Rural-Urban Differences in Late-Stage Breast Cancer: Do Associations Differ by Rural-Urban Classification System?

Sandi L Pruitt, Ph.D[1,2,*], **Jan M Eberth, Ph.D**[3,4], **E Scott Morris, MS**[1,5], **David B Grinsfelder, MPH**[1], and **Erica L Cuate, MPH**[1]

[1]Department of Clinical Sciences, University of Texas Southwestern Medical Center, Dallas, TX U.S.A

[2]Harold C. Simmons Comprehensive Cancer Center, Dallas, TX U.S.A

[3]Department of Epidemiology and Biostatistics, Arnold School of Public Health, University of South Carolina, Columbia, SC U.S.A

[4]Cancer Prevention and Control Program, University of South Carolina, Columbia, SC U.S.A

[5]School of Economic, Political, and Policy Sciences, University of Texas Dallas, Dallas TX U.S.A

Abstract

Introduction—Rural residence is associated with later stage of breast cancer diagnosis in some but not all prior studies. The lack of a standardized definition of rural residence may contribute to these mixed findings. We characterize and compare multiple definitions of rural vs. non-rural residence to provide guidance regarding choice of measures and to further elucidate rural disparities in breast cancer stage at diagnosis.

Methods—We used Texas Cancer Registry data of 120,738 female breast cancer patients 50 years old diagnosed between 1995–2009. We defined rural vs. non-rural residence using 7 different measures and examined their agreement using Kappa statistics. Measures were defined at various geographic levels: county, ZIP code, census tract, and census block group. Late-stage was defined as regional or distant disease. For each measure, we tested the association of rural residence and late-stage cancer with unadjusted and adjusted logistic regression. Covariates included: age; patient race/ethnicity; diagnosis year; census block group-level mammography capacity; and census tract-level percent poverty, percent Hispanic, and percent Black.

Results—We found moderate to high levels of agreement between measures of rural vs. non-rural residence. For 72.9% of all patients, all 7 definitions agreed as to rural vs. non-rural residence. Overall, 6 of 7 definitions demonstrated an adverse association between rural residence and late-stage disease in unadjusted and adjusted models (Adjusted OR Range = 1.09–1.14).

Discussion—Our results document a clear rural disadvantage in late-stage breast cancer. We contribute to the heterogeneous literature by comparing varied measures of rural residence. We recommend use of the census tract-level Rural Urban Commuting Area Codes in future cancer outcomes research where small area data are available.

*(Address for correspondence) University of Texas Southwestern Medical Center, 5323 Harry Hines Blvd E1.410D Dallas, TX 75390-9169, Phone: 01-214-648-3257 Fax: 01-214-648-3232 ; Email: Sandi.Pruitt@utsouthwestern.edu

Keywords

breast cancer; rural; health disparities; measurement; cancer stage

INTRODUCTION

Among U.S. women, breast cancer is the leading cause of cancer and the second leading cause of cancer deaths [1]. Stage at diagnosis is the most significant prognostic factor for survival. Five-year relative survival is 98% for localized disease but falls to only 24% for distant disease [1].

Living in a rural area has been associated with later stage of breast cancer diagnosis in some but not all prior studies [2,3]. It is not known whether the mixed findings to date result from varying definitions of urban vs. rural or from other aspects of study design (e.g. different populations, covariates, or model specifications). A recent literature review examining rural residence and cancer outcomes found wide heterogeneity in the use of definitions of rural vs. non-rural residence and describes the literature as "nascent and methodologically inconsistent." (p. 1657) [2]. Another recent study compared breast cancer stage at diagnosis between women living in urban vs. rural areas. This systematic review and meta-analysis of 21 studies, despite documenting mixed findings in the extant literature, concluded that patients living in rural areas were more likely to be diagnosed with late-stage breast cancer [3]. This review also noted wide variation in definitions of "rural," [3] as have previous reviews [4].

In the cancer literature, researchers typically rely upon methods developed by government agencies for defining "rural" residence [2]. These methods are relatively simple to calculate and use publicly available data. But to date, there are no methodologic studies available that compare and contrast available methods. Our study aims to fill these gaps in the literature. In this study, we characterize and compare multiple definitions of rural vs. non-rural residence in a study of breast cancer patients in Texas to 1) provide guidance to cancer outcomes researchers regarding choice of measures and to 2) further elucidate rural-urban disparities in breast cancer stage at diagnosis.

METHODS

Data and Sample

Data were obtained from the Texas Cancer Registry (TCR), a North American Association of Central Cancer Registries (NAACCR) gold-certified population-based cancer registry. Female adult (50 years old) breast cancer patients diagnosed between 1995–2009 were eligible for inclusion. Data were limited to women aged 50 and older because breast cancer screening, a primary determinant of late-stage cancer diagnosis, is recommended for women aged 50 years and older by the U.S. Preventive Services Task Force [5]. Data were limited to patients with a first primary breast cancer and patients with known cancer stage. Data were geocoded by TCR. To ensure accuracy, we limited inclusion to patients with addresses geocoded at the street address level, representing the highest degree of geocoding certainty. Using latitude and longitude of patient residence at diagnosis, we spatially joined the

patients in the sample to their respective counties, ZIP codes, census tracts, and block groups. Then, we merged in 2000 U.S. Census data as provided by Geolytics Inc. (Census 2000, Long Form [SF3], GeoLytics, Inc., East Brunswick, NJ, 2012.) including population values at the county, ZIP code, census tract, and block group levels. We obtained data on location and number of mammography machines in the year 2000 obtained from a Freedom of Information Act request to the Food and Drug Administration (requested July 2008, received March 2009). This study was approved by UT Southwestern Medical Center, Texas Cancer Registry, and the Texas Department of Health and Social Services Institutional Review Boards.

Urban-Rural Classification Methods

We compared seven different urban/rural measures, defined at block group, census tract, ZIP code, and county levels. Attributes of each measure are provided in Table 1 and described briefly below. All measures were developed by governmental agencies based on one of multiple taxonomies. The underlying taxonomies incorporate factors such as population size, density, urbanization, and commuting patterns. Measures were downloaded from governmental agency websites (see references in Table 1) in pre-specified formats. Therefore, as succinctly described in a recent review [2], many of these measures were developed by transforming underlying, granular block-level classifications (the smallest geospatial unit in Census data) to larger units. Thus, for example, an urban census tract may actually include some rural blocks. Some degree of this problem, known as *overbounding* or *underbounding* is inherent to any definition of rurality [6]. All measures except the block group measure were already classified at their representative geospatial unit (e.g. census tract). We classified block groups using the location of the block group centroid. Thus, for example, all block groups with a centroid within an urbanized area (according to U.S. Census classification, UA/UC shapefile) were considered as such. Also of note, ZIP Code configurations change frequently, do not represent Census area spatial boundaries, and do not necessarily represent polygons [7–9]. Thus, to convert ZIP codes to spatial units, both ZIP code measures (FAR and zip-RUCA) applied different source mapping data and methods, as described elsewhere [10,11]. Comprehensive reviews of the intricacies of measurement (e.g. definitions underlying each method, the difference between non-metropolitan vs. rural) are available elsewhere [2,6,12–14].

For each measure, we applied multiple categorizations of urban-rural residence identified from measure documentation [10,11,15–20] and previous literature [21]. Table 1 denotes the binary (rural vs. non-rural) categorization for each measure and any alternative categorizations (e.g. large metropolitan, small metropolitan). Alternative categorizations are further described in Appendix 1.

Outcome and Covariates

The outcome was late-stage cancer defined using the Surveillance Epidemiology and End Result (SEER) summary stage variable as: late (regional or distant) vs. early (in situ or localized).

