184466]
10. Howe HL, Edwards BK, Young JL, Shen T, West DW, Hutton M, et al. A Vision for Cancer Incidence Surveillance in the United States. *Cancer Causes Control*. 2003; 14(7):663–672. [PubMed: 14575364]
11. Ryerson AB, Massetti GM. CDC's Public Health Surveillance of Cancer. *Prev Chronic Dis*. 2017; 14:E39.doi: 10.5888/pcd14.160480 [PubMed: 28520553]

12. North American Association of Central Cancer Registries. Springfield, IL: North American Association of Central Cancer Registries; Registry Research Information Database Search. Available from: <http://rridb.naaccr.org/>
13. National Program of Cancer Registries. Atlanta, GA: Center for Disease Control and Prevention; CDC Cancer Registry Data Access for Research Project. Available from: https://www.cdc.gov/cancer/npcr/data_access/
14. Wilson RJ, Ryerson AB, Singh SD, King JB. Cancer Incidence in Appalachia, 2004–2011. *Cancer Epidemiol Biomarkers Prev.* 2016 Feb; 25(2):250–8. Epub 2016 Jan 27. DOI: 10.1158/1055-9965.EPI-15-0946 [PubMed: 26819264]
15. Zahnd WE, Jenkins WD, Mueller-Luckey GS. Cancer Mortality in the Mississippi Delta Region: Descriptive Epidemiology and Needed Future Research and Interventions. *J Health Care Poor Underserved.* 2017; 28(1):315–328. DOI: 10.1353/hpu.2017.0025.
16. Soni A, Simon K, Cawley J, Sabik L. Effect of Medicaid Expansions of 2014 on Overall and Early-Stage Cancer Diagnoses. *Am J Public Health.* 2018 Feb; 108(2):216–218. Epub 2017 Dec 21. DOI: 10.2105/AJPH.2017.304166 [PubMed: 29267058]
17. North American Association of Central Cancer Registries. Springfield, IL: North American Association of Central Cancer Registries; About NAACCR. Available from: <https://www.naaccr.org/about-naaccr/>
18. SEER*Stat Database: NAACCR Incidence Data - CiNA Analytic File -, Public Use (which includes data from CDC's National Program of Cancer Registries (NPCR), CCCR's Provincial and Territorial Registries, and the NCI's Surveillance, Epidemiology and End Results (SEER) Registries), certified by the North American Association of Central Cancer Registries (NAACCR) as meeting high-quality incidence data standards for the specified time periods submitted December 2015.
19. North American Association of Central Cancer Registries. Springfield, IL: North American Association of Central Cancer Registries; CiNA Deluxe for Researchers. Available from: <https://www.naaccr.org/cina-deluxe-for-researchers/>
20. Surveillance, Epidemiology, and End Results (SEER) Program (www.seer.cancer.gov) SEER*Stat Database: Incidence - SEER 18 Regs Research Data + Hurricane Katrina Impacted Louisiana Cases, Nov 2016 Sub (2000–2014) <Katrina/Rita Population Adjustment> - Linked To County Attributes - Total U.S., 1969–2015 Counties, National Cancer Institute, DCCPS, Surveillance Research Program, released April 2017, based on the November 2016 submission.
21. 2001–2014 Database: National Program of Cancer Registries and Surveillance, Epidemiology, and End Results SEER*Stat Database: NPCR and SEER Incidence – USCS 2001–2014 Public Use Research Database, United States Department of Health and Human Services, Centers for Disease Control and Prevention and National Cancer Institute. Released August, 2017, based on the November 2016 submission. Accessed at www.cdc.gov/cancer/npcr/public-use.
22. Center for Disease Control and Prevention. Atlanta, GA: Center for Disease Control and Prevention; United States Cancer Statistics. Available from: <https://www.cdc.gov/rdc/b1datatype/Dt131.htm>
23. Lichter DT. Immigration and the New Racial Diversity in Rural America. *Rural Sociol.* 2012; 77(1):3–35. DOI: 10.1111/j.1549-0831.2012.00070.x [PubMed: 26478602]
24. U.S. Census Bureau. American Community Survey (ACS). Suitland, MD: U.S. Census Bureau; Available from <https://www.census.gov/programs-surveys/acs/data.html>
25. Kuo TM, Mobley LR. How generalizable are the SEER registries to the cancer populations of the USA? *Cancer Causes Control.* 2016; 27(9):1117–26. DOI: 10.1007/s10552-016-0790-x [PubMed: 27443170]
26. Center for Disease Control and Prevention. Atlanta, GA: MMWR Rural Health Series Center for Disease Control and Prevention; Available from https://www.cdc.gov/mmwr/rural_health_series.html
27. Nattinger AB, McAuliffe TL, Schapira MM. Generalizability of the surveillance, epidemiology, and end results registry population: factors relevant to epidemiologic and health care research. *J Clin Epidemiol.* 1997; 50(8):939–45. [PubMed: 9291879]

28. National Cancer Institute Division of Cancer Control & Population Sciences. Rockville, MD: National Cancer Institute; 2017. Overview and Highlights. Available from: <https://cancercontrol.cancer.gov/bb/index.html>
29. National Cancer Institute Surveillance, Epidemiology, and End Results Program. Rockville, MD: National Cancer Institute; Goals of the SEER Program. Available from: <https://seer.cancer.gov/about/goals.html>
30. National Cancer Institute Division of Cancer Control & Population Sciences. Rockville, MD: National Cancer Institute; Publicly Accessible Data. Available from: <https://healthcaredelivery.cancer.gov/initiatives/public.html>
31. Liang PS, Mayer JD, Wakefield JD, Ko CW. Temporal Trends in Geographic and Sociodemographic Disparities in Colorectal Cancer Among Medicare Patients, 1973-2010. *J Rural Health*. 2017; 33(4):361–370. [PubMed: 27578387]
32. Camacho FT, Tan X, Alcalá HE, Shah S, Anderson RT, Balkrishnan R. Impact of patient race and geographical factors on initiation and adherence to adjuvant endocrine therapy in medicare breast cancer survivors. *Medicine (Baltimore)*. 2017; 96(24):e7147. [PubMed: 28614244]
33. Shirvani SM, Jiang J, Likhacheva A, Hoffman KE, Shaitelman SF, Caudle A, Buchholz TA, Giordano SH, Smith BD. Trends in Local Therapy Utilization and Cost for Early-Stage Breast Cancer in Older Women: Implications for Payment and Policy Reform. *Int J Radiat Oncol Biol Phys*. 2016; 95(2):605–16. [PubMed: 27034179]
34. Fu P, Pannerselvam A, Clifford B, Dowlati A, Ma PC, Zeng G, Halmos B, Leidner RS. Simpson's paradox - aggregating and partitioning populations in health disparities of lung cancer patients. *Stat Methods Med Res*. 2015 Dec; 24(6):937–48. [PubMed: 22246415]
35. Hernan MA, Clayton D, Keiding N. The Simpson's paradox unraveled. *Int J Epidemiol*. 2011 Jun; 40(3):780–5. DOI: 10.1093/ije/dyr041 [PubMed: 21454324]
36. Duan N, Meng XL, Lin JY, Chen C, Alegria M. Disparities in Defining Disparities: Statistical Conceptual Frameworks. *Stat Med*. 2008 Sep 10; 27(20):3941–56. DOI: 10.1002/sim.3283 [PubMed: 18626925]
37. Probst JC, Moore CG, Glover SH, Samuels ME. Person and place: the compounding effects of race/ethnicity and rurality on health. *Am J Public Health*. 2004 Oct; 94(10):1695–703. [PubMed: 15451735]
38. Cole L, Polfus L, Peters ES. Examining the incidence of human papillomavirus-associated head and neck cancers by race and ethnicity in the U.S., 1995-2005. *PLoS One*. 2012; 7(3):e32657. Epub 2012 Mar 20. doi: 10.1371/journal.pone.0032657 [PubMed: 22448226]
39. Huang B, Luo A, Durbin EB, et al. Incidence of CNS tumors in Appalachian children. *J Neurooncol*. 2017 May; 132(3):507–512. Epub 2017 Mar 11. DOI: 10.1007/s11060-017-2403-2 [PubMed: 28285334]

Summary of Multi-State, Population-Based Cancer Registry Data Sources for U.S. Rural Cancer Surveillance Research