We included several covariates in our adjusted models in order to control for confounding factors found to be associated with urban/rural residence and/or stage at diagnosis in previous studies [3,22,23]. Patient-level covariates included: race (non-Hispanic white, non-Hispanic black, Hispanic, other, and unknown); age (50–59, 60–69, 70–79, 80); and diagnosis year (1995–97, 1998-00, 2001–03, 2004–06, 2007–09). Neighborhood-level covariates included: census tract-level percent poverty, percent black, and percent Hispanic; and block group-level mammography capacity.

We measured mammography capacity using the two-step floating catchment area (2SFCA) method as first described by Luo et al. [24] and later applied to mammography by Eberth et al. [25] In essence, this measure accounts for both the supply (i.e., number of available mammography machines) and demand (i.e., number of women 50 years based on 2000 U.S. census data) for mammography services at the block group level. Step 1 of the 2SFCA is a facility specific machine-to-population ratio and step 2 is a block group specific ratio that sums over all the facilities that fall within 60 minutes from the population-weighted block group centroid. The resulting spatial accessibility score is then categorized into 3 levels of capacity (poor, adequate, and excess) based on the expected number of machines needed to meet the biennial screening goal of 81% set by HealthyPeople 2020 [26].

Analysis

We first describe characteristics of breast cancer patients in our sample. We compared the distribution of covariates by rural vs. non-rural residence using the county-based definitions using chi-square and t-test statistics. We describe covariates using the county-based definitions rather than comparing covariates across all 7 measures to simplify our presentation of data.

We next compared the 7 measures of rural residence. Using descriptive statistics and maps, we examined the classification of Texas geographies and total population defined as rural or non-rural using each measure. We generated four maps, one each for: county-level measures (n=3), ZIP-code measures (n=2), and census tract (n=1) and block group (n=1) measures. For each urban/rural measure, we compared the distribution of late vs. early stage cancers by rural vs. non-rural residence using chi-square statistics.

Among breast cancer cases, we assessed agreement between each binary urban/rural measure using percent agreement, kappa, and prevalence- and bias-adjusted kappa (PABAK). Kappa adjusts for the amount of agreement expected to occur by chance alone [27]. To correct for the tendency of kappa to be highly dependent on the prevalence of the condition in the population, we also report PABAK [28]. PABAK assumes fifty percent prevalence of the condition and absence of any bias, thereby reflecting an ideal situation and ignoring the prevalence and bias present in the "real world". We judge the adequacy of kappa and PABAK following guidelines suggested by Landis and Koch: poor: 0–0.20; fair: 0.21–0.40; moderate: 0.41–0.60; substantial: 0.61–0.80; and nearly perfect agreement: 0.81 to 1.0 [29].

Finally, to measure the association of rural residence and late-stage disease, we fitted a series of unadjusted and adjusted logistic regression models. Adjusted models controlled for all

covariates described above. We fitted binary measures of rural vs. non-rural residence and categorical measures as described in Table 1. For the categorical analyses of each measure we set the most urban category as the referent. In sensitivity analyses, to adjust for the non-independence of observations within each "level" (e.g. census tract, county), we added random intercepts to each model defined at the block group, census tract, ZIP code, and county-level, as appropriate. For example, for the ZIP-code FAR measure, models included a random intercept at the ZIP code level. All analyses were conducted in STATA Version 13.0 (StataCorp. 2013. *Stata Statistical Software: Release 13*. College Station, TX: StataCorp LP). Maps were created in ArcGIS Version 10.1 (ESRI 2011. ArcGIS Desktop: Release 10.1 Redlands, CA: Environmental Systems Research Institute).

RESULTS

In all, there were 120,738 breast cancer patients included in this study. Comparison of patients by rural vs. non-rural residence (as defined by county-based methods of UI, NCHS, and RUCC) are provided in Table 2. Distribution of all covariates differed significantly between rural and non-rural patients. Notably, rural patients were more likely to be white, live in areas of poor mammography capacity and resided in neighborhoods where a greater percent of the population lived in poverty. In additional analyses [not shown], we compared covariates across all 7 measures and the results of this analysis were comparable.

We examined the classification of Texas geographies and total population of Texas using each of the 7 rural vs. non-rural definitions (Appendix 2). For all definitions, a greater percent of total area in miles was considered rural (ranging from 53.5–97.3%, depending on definition). However a greater percent of the population lived in non-rural areas (ranging from 81.6–94.5% in non-rural and 5.5–18.4% in rural areas depending on definition).

Of all breast cancer patients, 88% (n=106,706) were diagnosed with early stage disease whereas 11.6% (n=14,032) were diagnosed with late disease (Appendix 2). The number of patients defined as residing in a rural area varied considerably by definition. For example, the percent residing in rural areas ranged from 3.9% (zip code FAR) to 16.6% (tract RUCA). With the exception of the block group (UA/UC) definition, for all other definitions, rural residence (vs. non-rural) was significantly associated with a greater number of late- (vs. early) stage breast cancer cases (p<.05).

Figure 1 illustrates rural vs. non-rural areas as defined by measure. All 3 county-based definitions of rural vs. non-rural status agreed (1a). The two zip code definitions varied considerably, with a greater number of zip codes defined as urban when using the FAR vs. the zip-code RUCA approach (1b). There was only 1 census tract-based definition (1c). For block groups, those classified as within an urbanized area or urban cluster are considered non-rural in the binary comparison; but for comparison purposes, we differentiate these categories in the map (1d).

Overall, we found moderate to high levels of agreement between varying definitions of rural vs. non-rural residence among breast cancer cases (Table 3). As previously noted, all county-based definitions demonstrated perfect agreement. In all, for 72.9% of all patients, all 7

definitions agreed as to rural vs. non-rural residence. Percent agreement between all other comparisons ranged from 78.2–95.7%. PABAKs indicate nearly perfect agreement for 4 of the comparisons, substantial agreement for 5 comparisons, and moderate agreement for one comparison: tract-RUCA vs. block group.

Lastly, we examined the association of rural residence and late-stage disease using unadjusted and adjusted logistic regression. Overall, binary measures (rural vs. non-rural) indicate a generally consistent and moderate association between rural residence and late-stage breast cancer diagnosis (Table 4, Method 1). For all definitions with the exception of block group, unadjusted odds ratios of rural vs. non-rural residence were significant and ranged from 1.18 (tract-RUCA) to 1.21 (county definitions). In fully adjusted models, this adverse association is modestly attenuated but odds ratios remain significantly associated with late-stage diagnosis and range from 1.09 (tract-RUCA) to 1.14 (County definitions). Notably, rural residence as defined by the binary block group definition was not associated with late-stage diagnosis in either unadjusted or adjusted models.

Associations using various categorizations of urban/rural status are provided in Table 4 (Methods 2 and 3). Overall, findings suggest that compared to those in the most urban areas, those living in more rural areas are more likely to be diagnosed at late stage. However, associations were not always linear across categorizations; and not all categories were significant. Notably, when examining a secondary classification of block group residence, differences in odds of late-stage disease emerged. Specifically, compared to patients living in an urbanized area (the most urban), those living urban clusters were more likely to be diagnosed at late stage in adjusted models (aOR: 1.14; 95% CI: 1.08–1.21). However, residence in a rural area was not significantly associated with late-stage disease.

In sensitivity analyses, we fitted all models with a random intercept (i.e. we fitted 2-level hierarchical models) at the relevant geographic level (county, zip-code, census tract, or block group) to adjust for hierarchical clustering. Random intercept models are presented in Appendix 3. We observed no changes in the direction of effects and only very minor changes in the magnitude and significance of effect sizes displayed in Table 4, likely reflecting the impact of underestimated standard errors in the original analysis. These changes do not substantively alter the study conclusions.

DISCUSSION

Comparisons of Urban/Rural Measures

Overall, our study found moderate to high levels of agreement between measures of rural vs. non-rural residence. For 72.9% of our breast cancer cases, all 7 definitions agreed as to rural vs. non-rural residence. In general, comparisons among larger units (e.g. county-zip) demonstrated higher agreement. Comparisons using block group measures demonstrated the lowest agreement. The tract and ZIP versions of RUCA also demonstrated high agreement.

Overall, 6 of 7 definitions demonstrated a positive and adverse association between rural residence and late-stage disease in unadjusted and adjusted models.While contradicting some prior U.S. studies,[30,31] our results confirm a multiple other U.S. studies documenting

rural disadvantage in late-stage breast cancer [23,32–34]. Indeed, our observed point estimates for a rural disadvantage in regional or late stage disease (adjusted OR ranged from 1.09–1.14) are very closely aligned to those observed in a recent meta-analysis. In a random effects model of 21 high-quality studies, rural (vs. urban) breast cancer patients had 1.19 higher odds (95% CI: 1.12–1.27) of late-stage and a 1.07 higher odds (95% CI: 1.04–1.10) of regional or late-stage breast cancer, respectively [3].

Few of the categorical measures demonstrated significant linear tests of trend across categories. While the majority of existing research has used binary measures of rural vs. non-rural, several previous studies have noted the lack of linear relationships across categories [35–38]. However, unlike a handful of prior studies from Illinois, we found no evidence of a "rural reversal."[11,1237] One of these studies found an "urban disadvantage" such that percent late-stage breast cancer was highest among Chicago residents, lower among those in Chicago suburbs, other metro areas, and large towns, but rates were also somewhat elevated among those living in small rural and isolated rural zip codes [36]. In this study, Chicago was compared to the remainder of Illinois, and findings to some degree may reflect the vulnerable and economically disadvantaged populations in inner-city Chicago. Regardless, such findings do raise some unanswered questions about the use of binary categories of urban vs. rural when studying geographic disparities in cancer outcomes.

Notably, only one binary measure of rural status was not associated with late-stage disease. The block group-based comparison demonstrated no difference in disease stage between rural vs. non-rural women. However, differences between block-group-based categories emerged when this measure was examined separately by urbanized area, urban cluster, and rural area. In this comparison, compared to urbanized areas (the most urban), those living in urban clusters, but not those in rural areas, were more likely to be diagnosed with late-stage disease (aOR: 1.14 95% CI: 1.08–1.21). The reasons for this are not clear. Compared to the other measures, the block group (BG) "rural" measure encompasses the largest amount of area and among the largest population (Appendix 2, Figure 1). This could result in more heterogeneity in this "rural" classification, resulting in the lack of an observed rural disparity. Just 12.6% of women with breast cancer in this study live in "urbanized clusters" (Appendix 2); many of whom (59%) are classified as "rural" using a county-based definition [data not shown]. While both are considered urban in our binary classification per standard practice, areas and clusters differ significantly in population density which ranges from 50,000 (urbanized areas) to 2,500–49,999 (urban clusters) residents per square mile [13]. When considering urban clusters as rural, rural (vs urban) women were more likely to be diagnosed with late-stage disease, consistent with findings using other measures.

The mechanisms driving rural disparity in breast cancer diagnosis stage are still uncertain. Possible hypothesized mechanisms include lack of insurance, poor spatial access to health care, including mammography access, lack of knowledge about screening guidelines, and lower use of mammography [33,39–43]. It will be important for future studies to examine the mechanisms and pathways (i.e., mediation models) that explain the effects of rurality on stage at breast cancer diagnosis, and ultimately, to develop interventions that target the "root cause(s)" of rural disparities in breast cancer stage.

Recommendations and Future Research

The lack of consensus on how best to measure rural residence and the resulting diversity of methodologies was noted in both recent reviews of the cancer literature [2,3]. Our study, together with the meta-analysis by Nguyen-Pham et al. [3], provides some assurances in this regard. Namely, our results suggest that the use of different urban/rural measures is unlikely to be responsible for the mixed literature, at least in the limited case of breast cancer stage at diagnosis. It seems more likely that mixed findings result from different populations, geographies, settings, time periods, designs, and analytic methods applied in prior studies. However, the recent meta-analysis noted no clear patterns in findings across study differences such as sample size, year, patient age, cancer staging method, or control for covariates [3]. Clearer patterns may emerge as the evidence-base grows and we gain more insight from *national* samples/administrative databases, which will be greatly facilitated if studies are directly comparable.

Consistent use of a standardized measure of rural residence would facilitate comparisons across studies. Recently, Meilleur et al. [2] recommended use of the Rural Urban Commuting Area (RUCA) measure. This recommendation was based on conceptual considerations as well as a comprehensive and thoughtful review of the cancer outcomes literature. Our methodologic study builds from this review and extends the nascent evidence base regarding measurement of rural residence in the cancer literature. Given our findings—relatively high agreement and the largely consistent rural disparity in late-stage disease across measures— we support Meilleur et al.'s [2] recommendation of the RUCA. The RUCA measure incorporates both population density and travel distance, can be classified in multiple ways (i.e. with 2, 3, or 4 categories, including others not analyzed here) [20], and lends itself easily to multilevel regression analyses [2]. Given the known spatial and temporal biases introduced by use of ZIP codes [7,8,44], and the ongoing convention of measuring neighborhood factors (e.g. percent poverty) using census tracts [44–46] we suggest researchers opt for census-tract RUCA if data permit.

While we compared some of the most frequently used measures, there are alternative ways to measure urban/rural status. Alternative approaches could alter findings in unknown ways. For example, it has been shown that traditional analysis using geopolitical boundaries may mask or attenuate differences between urban and rural populations [47]. Some researchers develop their own measures. For example, Robert et al. classified areas as 100% rural, mixed urban/rural, or 100% urban using percent of census tract or ZIP code residents living in census-defined urban areas and places [48].

Limitations and strengths

Our study faces several limitations. First, our results may not be generalizable to breast cancer patients living outside of Texas who may differ in unknown ways such as those facing different exposures to rural or urban environments. As one example, Texas has a large rural land mass compared to other states. Thus, individuals from other states living in rural areas may face much shorter travel times to reach an urban area and the healthcare facilities therein. Second, we faced data limitations. Data on multiple factors that might confound the association of rural residence and late-stage diagnosis were not available. Such factors might

include primary care utilization, insurance status, or mammography use. Third, we included multiple variables in the adjusted models which may be on the causal pathway (e.g. mammography capacity). In doing so, we may have underestimated the association between rural residence and late-stage breast cancer. Finally, our study examines rural disparities in a single outcome—breast cancer stage at diagnosis; and we should not assume that results will hold across other cancer outcomes.

Conclusions

Our results document a clear rural disadvantage in late-stage diagnosis of breast cancer. Thoughtful and consistent use of RUCA codes to measure rural residence will facilitate comparisons across the currently heterogeneous cancer outcomes literature. In turn, these comparisons will help to elucidate the impact of rural residence on cancer outcomes across the cancer continuum. Continued research will be needed to elucidate the potential mechanisms such as healthcare access that underlie observed rural disparities.

Acknowledgments

References

1. American Cancer Society. Cancer Facts and Figures. [Accessed on 3-18-14] http://www.cancer.org/acs/groups/content/@epidemiologysurveilance/documents/document/acspc-036845.pdf

2. 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 Oct; 22(10):1657–1667. [PubMed: 24097195]

3. Nguyen-Pham S, Leung J, McLaughlin D. Disparities in breast cancer stage at diagnosis in urban and rural adult women: a systematic review and meta-analysis. Ann Epidemiol. 2014 Mar; 24(3):228–235. [PubMed: 24462273]

4. Monroe AC, Ricketts TC, Savitz LA. Cancer in rural versus urban populations: a review. The Journal of rural health : official journal of the American Rural Health Association and the National Rural Health Care Association. 1992 Summer;8(3):212–220.