Table 1

	North American Association of Central Cancer Registries (NAACCR) Public Use Data Set ¹⁸	NAACCR Cancer in North America Deluxe File ¹⁹	Surveillance Epidemiology and End Results (SEER) Dataset ²⁰	National Program of Cancer Registries (NPCR) Public Use Dataset ²¹	United States Cancer Statistics Restricted Access Dataset ²²
Accessibility (<i>no costs associated with any data set</i>)	<ul style="list-style-type: none"> Remotely accessible Data available for rate, ratio, and trend analysis only 	<ul style="list-style-type: none"> Remotely accessible NAACCR members only Active consent from individual registries needed for access to area variables (e.g. county) 	<ul style="list-style-type: none"> Remotely accessible 	<ul style="list-style-type: none"> Remotely accessible Data available for rate, ratio, and trend analysis only 	<ul style="list-style-type: none"> Accessible through federal data centers only Proposal approval needed
Source Registries	NPCR; SEER	NPCR; SEER	SEER	NPCR; SEER	NPCR; SEER
US Population Represented	<ul style="list-style-type: none"> Up to 93% Registries who meet data quality standards for given year 	<ul style="list-style-type: none"> Up to 100%; coverage variable by year Registries who meet data quality standards 	<ul style="list-style-type: none"> 28% (SEER 18) 	<ul style="list-style-type: none"> 96% 	<ul style="list-style-type: none"> 100%
Geographic Scale	<ul style="list-style-type: none"> State/registry County level measures of rurality and poverty Specific counties are not identifiable 	<ul style="list-style-type: none"> County Census tract level rural-urban metric available with approval 	<ul style="list-style-type: none"> State/registry County 	<ul style="list-style-type: none"> State 	<ul style="list-style-type: none"> County (restricted) States may suppress county-specific data
Demographic Variables Available (<i>in addition to sex, race, Hispanic ethnicity, age group</i>)	<ul style="list-style-type: none"> No additional 	<ul style="list-style-type: none"> Insurance status Birth country/state 	<ul style="list-style-type: none"> Insurance status Marital status Specific age 	<ul style="list-style-type: none"> No additional 	<ul style="list-style-type: none"> No additional
Tumor Characteristics Available (<i>in addition to site, histology, grade, laterality, tumor size</i>)	<ul style="list-style-type: none"> SEER summary stage groupings 	<ul style="list-style-type: none"> SEER and AJCC stage groupings Some site specific factors[†] 	<ul style="list-style-type: none"> SEER, AJCC, and TNM stage groupings 	<ul style="list-style-type: none"> SEER summary stage groupings 	<ul style="list-style-type: none"> SEER summary stage groupings

	North American Cancer Registries (NAACCR) Public Use Data Set ¹⁸	NAACCR Cancer in North America Deluxe File ¹⁹	Surveillance Epidemiology and End Results (SEER) Dataset ²⁰	National Program of Cancer Registries (NPCR) Public Use Dataset ²¹	United States Cancer Statistics Restricted Access Dataset ²²
			<ul style="list-style-type: none"> Some site specific factors [†] 		<ul style="list-style-type: none"> Some site specific factors [†]
Treatment Variables Available	None	<ul style="list-style-type: none"> Course of first treatment Surgery code Reason for no surgery Extent of lymph node examination Extent of nodes positive Receipt of radiation, chemotherapy, hormone therapy 	<ul style="list-style-type: none"> Course of first treatment Surgery code Reason for no surgery Extent of lymph node examination Extent of nodes positive 	None	None
Outcome Variables Available	None	<ul style="list-style-type: none"> Vital status Cause of death Survival months 	<ul style="list-style-type: none"> Vital status Cause of death Survival months 	None	None
Strengths	<ul style="list-style-type: none"> Dichotomous rural-urban variable; 	<ul style="list-style-type: none"> Centralization of data request Data can be exported for analyses using non-SEER*Stat software Treatment variables are accessible Multiple rural-urban variables may be available 	<ul style="list-style-type: none"> County scale allows for utilization of rural-urban measures Underserved populations are oversampled Data can be exported for analyses using non-SEER*Stat software Includes extensive data on cancer incidence and survival Historical coverage from 	<ul style="list-style-type: none"> Large population represented 	<ul style="list-style-type: none"> Potential to evaluate cancer incidence and staging across entire country