5. U. S. Preventive Services Task Force. Screening for breast cancer: U.S. Preventive Services Task Force recommendation statement. Ann Intern Med. 2009 Nov 17; 151(10):716–726. W-236. [PubMed: 19920272]

6. Larson, EH.; Hart, LG. Rural methods health workforce analysis. In: Larson, EH.; Johnson, KE.; Norris, TE.; Lishner, DM.; Rosenblatt, RA.; Hart, LG., editors. State of the health workforce in rural America: State profiles and comparisons. Seattle, WA: WWAMI Rural Health; 2003.

7. Krieger N, Waterman P, Chen JT, Soobader MJ, Subramanian SV, Carson R. Zip code caveat: bias due to spatiotemporal mismatches between zip codes and US census-defined geographic areas--the Public Health Disparities Geocoding Project. Am J Public Health. 2002; 92(7):1100–1102. [PubMed: 12084688]

8. Grubesic TH, Matisziw TC. On the use of ZIP codes and ZIP code tabulation areas (ZCTAs) for the spatial analysis of epidemiological data. Int J Health Geogr. 2006; 5:58. [PubMed: 17166283]

9. Grubesic TH. Zip codes and spatial analysis: Problems and prospects. Socio-Economic Planning Sciences. 2008; 42(2):129–149.

10. U.S. Department of Agriculture Economic Research Service. [Accessed on 3-25-14] Frontier and Remote Area Codes. Overview. http://www.ers.usda.gov/data-products/frontier-and-remote-area-codes.aspx#.UzHv_IXDtbp

11. WWAMI Rural Health Research Center UoW. [Accessed on 3-25-14] RUCA ZIP Version 2. RUCA Data. http://depts.washington.edu/uwruca/ruca-download.php ZIP Code RUCA Approximation http://depts.washington.edu/uwruca/ruca-approx.php

12. Atav AS, Darling R. Comparison of coding schemes for rural-urban designations with New York state counties and birth outcomes as exemplars. Online Journal of Rural Nursing and Health Care. 2012; 12(1):29–39.

13. Hart LG, Larson EH, Lishner DM. Rural definitions for health policy and research. Am J Public Health. 2005 Jul; 95(7):1149–1155. [PubMed: 15983270]

14. Isserman AM. In the national interest: Defining rural and urban correctly in reserach and public policy. International Regional Science Review. 2005; 28(4):465–499.

15. Centers for Disease Control and Prevention. [Accessed 3-25-14] NCHS Urban-Rural Classification Scheme for Counties. 2006. http://www.cdc.gov/nchs/data_access/urban_rural.htm

16. U.S. Department of Agriculture Economic Research Service. [Accessed on 3-25-14] Urban Influence codes for U.S. counties. 2003. http://www.ers.usda.gov/data-products/urban-influence-codes.aspx#.UzH3V4XDtbp Accessed on 3-25-14 Documentation at http://www.ers.usda.gov/data-products/urban-influence-codes/documentation.aspx#.UzH32YXDtbo

17. U.S. Department of Agriculture Economic Research Service. [Accessed on 3-25-14] Rural-Urban Continuum Codes. 2003. http://www.ers.usda.gov/data-products/rural-urban-continuum-codes.aspx#.UzH444XDtbp

18. U.S. Department of Agriculture Economic Research Service. Rural-Urban Commuting Area Codes. Rural-Urban Commuting Area Codes. 2000. http://www.ers.usda.gov/data-products/rural-urban-commuting-area-codes.aspx#.UzH8o4XDtbp

19. U.S. Department of Health and Human Services Centers for Disease Control and Prevention National Center for Health Statistics. NCHS Urban-rural classification scheme for counties. 2012

20. WWAMI Rural Health Research Center UoW. [Accessed on 3-27-14] RUCA Version 1.11 Using RUCA codes: some aggregation schemes for health-related work. https://depts.washington.edu/uwruca/ruca1/ruca-uses11.php

21. Singh GK, Williams SD, Siahpush M, Mulhollen A. Socioeconomic, Rural-Urban, and Racial Inequalities in US Cancer Mortality: Part I-All Cancers and Lung Cancer and Part II-Colorectal, Prostate, Breast, and Cervical Cancers. Journal of cancer epidemiology. 2011; 2011:107497. [PubMed: 22496688]

22. Dai D. Black residential segregation, disparities in spatial access to health care facilities, and late-stage breast cancer diagnosis in metropolitan Detroit. Health Place. 2010 Sep; 16(5):1038–1052. [PubMed: 20630792]

23. Henry KA, Sherman R, Farber S, Cockburn M, Goldberg DW, Stroup AM. The joint effects of census tract poverty and geographic access on late-stage breast cancer diagnosis in 10 US States. Health Place. 2013 May.21:110–121. [PubMed: 23454732]

24. Luo W, Wang F. Measures of spatial accessibility to healthcare in a GIS environment: Synthesis and a case study in Chicago region. Environment and Planning B. 2003; 30(6):865–884.

25. Eberth JM, Eschbach K, Morris JS, Nguyen HT, Hossain MM, Elting LS. Geographic Disparities in Mammography Capacity in the South: A Longitudinal Assessment of Supply and Demand. Health Services Research. 2013 n/a-n/a.

26. Healthy People 2020. United States Department of Health and Human Services. 2nd. Washington, DC: U.S. Government Printing Office; 2010.

27. Cohen J. A coefficient of agreement for nominal scales. Educational and Psychological Measurement. 1960; 20(1):37–46.

28. Byrt T, Bishop J, Carlin JB. Bias, prevalence and kappa. J Clin Epidemiol. 1993 May; 46(5):423–429. [PubMed: 8501467]

29. Landis JR, Koch GG. The measurement of observer agreement for categorical data. Biometrics. 1977 Mar; 33(1):159–174. [PubMed: 843571]

30. Blair SL, Sadler GR, Bristol R, Summers C, Tahar Z, Saltzstein SL. Early cancer detection among rural and urban Californians. BMC Public Health. 2006; 6:194. [PubMed: 16869975]

31. Celaya MO, Berke EM, Onega TL, et al. Breast cancer stage at diagnosis and geographic access to mammography screening (New Hampshire, 1998–2004). Rural and remote Health. 2010 Apr-Jun; 10(2):1361. [PubMed: 20438282]

32. Higginbotham JC, Moulder J, Currier M, Rural V. urban aspects of cancer: first-year data from the Mississippi Central Cancer Registry. Family & community Health. 2001 Jul; 24(2):1–9. [PubMed: 11373161]

33. Liff JM, Chow WH, Greenberg RS. Rural-urban differences in stage at diagnosis. Possible relationship to cancer screening. Cancer. 1991 Mar 1; 67(5):1454–1459. [PubMed: 1991313]

34. McElroy JA, Remington PL, Gangnon RE, Hariharan L, Andersen LD. Identifying geographic disparities in the early detection of breast cancer using a geographic information system. Prev Chronic Dis. 2006 Jan.3(1):A10. [PubMed: 16356363]

35. Cossman RE, Cossman JS, Cosby AG, Reavis RM. Reconsidering the rural-urban continuum in rural health research: A test of stable relationships using mortality as a health measure. Population Research and Policy Review. 2008

36. McLafferty S, Wang F. Rural reversal? Rural-urban disparities in late-stage cancer risk in Illinois. Cancer. 2009 Jun 15; 115(12):2755–2764. [PubMed: 19434667]

37. McLafferty S, Wang F, Luo L, Butler J. Rural - urban inequalities in late-stage breast cancer: spatial and social dimensions of risk and access. Environment and planning. B, Planning & design. 2011; 38(4):724–740. [PubMed: 23335830]

38. Hawley ST, Chang S, Risser D, Zhang Q. Colorectal cancer incidence and mortality in Texas 1990–1992: a comparison of rural classifications. The Journal of rural health : official journal of the American Rural Health Association and the National Rural Health Care Association. 2002 Fall;18(4):536–546.