	North American Cancer Registries (NAACCR) Public Use Data Set ¹⁸	NAACCR Cancer in North America Deluxe File ¹⁹	Surveillance Epidemiology and End Results (SEER) Dataset ²⁰	National Program of Cancer Registries (NPCR) Public Use Dataset ²¹	United States Cancer Statistics Restricted Access Dataset ²²
			<ul style="list-style-type: none"> as far back as 1973 for some registries For a fee, SEER data linked to Medicare and Medicare-associated survey data are available from NCI 		
Limitations	<ul style="list-style-type: none"> Currently only available for rate, rate ratio, and trend calculation No treatment or survival data available 	<ul style="list-style-type: none"> State-specific approvals needed for county level release and census tract level rural-urban measurement variable 	<ul style="list-style-type: none"> Few treatment variables available Less generalizable to the entire population 	<ul style="list-style-type: none"> State geographic scale severely limits exploration of rural cancer control No treatment or survival data Only available for rate, rate ratio, and trend calculation 	<ul style="list-style-type: none"> No treatment data available Data only available at approved research data center Rural analysis may not be possible without state's approval

IRB=Institutional Review Board; DUA=Data Use Agreement; AJCC=American Joint Commission on Cancer; TNM=Tumor, Number of Lymph Nodes, Metastasis; NCHS=National Center for Health Statistics;

[†] Site-specific factors are cancer site specific variables that provide additional, clinically relevant information about the cancer tumor beyond stage, grade, or histology

Table 2
Rural and Urban Cancer Incidence Rates in NAACCR and SEER 18, 2009-2013, by cancer site

	NAACCR ^a		SEER 18	
	Rural IR (95% CI)	Urban IR (95% CI)	Rural IR (95% CI)	Urban IR (95% CI)
All cancers	446.4 (445.6-447.2) ^b	448.7 (448.3-449.1)	461.8 (459.9-463.7) ^b	449.1 (448.4-449.7)
Oral cavity and pharynx	12.2 (12.1-12.4) ^{b,c}	11.2 (11.1-11.3)	12.8 (12.4-13.1) ^{b,c}	10.9 (10.8-11.0)
Esophagus	5.1 (5.0-5.2) ^{b,c}	4.6 (4.5-4.6)	4.9 (4.6-5.1) ^{b,c}	4.2 (4.2-4.3)
Stomach	5.8 (5.7-5.9) ^{b,c}	6.9 (6.8-6.9)	6.2 (6.0-6.4) ^{b,c}	7.6 (7.5-7.6)
Colon and rectum	43.9 (43.7-44.2) ^{b,c}	40.1 (39.9-40.2)	46.0 (45.4-46.6) ^{b,c}	40.4 (40.2-40.6)
Liver and intrahepatic duct	6.3 (6.2-6.4) ^{b,c}	7.9 (7.9-8.0)	6.8 (6.6-7.0) ^{b,c}	8.7 (8.6-8.8)
Pancreas	11.9 (11.8-12.1) ^{b,c}	12.5 (12.4-12.5)	12.3 (12.0-12.6)	12.5 (12.0-12.6)
Larynx	4.4 (4.3-4.4) ^{b,c}	3.4 (3.4-3.4)	4.7 (4.5-4.9) ^{b,c}	3.0 (2.9-3.0)
Lung and bronchus	70.2 (69.9-70.6) ^{b,c}	61.2 (61.0-61.3)	72.2 (71.5-73.0) ^{b,c}	55.6 (55.3-55.8)
Melanoma	19.9 (19.8-20.1) ^b	20.1 (20.1-20.2)	22.5 (22.0-22.9) ^b	21.8 (21.7-22.0)
Breast (female)	113.4 (112.9-114.0) ^{b,c}	124.8 (124.5-125.1)	115.2 (113.8-116.5) ^{b,c}	126.4 (125.9-126.9)
Cervix uteri	8.4 (8.2-8.5) ^{b,c}	7.6 (7.5-7.6)	8.5 (8.1-8.9) ^{b,c}	7.4 (7.3-7.5)
Corpus and uterus, NOS	25.0 (24.7-25.2) ^{b,c}	25.7 (25.5-25.8)	24.6 (24.0-25.2) ^b	25.6 (25.4-25.8)
Ovary	11.3 (11.1-11.5) ^{b,c}	11.7 (11.6-11.8)	11.4 (11.0-11.8) ^{b,c}	12.0 (11.8-12.1)
Prostate	114.1 (113.5-114.7) ^{b,c}	124.5 (124.3-124.8)	117.8 (116.5-119.2) ^{b,c}	131.9 (131.3-132.4)
Testis	5.5 (5.3-5.6)	5.5 (5.4-5.5)	5.7 (5.4-6.1) ^{b,c}	5.7 (5.5-5.8)
Urinary Bladder	20.9 (20.7-21.1) ^b	20.6 (20.5-20.7)	21.0 (20.6-21.4) ^b	20.0 (19.9-20.2)
Kidney and renal pelvis	16.9 (16.7-17.0) ^{b,c}	15.9 (15.8-16.0)	17.3 (16.9-17.7) ^{b,c}	15.5 (15.4-15.6)
Brain and other nervous system	6.7 (6.6-6.8) ^b	6.6 (6.5-6.6)	6.8 (6.5-7.0) ^{b,c}	6.4 (6.3-6.5)
Thyroid	12.0 (11.8-12.1) ^{b,c}	14.4 (14.3-14.4)	12.4 (12.0-12.7) ^{b,c}	14.1 (14.0-14.2)
Hodgkin lymphoma	2.6 (2.6-2.7)	2.7 (2.7-2.7)	2.6 (2.5-2.8)	2.6 (2.6-2.7)
Non-Hodgkin lymphoma	18.2 (18.1-18.4) ^{b,c}	19.2 (19.1-19.3)	18.9 (18.6-19.3) ^b	19.7 (19.5-19.8)
Myeloma	5.9 (5.8-6.0) ^{b,c}	6.5 (6.4-6.5)	6.4 (6.2-6.7) ^b	6.6 (6.5-6.7)
Leukemia	13.5 (13.4-13.6) ^{b,c}	13.3 (13.3-13.3)	14.6 (14.2-14.9) ^{b,c}	13.6 (13.5-13.7)

IR=incidence rate; 95% CI= 95% confidence intervals;

^aRates are expressed per 100,000 population and age-adjusted to the 2000 US standard population.

^bThe rate ratio indicates that the rural rate is statistically significantly different than the urban rate (p<0.05).

^cThe rate ratio indicates that the rural rate is statistically significantly different than the urban rate and differs by at least 5% (p<0.05).

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Table 3
Rural-Urban Cancer Incidence Rates in NAACCR and SEER 18 by Race/Ethnicity, 2009-2013

	Non-Hispanic White						Non-Hispanic Black						Hispanic					
	NAACCR			SEER 18			NAACCR			SEER 18			NAACCR			SEER 18		
	Rural IR (95% CI)	Urban IR (95% CI)		Rural IR (95% CI)	Urban IR (95% CI)		Rural IR (95% CI)	Urban IR (95% CI)		Rural IR (95% CI)	Urban IR (95% CI)		Rural IR (95% CI)	Urban IR (95% CI)		Rural IR (95% CI)	Urban IR (95% CI)	
All cancers	449.5 (448.6-450.4) b	464.4 (463.9-464.8)		472.6 (470.4-474.8) b	480.6 (479.8-481.5)		458.7 (455.5-461.9) b	471.4 (470.3-472.6)		468.0 (460.9-475.1) b	485.7 (483.4-487.9)		322.7 (319.0-326.5) b,c	350.9 (349.9-351.9)		331.1 (323.7-338.6) b,c	347.6 (346.0-349.3)	
Oral cavity and pharynx	12.5 (12.3-12.7) b	12.2 (12.2-12.3)		13.3 (12.9-13.6) b,c	12.5 (12.4-12.6)		11.3 (10.8-11.8) b,c	9.3 (9.2-9.5)		11.1 (10.1-12.2) b,c	9.3 (9.0-9.6)		6.6 (6.1-7.1) b,c	7.2 (7.0-7.3)		7.1 (6.1-8.3)	6.7 (6.5-6.9)	
Esophagus	5.1 (5.0-5.2) b	5.0 (4.9-5.0)		4.9 (4.7-5.1)	4.8 (4.7-4.9)		5.8 (5.5-6.2) b,c	4.5 (4.4-4.6)		6.6 (5.8-7.5) b,c	4.4 (4.2-4.6)		2.9 (2.6-3.4)	2.8 (2.7-2.9)		2.8 (2.1-3.6)	2.8 (2.6-2.9)	
Stomach	5.2 (5.1-5.3) b,c	5.5 (5.5-5.6)		5.5 (5.3-5.7) b,c	5.8 (5.8-5.9)		10.4 (9.9-10.9)	10.8 (10.6-10.9)		10.7 (9.6-11.8)	10.8 (10.5-11.2)		8.9 (8.3-9.6) b,c	10.2 (10.0-10.4)		8.4 (7.2-9.7) b,c	10.8 (10.5-11.1)	
Colon and rectum	43.2 (43.0-43.5) b,c	39.5 (39.4-39.6)		45.7 (45.0-46.3) b,c	40.3 (40.1-40.6)		53.8 (52.7-54.9) b,c	48.8 (48.4-49.2)		56.5 (54.0-59.0) b,c	51.6 (50.9-52.3)		36.6 (35.3-37.9) b,c	25.6 (35.2-35.9)		37.8 (35.3-40.5) b,c	34.7 (34.2-35.2)	
Liver and intrahepatic duct	5.8 (5.7-5.9) b,c	6.4 (6.3-6.4)		6.2 (6.0-6.4) b,c	6.7 (6.6-6.8)		6.8 (6.4-7.1) b,c	10.6 (10.5-10.7)		6.3 (5.8-7.4) b,c	10.7 (10.4-11.0)		13.4 (12.6-14.2)	13.0 (12.8-13.2)		13.3 (11.9-14.9)	13.0 (12.7-13.4)	
Pancreas	11.8 (11.6-11.9) b	12.4 (12.3-12.4)		12.3 (12.0-12.6)	12.6 (12.5-12.8)		15.1 (14.5-15.7) b	15.8 (15.6-16.0)		14.0 (12.8-15.3) b,c	16.1 (15.7-16.5)		10.7 (10.0-11.5)	11.2 (11.0-11.4)		10.4 (9.1-11.9)	11.2 (10.9-11.5)	
Larynx	4.3 (4.3-4.4) b,c	3.5 (3.4-3.6)		4.8 (4.6-5.0) b,c	3.2 (3.1-3.3)		6.1 (5.8-6.5) b,c	4.6 (4.5-4.7)		6.3 (5.5-7.1) b,c	4.6 (4.4-4.8)		2.6 (2.2-2.9)	2.6 (2.5-2.7)		2.8 (2.1-3.6)	2.2 (2.1-2.3)	
Lung and bronchus	71.7 (71.3-72.0) b,c	65.4 (65.3-65.3)		75.5 (74.6-76.3) b,c	61.5 (61.2-61.8)		70.1 (68.9-71.4) b,c	67.0 (66.5-67.4)		73.3 (70.5-76.2) b,c	66.2 (65.4-67.1)		36.7 (35.4-38.0) b,c	32.3 (31.9-32.6)		36.8 (34.3-39.5) b,c	29.6 (29.1-30.1)	