39. Jackson MC, Davis WW, Waldron W, McNeel TS, Pfeiffer R, Breen N. Impact of geography on mammography use in California. Cancer Causes Control. 2009 Oct; 20(8):1339–1353. [PubMed: 19449107]

40. Bryant H, Mah Z. Breast cancer screening attitudes and behaviors of rural and urban women. Prev Med. 1992 Jul; 21(4):405–418. [PubMed: 1409484]

41. Carr WP, Maldonado G, Leonard PR, et al. Mammogram utilization among farm women. The Journal of rural health : official journal of the American Rural Health Association and the National Rural Health Care Association. 1996; 12(4 Suppl):278–290.

42. Casey MM, Thiede Call K, Klingner JM. Are rural residents less likely to obtain recommended preventive healthcare services? Am J Prev Med. 2001 Oct; 21(3):182–188. [PubMed: 11567838]

43. Wang F, McLafferty S, Escamilla V, Luo L. Late-stage breast cancer diagnosis and health care access in Illinois. Prof Geogr. 2008; 60(1):54–69. [PubMed: 18458760]

44. Krieger N, Waterman PD, Chen JT, Soobader MJ, Subramanian SV. Monitoring socioeconomic inequalities in sexually transmitted infections, tuberculosis, and violence: geocoding and choice of area-based socioeconomic measures--the public health disparities geocoding project (US). Public health reports. 2003; 118(3):240–260. [PubMed: 12766219]

45. Doubeni CA, Jambaulikar GD, Fouayzi H, et al. Neighborhood socioeconomic status and use of colonoscopy in an insured population--a retrospective cohort study. PLoS One. 2012; 7(5):e36392. [PubMed: 22567154]

46. Krieger N, Chen JT, Waterman PD, Soobader MJ, Subramanian SV, Carson R. Choosing area based socioeconomic measures to monitor social inequalities in low birth weight and childhood lead poisoning: The Public Health Disparities Geocoding Project (US). Journal of Epidemiology and Community Health. 2003; 57(3):186–199. [PubMed: 12594195]

47. Hewitt, M. Defining "rural" areas: Impact on health care policy and research. Washington DC: Government Printing Office; 1989. Available at: http://govinfo.library.unt.edu/ota/Ota_2/DATA/ 1989/8912.PDF [Accessed on 3-27-14]

48. Robert SA, Strombom I, Trentham-Dietz A, et al. Socioeconomic risk factors for breast cancer: distinguishing individual- and community-level effects. Epidemiology. 2004 Jul; 15(4):442–450. [PubMed: 15232405]

49. U.S. Department of Commerce. United States Census Bureau. Cartographic Boundary Files - Urban Areas. Census 2000 version http://www.census.gov/geo/maps-data/data/cbf/cbf_ua.html Accessed 3-28-14 Downloaded from: http://www2.census.gov/geo/tiger/PREVGENZ/ua/ua00shp/ua99_d00_shp.zip

Appendix 1

Alternative categorizations of urban/rural measures

Urban/Rural measures by geographic unit and measure source documentation	Alternative categorizations for urban/rural measures	Number of breast cancer patients in each category	
County		**N (%)**	
UI: Urban Influence Code[16]	**Four categories:** 1. "Large metro" (code 1) 2. "Small metro" (code 2) 3. "Non-metro adjacent to metro" (codes 3–5) 4. "Noncore/non-adjacent areas" (codes 6–12)	77,946 (64.6) 28,760 (23.8) 7,592 (6.3) 6,440 (5.3)	
NCHS: National Center for Health Statistics Urban-Rural Classification Scheme[15,19]	**Six categories:** 1. "Large central metropolitan" (code 1) 2. "Large fringe metro" (code 2) 3. "Medium metro" (code 3) 4. "Small metro" (code 4) 5. "Micropolitan" (code 5) 6. "Noncore" counties (code 6)	55,940 (46.3) 22,006 (18.2) 14,660 (12.1) 14,100 (11.7) 7,857 (6.5) 6,175 (5.1)	
RUCC: Rural Urban Continuum Code[17], also known as Beale's codes	**Five categories[21]:** 1. "Large metropolitan" (code 1) 2. "Medium metropolitan" (code 2) 3. "Small metropolitan" (code 3) 4. "Urban nonmetropolitan" (codes 4–7) 5. "Rural nonmetropolitan" (codes 8–9)	77,946 (64.6) 13,021 (10.8) 15,739 (13.0) 13,369 (11.1) 663 (0.6)	
ZIP Code			
FAR: Frontier and Remote Area Code[10]	**Five categories:** 1. No FAR categorization (non-rural category) 2. "Level 1" 3. "Level 2" 4. "Level 3" 5. "Level 4"	11,6073 (96.1) 1,731 (1.4) 1,010 (0.8) 1,386 (1.2) 538 (0.5)	
zRUCA: ZIP-Code Rural Urban Commuting Area[11]	**Four categories /"Categorization A" :** 1. "Urban focused" (codes 1.0, 1.1, 2.0, 2.1, 3.0, 4.1, 5.1, 7.1, 8.1, 10.1) 2. "Large Rural City/Town" (Micropolitan) focused (codes 4.0, 4.2, 5.0, 5.2, 6.0, 6.1) 3. "Small rural town focused" (codes 7.0, 7.2, 7.3, 7.4, 8.0, 8.2, 8.3, 8.4, 9.0, 9.1, 9.2) 4. "Isolated small rural town focused" (codes 10.0, 10.2, 10.3, 10.4, 10.5, 10.6) **Three categories/"Categorization B":** 1. "Urban" (1.0, 1.1, 2.0, 2.1, 3.0, 4.1, 5.1, 7.1, 8.1, 10.1) 2. "Large rural city/town" (4.0, 4.2, 5.0, 5.2, 6.0, 6.1); 3. "Small and isolated small rural town" (7.0, 7.2, 7.3, 7.4, 8.0, 8.2, 8.3, 8.4, 9.0, 9.1, 9.2, 10.0, 10.2, 10.3, 10.4, 10.5, 10.6)	**ZIP** 103,286 (85.6) 9,930 (8.2) 5,141 (4.3) 2,381 (2.0) 103,286 (85.6) 9,930 (8.2) 7,522 (6.2)	**Tract** 100,708 (83.4) 12,553 (10.4) 1,509 (1.3) 5,968 (4.9) 100,708 (83.4) 12,553 (10.4) 7,477 (6.2)
Census Tract			
tRUCA: Tract Rural Urban Commuting Area[18]	Same as z-RUCA, above	See above	
Block Group			
UA/UC: Urban Area/Urban Cluster[49]	**Three categories:** 1. "Urbanized area" 1. "Urban Cluster"	86,884 (72.2) 15,212 (12.6) 18,642 (15.4)	

Urban/Rural measures by geographic unit and measure source documentation	Alternative categorizations for urban/rural measures	Number of breast cancer patients in each category
County		**N (%)**
	2. All other areas are designated as rural.	

RUCA Binary categorization is known as "Categorization C"

Appendix 2

Urban/Rural Classification of Texas Geography, Population, and Breast Cancer Cases using Multiple Classification Methods

	Texas Geography and Population			Texas Breast Cancer Cases			
	Number of Geographies n	Area %	Population %	Total (n=120,738) n (%)	Early (n=110,168) n (%)	Late (n=10,570) n (%)	p
Binary Classifications							
County[a]							
Urban	77	26.8	86.1	106,706 (88.4)	97565 (91.4)	9141 (8.6)	***
Rural	177	73.3	13.9	14,032 (11.6)	12603 (89.8)	1429 (10.2)	
ZIP Code							
Frontier and Remote Area Code (FAR)							
Urban	1308	46.5	94.5	116,073 (96.1)	106010 (91.3)	10063 (8.7)	***
Rural	587	53.5	5.5	4,665 (3.9)	4158 (89.1)	507 (10.9)	
Rural Urban Commuting Area (zRUCA)							
Urban	985	18.5	81.6	103,286 (85.6)	94,457 (91.5)	8829 (8.6)	***
Rural	910	81.5	18.4	17,452 (14.5)	15,711 (90.0)	1741 (10.0)	
Census Tract							
Rural Urban Commuting Area (tRUCA)							
Urban	3524	23.5	83.6	100,708 (83.4)	92117 (91.5)	8591 (8.5)	***
Rural	864	76.5	16.4	20,030 (16.6)	18051 (90.1)	1979 (9.9)	
Block/Block Group							
Urbanized Area/ Urban Cluster (UA/UC)							
Urban	10798[b]	2.7	82.5	102,096 (85.6)	93097 (91.2)	8999 (8.8)	
Rural	3684	97.3	17.5	18,642 (15.4)	17071 (91.6)	1571 (8.4)	
Secondary Classification							
Block/Block Group							
Urban							

	Texas Geography and Population			Texas Breast Cancer Cases			
	Number of Geographies n	Area %	Population %	Total (n=120,738) n (%)	Early (n=110,168) n (%)	Late (n=10,570) n (%)	p
UA	9427	2.1	71.0	86884 (72.0)	79411 (91.4)	7473 (8.6)	***
UC	1371	0.6	11.5	15212 (12.6)	13686 (90.0)	1526 (10.0)	
Rural	3684	97.3	17.5	18,642 (15.4)	17071 (91.6)	1571 (8.4)	

*
 p<.01
**
 p<.001
***
 p<.0001

[a] All three county-based methods agree: Rural Urban Continuum Code (RUCC), Urban Influence Code (UI), and National Center for Health Statistics Urban-Rural Classification Scheme for Counties (NCHS).

[b] Texas geographies, areas, and population are calculated using block groups by assigning the block centroid to nearest block group.

Appendix 3

Unadjusted and adjusted association of urban/rural status and late-stage breast cancer across 7 classification methods (n=120,738) calculated from random intercept models (with random intercepts defined at the county, ZIP-code, census tract, and block group levels, as appropriate).

	Crude Odds Ratio (OR) and 95% CI			Adjusted Odds Ratio (aOR) and 95% CI		
	Method 1 (Binary Rural vs. non-rural)	Method 2	Method 3	Method 1 (Binary Rural vs. non-rural)	Method 2	Method 3
County						
UI 1 (most urban)	1	1	-	1	1	-
2	**1.21 (1.10–1.34)**	**1.15 (1.02–1.31)**	-	**1.14 (1.03–1.26)**	1.06 (0.94–1.20)	-
3	-	**1.28 (1.11–1.46)**	-	-	**1.18 (1.03–1.35)**	-
4 (most rural)	-	**1.33 (1.17–1.52)** ***	-	-	**1.18 (1.03–1.35)** **	-
RUCC 1 (most urban)	1	1	-	1	1	-
2	**1.21 (1.10–1.34)**	**1.31 (1.10–1.54)**	-	**1.14 (1.03–1.26)**	1.11 (0.94–1.32)	-
3	-	1.08 (0.94–1.24)	-	-	1.03 (0.90–1.19)	-
4	-	**1.30 (1.16–1.46)**	-	-	**1.17 (1.04–1.32)**	-
5 (most rural)	-	**1.38 (1.05–1.81)** *	-	-	**1.33 (1.01–1.76)** *	-
NCHS 1 (most urban)	1	1		1	1	
2	**1.21 (1.10–1.34)**	1.05 (0.85–1.29)	-	**1.14 (1.03–1.26)**	1.16 (0.94–1.42)	-
3	-	**1.33 (1.06–1.67)**	-	-	1.24 (0.99–1.55)	-
4	-	1.11 (0.90–1.38)	-	-	1.15 (0.93–1.42)	-
5	-	**1.41 (1.14–1.73)**	-	-	**1.36 (1.10–1.68)**	-
6 (most rural)	-	**1.30 (1.06–1.60)** ***	-	-	**1.28 (1.04–1.58)** **	-
ZIP-Code						

	Crude Odds Ratio (OR) and 95% CI			Adjusted Odds Ratio (aOR) and 95% CI		
	Method 1 (Binary Rural vs. non-rural)	Method 2	Method 3	Method 1 (Binary Rural vs. non-rural)	Method 2	Method 3
z-RUCA 1 (most urban)	1	1	1	1	1	1
2	**1.19 (1.11–1.28)**	**1.24 (1.13–1.37)**	**1.24 (1.13–1.37)**	**1.14 (1.06–1.23)**	**1.18 (1.08–1.30)**	**1.18 (1.08–1.29)**
3	-	**1.16 (1.03–1.30)**	**1.15 (1.05–1.27)** [***]	-	1.08 (0.97–1.21)	1.10 (1.00–1.21)
4 (most rural)	-	1.15 (0.98–1.34)	-	-	1.14 (0.97–1.34)	-
FAR 1 (most urban)	1	1	-	1	1	-
2	**1.27 (1.13–1.43)**	**1.26 (1.02–1.55)**	-	**1.18 (1.05–1.32)**	1.15 (0.95–1.38)	-
3	-	**1.40 (1.09–1.78)**	-	-	**1.29 (1.03–1.62)**	-
4	-	**1.30 (1.06–1.58)**	-	-	**1.21 (1.00–1.46)**	-
5 (most rural)	-	1.06 (0.78–1.45)	-	-	1.02 (0.76–1.39)	-
Tract						
t-RUCA 1 (most urban)	1	1	1	1	1	1
2	**1.18 (1.11–1.25)**	**1.17 (1.09–1.26)**	**1.17 (1.09–1.26)**	**1.09 (1.02–1.16)**	**1.08 (1.01–1.16)**	**1.08 (1.01–1.16)**
3	-	**1.28 (1.06–1.55)**	**1.18 (1.08–1.29)** [***]	-	1.19 (1.00–1.43)	**1.12 (1.02–1.22)** [*]
4 (most rural)	-	**1.16 (1.05–1.28)** [**]	-	-	1.09 (0.99–1.21) [*]	-
Block/Block Group						
UA/UC 1 (most urban)	1	1	-	1	1	-
2	0.96 (0.90–1.01)	**1.20 (1.13–1.27)**	-	1.02 (0.95–1.08)	**1.16 (1.08–1.24)**	-
3 (most rural)	-	0.99 (0.93–1.05)	-	-	1.06 (0.99–1.14)	-

Adjusted models include the following covariates: race/ethnicity, age, year of diagnosis, census tract percent living in poverty, census tract percent black, and census tract percent Hispanic, and mammography access. Bold text indicates statistical significance of OR at p<.05.