Melanoma	22.1 (21.9-22.3) b,c	26.2 (26.1-26.3)		25.9 (25.4-26.4) b,c	31.7 (31.4-31.9)		1.3 (1.1-1.5) b,c	1.0 (0.9-1.1)		1.5 (1.1-2.0) b,c	1.0 (0.9-1.1)		5.5 (5.0-6.0) b,c	4.5 (4.3-4.6)		4.8 (3.9-5.7)	4.5 (4.3-4.7)	
Breast (female)	114.3 (113.6-114.9) b,c	131.4 (131.1-131.7)		116.7 (115.2-118.2) b,c	137.3 (136.7-138.0)		118.2 (116.0-120.4) b,c	125.5 (125.8-126.3)		116.1 (111.4-121.0) b,c	129.8 (128.3-131.3)		81.6 (79.1-84.2) b,c	92.5 (91.8-93.1)		93.0 (91.9-94.1)	91.3 (86.1-96.8)	
Cervix uteri	8.0 (7.8-8.2) b,c	6.9 (6.8-6.9)		8.3 (7.9-8.8) b,c	6.8 (6.6-6.9)		10.8 (10.1-11.5) b,c	9.8 (9.5-10.0)		10.7 (9.3-12.3) b,c	9.0 (8.6-9.4)		9.4 (8.6-10.2) b	9.8 (9.7-10.1)		6.6 (5.4-8.1) b,c	9.5 (9.2-9.9)	
Corpus and uterus, NOS	25.3 (25.0-25.6) b	26.5 (26.4-26.7)		24.9 (24.2-25.6) b,c	26.9 (26.6-27.2)		23.3 (22.4-24.3) b,c	25.2 (24.8-25.5)		23.0 (20.9-25.1) b,c	25.2 (24.6-25.9)		18.9 (17.8-20.2) b,c	21.8 (21.5-22.2)		17.9 (15.6-20.4)	21.7 (21.1-22.2)	
Ovary	11.5 (11.3-11.7) b,c	12.3 (12.2-12.4)		11.6 (11.2-12.1) b,c	12.9 (12.7-13.1)		9.3 (8.7-9.9)	9.5 (8.6-9.9)		9.7 (8.4-11.2)	9.9 (9.5-10.4)		10.5 (9.6-11.5)	10.3 (10.0-10.5)		10.9 (9.1-12.9)	10.6 (10.3-11.0)	
Prostate	108.1 (107.5-108.7) b,c	116.2 (115.8-116.5)		113.9 (112.4-115.3) b,c	128.4 (127.8-129.1)		190.5 (187.4-193.6) b	199.0 (197.9-200.2)		196.8 (189.7-204.1) b,c	212.2 (209.9-214.5)		78.1 (75.3-80.9) b,c	106.7 (105.8-107.6)		84.5 (79.1-90.2) b,c	109.0 (107.6-110.5)	
Testis	6.2 (6.0-6.4) b,c	6.9 (6.8-7.0)		6.5 (6.1-7.0) b,c	7.4 (7.2-7.6)		1.4 (1.1-1.6)	1.4 (1.3-1.5)		1.1 (0.7-1.7)	1.5 (1.4-1.7)		3.8 (3.4-4.2) b,c	4.6 (4.4-4.7)		4.8 (3.9-5.9)	5.0 (4.8-5.2)	
Urinary bladder	22.1 (21.9-22.3) b	23.3 (23.2-23.4)		22.6 (22.2-23.1) b,c	23.9 (23.7-24.1)		10.5 (10.0-11.0)	12.1 (11.9-12.3)		11.3 (10.1-12.5) b,c	13.1 (12.7-13.5)		10.3 (9.6-11.0) b,c	11.4 (11.2-11.6)		11.2 (9.8-12.7)	11.2 (10.9-11.5)	
Kidney and renal pelvis	16.8 (16.6-16.9) b,c	16.1 (16.1-16.2)		17.6 (17.2-18.0) b,c	16.1 (15.9-16.2)		17.5 (16.9-18.2)	18.0 (17.7-18.2)		17.3 (16.0-18.7) b,c	18.8 (18.4-19.2)		16.6 (15.7-17.4) b	15.8 (15.6-16.0)		14.7 (13.2-16.4)	15.5 (15.2-15.9)	
Brain and other nervous system	7.2 (7.0-7.3) b	7.5 (7.4-7.5)		7.4 (7.1-7.7)	7.7 (7.6-7.8)		3.8 (3.5-4.1)	4.2 (4.1-4.3)		4.2 (3.6-4.9)	4.3 (4.1-4.5)		4.7 (4.3-5.1) b,c	5.2 (5.1-5.3)		5.0 (4.2-5.9)	5.0 (4.8-5.1)	
Thyroid	12.8 (12.6-12.9) b,c	15.7 (15.6-15.8)		13.3 (12.9-13.7) b,c	15.9 (15.8-16.1)		6.7 (6.3-7.1) b,c	9.2 (9.0-9.3)		6.5 (5.7-7.4) b,c	8.7 (8.5-9.0)		9.9 (8.5-9.6) b,c	12.8 (12.6-13.0)		10.5 (9.4-11.8) b,c	12.0 (11.7-12.3)	
Hodgkin lymphoma	2.8 (2.7-2.9)	3.0 (2.9-3.0)		2.8 (2.6-3.0) b,c	3.1 (3.0-3.1)		2.4 (2.2-2.6) b,c	2.7 (2.6-2.8)		2.7 (2.2-3.2)	2.7 (2.6-2.9)		2.0 (1.7-2.2) b,c	2.4 (2.3-2.5)		1.8 (1.3-2.3)	2.2 (2.1-2.4)	
Non-Hodgkin lymphoma	18.7 (18.6-18.9) b,c	20.1 (20.0-20.2)		19.8 (19.4-20.3) b,c	21.1 (21.0-21.3)		12.5 (11.9-13.0) b,c	14.6 (14.4-14.8)		12.7 (11.5-13.9) b,c	15.2 (14.8-16.0)		14.9 (14.1-15.7) b,c	17.4 (17.2-17.6)		14.0 (12.5-15.6) b,c	18.0 (17.7-18.4)	