[***] linear test of trend p<.001

[**] linear test of trend p<.01

[*] linear test of trend p<.05
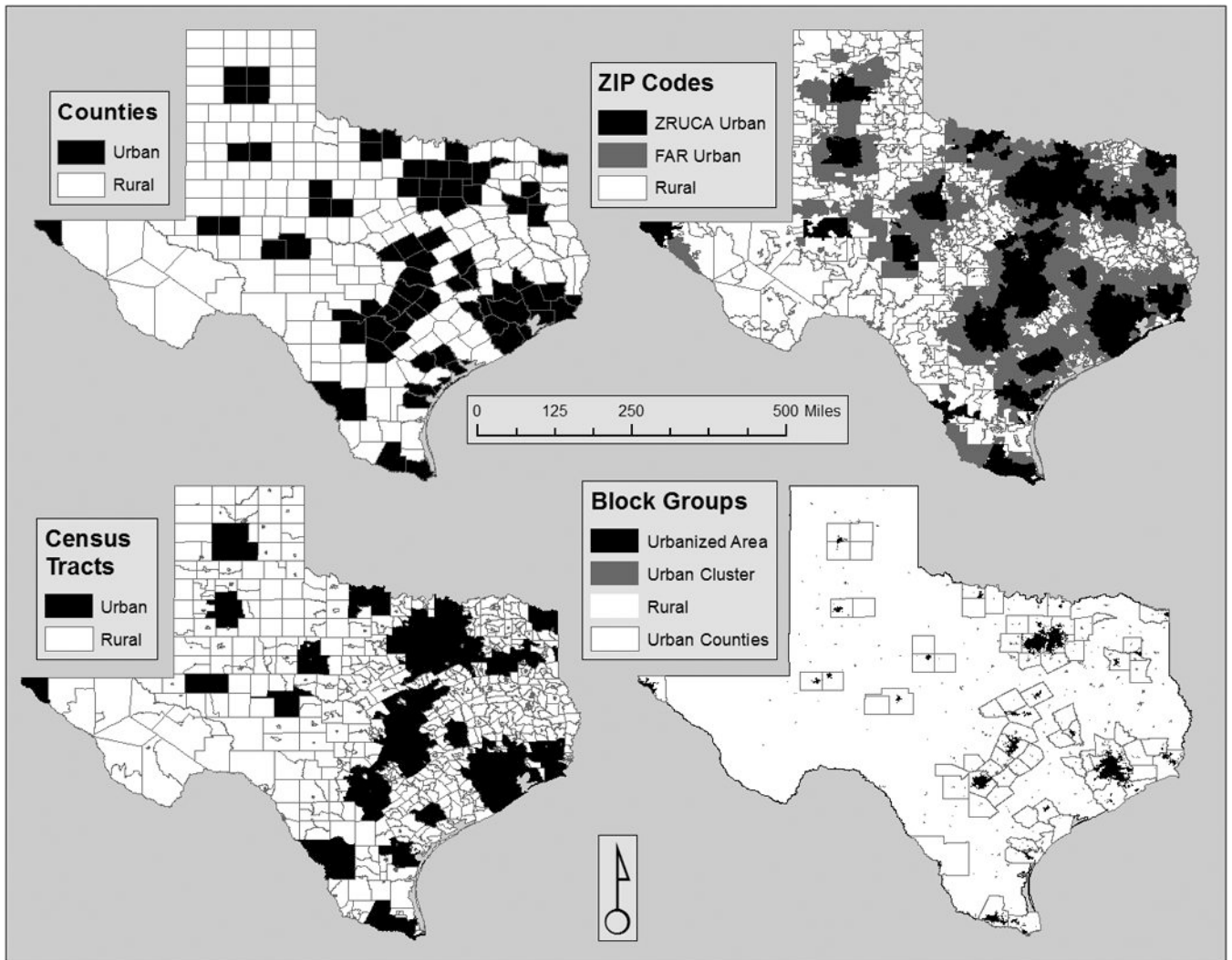
Fig. 1.
Urban vs. Rural Classifications by County (NCHS, RUCC, UI), ZIP Code (ZIP-RUCA, FAR), Census Tract (RUCA), and Block/Block Group (UA/UC) Classification Methods
NCHS= National Center for Health Statistics; RUCC=Rural-Urban Commuting Code; UI=Urban Influence Codes FAR= Frontier and Remote Area Codes; RUCA=Rural Urban Commuting Area; ZIP-RUCA= ZIP-code RUCA; UA/UC=Urbanized area/Urban cluster

Table 1

Characteristics of Seven Classification Methods of Urban/Rural Status

Urban/Rural measures by geographic unit and measure source documentation	Urban/ Rural categorizations		How characterized and (number of categories)
County			
UI: Urban Influence Code[16]	**1**	**Binary** Non-rural: 1–2 Rural: 3–12	Population size, proximity (measured with adjacency and percentage of persons commuting) to metropolitan areas (12)
	2	**Four categories:** 1 (1) [most urban], 2 (2), 3 (3–5), 4 (6–12) [most rural]	
NCHS: National Center for Health Statistics Urban-Rural Classification Scheme[15,19]	**1**	**Binary** Non-rural: 1–4 Rural: 5–6	Population size, proximity to metropolitan areas (6)
	2	**Six categories:** 1 (most urban) - 6 (most rural)	
RUCC: Rural Urban Continuum Code[17], also known as Beale's codes	**1**	**Binary** Non-rural: 1–3 Rural: 4–9	Population size, proximity to an urbanized area and adjacency to metropolitan areas (9)
	2	**Five categories:** 1([most urban]; 1); 2(2); 3(3); 4(4–7); 5([most rural]; 8–9)	
ZIP Code			
FAR: Frontier and Remote Area Code[10]	**1**	**Binary** Non-rural: No FAR Indicators Rural: Any of 4 FAR Indicators	Degree of geographic remoteness characterized by population size and car travel time to nearby urban areas (5[a])
	2	**Five categories:** 0 (most urban/no FAR indicators); 1; 2; 3; 4 (most rural/remote)	
zRUCA: ZIP-Code Rural Urban Commuting Area[11]	**1**	**Binary** Non-rural: 1, 1.1, 2, 2.1, 3, 4.1, 5.1, 7.1, 8.1, 10.1 Rural: (4.0, 4.2, 5.0, 5.2, 6.0, 6.1, 7.0, 7.2, 7.3, 7.4, 8.0, 8.2, 8.3, 8.4, 9.0, 9.1, 9.2, 10.0, 10.2, 10.3, 10.4, 10.5, 10.6)	Population density, proximity to an Urban Area, daily commuting patterns (33)
	2	**Four categories:** 1 ([most urban];1.0, 1.1, 2.0, 2.1, 3.0, 4.1, 5.1, 7.1, 8.1, 10.1); 2 (4.0, 4.2, 5.0, 5.2, 6.0, 6.1); 3 (7.0, 7.2, 7.3, 7.4, 8.0, 8.2, 8.3, 8.4, 9.0, 9.1, 9.2); 4 ([most rural];10.0, 10.2, 10.3, 10.4, 10.5, 10.6)	
	3	**Three categories:** 1 ([most urban]; 1.0, 1.1, 2.0, 2.1, 3.0, 4.1, 5.1, 7.1, 8.1, 10.1); 2 (4.0, 4.2, 5.0, 5.2, 6.0, 6.1); 3 ([most rural]; 7.0, 7.2, 7.3, 7.4, 8.0, 8.2, 8.3, 8.4, 9.0, 9.1, 9.2, 10.0, 10.2, 10.3, 10.4, 10.5, 10.6)	
Census Tract			
tRUCA: Tract Rural Urban Commuting Area[18]	Same as z-RUCA, above		Population density, urbanization, daily commuting patterns (33)
Block Group			
UA/UC: Urbanized Area/Urban Cluster[49]	**1**	**Binary** Non-rural: Urbanized Area or Urban Cluster	Population density (3)

Urban/Rural measures by geographic unit and measure source documentation	Urban/ Rural categorizations		How characterized and (number of categories)
		Rural: Neither Urbanized Area or Cluster	
	2	**Three categories:** 1 ([most urban]; Urbanized Area); 2 (Urban Cluster); 3 ([most rural]; Rural/Neither Cluster nor Area)	

[a]The 5 FAR categories are comprised of the 4 FAR indicators as well as the lack of an indicator (non-rural ZIP codes)