	Non-Hispanic White						Non-Hispanic Black						Hispanic			
	NAACCR			SEER 18			NAACCR			SEER 18			NAACCR		SEER 18	
	Rural IR (95% CI)	Urban IR (95% CI)	Rural IR (95% CI)	Urban IR (95% CI)	Rural IR (95% CI)	Urban IR (95% CI)	Rural IR (95% CI)	Urban IR (95% CI)	Rural IR (95% CI)	Urban IR (95% CI)	Rural IR (95% CI)	Urban IR (95% CI)	Rural IR (95% CI)	Urban IR (95% CI)	Rural IR (95% CI)	Urban IR (95% CI)
Myeloma	5.5 (5.4-5.6) <i>b</i>	5.7 (5.6-5.7)	5.9 (5.6-6.1) <i>b,c</i>	6.0 (5.9-6.1)			12.3 (11.8-12.8) <i>b,c</i>	13.2 (13.0-13.4)	13.9 (12.7-15.2)	13.7 (13.3-14.1)	5.6 (5.1-6.1) <i>b,c</i>	6.3 (6.2-6.5)	6.0 (5.0-7.1)	6.0 (5.9-6.3)		
Leukemia	13.8 (13.6-13.9) <i>b</i>	14.0 (13.9-14.1)	15.1 (14.7-15.5)	14.8 (14.6-14.9)			10.1 (9.6-10.6) <i>b</i>	10.5 (10.4-10.7)	10.5 (9.5-11.7)	11.2 (10.9-11.6)	10.0 (9.4-10.7)	10.7 (10.6-10.9)	11.2 (9.9-12.6)	10.7 (10.4-10.9)		