Table 2

Sample Characteristics of Breast Cancer Patients by Rural or non-Rural[a] Residence (n=120,738)

	Non-rural (n=106,706)	Rural[a] (n=14,032)	p
	N (%)	N (%)	(chi²)
Diagnosis Year			
1995–1997	16,567 (15.5)	2,158 (15.4)	.001
1998–2000	20,233 (19.0)	2,704 (19.3)	
2001–2003	22,090 (20.7)	2,999 (21.4)	
2004–2006	22,585 (21.2)	3,064 (21.8)	
2007–2009	25,231 (23.7)	3,107 (22.1)	
Age			
50–59	38,402 (36.0)	3,699 (26.4)	<.001
60–69	32,212 (30.2)	4,342 (30.9)	
70–79	24,404 (22.9)	3,850 (27.4)	
80	11,688 (11.0)	2,141 (15.3)	
Race/ethnicity			
White	77,625 (72.8)	11,412 (81.3)	<.001
Black	10,883 (10.2)	918 (6.5)	
Hispanic	15,490 (14.5)	1,622 (11.6)	
Other	2,188 (2.1)	56 (0.4)	
Unknown	520 (0.5)	24 (0.2)	
Neighborhood poverty			
<10%	57,489 (53.9)	1,994 (14.2)	<.001
10–19.9%	27,789 (26.0)	8,019 (57.2)	
20%	21,428 (20.1)	4,019 (28.6)	
Mammography capacity			
Poor	8,766 (8.2)	5,179 (36.9)	<.001
Adequate	584 (0.6)	2,374 (16.9)	
Excess capacity	97,356 (91.2)	6,479 (46.2)	
	Mean (SD)	Mean (SD)	(t-test)
Neighborhood Composition			
Perblack	11.7 (19.2)	8.7 (11.7)	<.001
Perhispanic	26.5 (25.6)	23.1 (23.2)	<.001

[a]Rural vs. non-rural residence defined using County-based methods (UI, NCHS, RUCC)

Table 3

Percent Agreement, Kappa, and Prevalence- and Bias-Adjusted Kappa (PABAK) between Seven Binary Urban-Rural Measures defined at the County, ZIP Code, Census Tract, and Block Group Levels

Comparisons by Geography and Measure	% Agreement	Kappa	PABAK
County-County			
UI-RUCC	100		
UI-NCHS	100		
RUCC-NCHS	100		
County-ZIP			
County-FAR	92.2	46.3	84.3
County-zRUCA	95.7	81.1	91.6
County-Tract			
County-tRUCA	91.9	66.7	83.8
County-Block Group			
County-BG	81.8	19.1	63.6
ZIP-ZIP			
FAR-zRUCA	89.4	38.3	78.8
ZIP-Tract			
FAR-tRUCA	86.9	31.7	73.8
zRUCA-tRUCA	93.0	73.1	85.9
ZIP-Block Group			
FAR-BG	83.5	8.7	66.9
zRUCA- BG	80.5	23.3	61.0
Tract-Block Group			
tRUCA- BG	78.2	19.0	56.4
All 7 Definitions agree	72.9	54.3	-

Note: all Kappa and PABAK values are p<.0001

Table 4

Unadjusted and adjusted association of urban/rural status and late-stage breast cancer across 7 classification methods (n=120,738)

	Crude Odds Ratio (OR) and 95% CI			Adjusted Odds Ratio (aOR) and 95% CI		
	Method 1 (Binary Rural vs. non-rural)	Method 2	Method 3	Method 1 (Binary Rural vs. non-rural)	Method 2	Method 3
County						
UI 1 (most urban)	1	1	-	1	1	-
2	1.21 (1.14–1.28)	1.14 (1.09–1.19)	-	1.14 (1.06–1.22)	1.06 (1.01–1.12)	-
3	-	1.23 (1.14–1.33)	-	-	1.17 (1.06–1.28)	-
4 (most rural)	-	1.28 (1.18–1.39) ***	-	-	1.17 (1.07–1.27) ***	-
RUCC 1 (most urban)	1	1	-	1	1	-
2	1.21 (1.14–1.28)	1.27 (1.19–1.35)	-	1.14 (1.06–1.22)	1.11 (1.03–1.18)	-
3	-	1.03 (0.97–1.10)	-	-	1.03 (0.96–1.10)	-
4	-	1.25 (1.18–1.33)	-	-	1.16 (1.08–1.25)	-
5 (most rural)	-	1.35 (1.05–1.72) *	-	-	1.31 (1.02–1.68) *	-
NCHS 1 (most urban)	1	1		1	1	-
2	1.21 (1.14–1.28)	0.94 (0.89–1.00)	-	1.14 (1.06–1.22)	1.06 (0.99–1.13)	-
3	-	1.23 (1.15–1.31)	-	-	1.13 (1.05–1.20)	-
4	-	1.01 (0.94–1.08)	-	-	1.04 (0.97–1.11)	-
5	-	1.28 (1.18–1.38)	-	-	1.24 (1.13–1.36)	-
6 (most rural)	-	1.18 (1.08–1.29) ***	-	-	1.15 (1.04–1.27) ***	-
ZIP-Code						
z-RUCA 1 (most urban)	1	1	1	1	1	1
2	1.19 (1.12–1.25)	1.21 (1.13–1.30)	1.21 (1.13–1.30)	1.12 (1.05–1.20)	1.15 (1.07–1.24)	1.15 (1.07–1.24)
3	-	1.16 (1.05–1.27)	1.15 (1.06–1.24) ***	-	1.07 (0.96–1.18)	1.08 (0.99–1.18)
4 (most rural)	-	1.13 (0.98–1.30)	-	-	1.10 (0.95–1.30)	
FAR 1 (most urban)	1	1	-	1	1	-
2	1.28 (1.17–1.41)	1.27 (1.09–1.48)	-	1.17 (1.06–1.30)	1.13 (0.97–1.33)	-

	Crude Odds Ratio (OR) and 95% CI			Adjusted Odds Ratio (aOR) and 95% CI		
	Method 1 (Binary Rural vs. non-rural)	Method 2	Method 3	Method 1 (Binary Rural vs. non-rural)	Method 2	Method 3
3	-	**1.38 (1.14–1.68)**	-	-	**1.26 (1.04–1.54)**	-
4	-	**1.32 (1.11–1.56)**	-	-	**1.22 (1.02–1.44)**	-
5 (most rural)	-	1.08 (0.81–1.45)	-	-	1.02 (0.76–1.37)	-
Tract						
t-RUCA 1 (most urban)	1	1	1	1	1	1
2	**1.18 (1.12–1.24)**	**1.17 (1.10–1.25)**	**1.17 (1.10–1.25)**	**1.09 (1.03–1.16)**	**1.08 (1.01–1.15)**	**1.08 (1.01–1.15)**
3	-	**1.29 (1.09–1.52)**	**1.18 (1.09–1.28)** ***	-	**1.20 (1.01–1.42)**	**1.11 (1.02–1.21)** *
4 (most rural)	-	**1.16 (1.06–1.27)** ***	-	-	1.09 (0.99–1.20) *	-
Block/Block Group						
UA/UC 1 (most urban)	1	1	-	1	1	-
2	0.95 (0.90–1.01)	**1.18 (1.12–1.26)**	-	1.01 (0.95–1.08)	**1.14 (1.08–1.21)**	-
3 (most rural)	-	0.98 (0.92–1.04)	-	-	1.05 (0.99–1.11)	-

Adjusted models include the following covariates: race/ethnicity, age, year of diagnosis, census tract percent black, and census tract percent Hispanic, and mammography access. Bold text indicates statistical significance of p<.05.

***
linear test of trend p<.001

**
linear test of trend p<.01

*
linear test of trend p<.05