IR=incidence rate; CI=confidence intervals;

^aRates are expressed per 100,000 population and age-adjusted to the 2000 US standard population.

^bThe rate ratio indicates that the rural rate is statistically significantly different than the urban rate (p<0.05).

^cThe rate ratio indicates that the rural rate is statistically significantly different than the urban rate and differs by at least 5% (p<0.05).

Table 4
Rural Sociodemographic Representation of SEER 18 and the NAACCR Public Use Dataset Compared to the United States Population

	U.S. Population N (%)	SEER 18 Counties* N (%)	NAACCR Counties N(%)
Total Population	46,251,252 (14.8%)	8,290,215 (10.6%)	43,488,697 (14.6%)
Race			
White	39,219,179 (84.8%)	6,719,792 (81.1%)	37,091,185 (84.4%)
Black	3,826,447 (8.3%)	821,617 (9.9%)	3,774,972 (8.6%)
Asian	390,535 (0.8%)	132,892 (1.6%)	369,080 (0.8%)
American Indian/Alaska Native	942,923 (2.0%)	153,626 (1.9%)	909,256 (2.1%)
Hawaiian/Other Pacific Islander	56,144 (0.1%)	36,852 (0.4%)	54,892 (0.1%)
Other	815,494 (1.8%)	203,881 (2.5%)	789,147 (1.8%)
Two or More Races	1,000,530 (2.2%)	221,558 (2.7%)	954,183 (2.2%)
Ethnicity			
Hispanic	3,545,281 (7.7%)	753,267 (9.1%)	3,414,228 (7.8%)
Non-Hispanic	42,705,971 (92.3%)	7,536,948 (90.9%)	40,528,487 (92.2%)
% Living in Poverty	8,036,682 (18.1%)	1,571,878 (19.7%)	7,742,685 (18.3%)
U.S. Census Region			
Northeast	4,668,382 (10.1%)	188,629 (2.3%)	4,255,118 (9.8%)
Midwest	15,213,564 (32.9%)	1,395,475 (16.8%)	13,471,782 (31.0%)
South	19,793,868 (42.8%)	4,386,393 (52.9%)	18,728,972 (43.1%)
West	6,575,438 (14.2%)	2,419,718 (29.2%)	6,270,453 (14.4%)

* One SEER 18 registry is not defined solely geographically (i.e. the Alaska Native registry in Alaska). Characterization of that registry is not represented by this table.