Am J Cancer Res 2015;5(7):2314-2319
www.ajcr.us /ISSN:2156-6976/ajcr0010402

Original Article

Does where you live play an important role in cancer incidence in the U.S.?

Amanda J Fogleman[1], Georgia S Mueller[1], Wiley D Jenkins[1,2]

[1]*Center for Clinical Research, School of Medicine, Southern Illinois University, Springfield, IL USA;* [2]*Simmons Cancer Institute, School of Medicine, Southern Illinois University, Springfield, IL USA*

Received May 19, 2015; Accepted June 11, 2015; Epub June 15, 2015; Published July 1, 2015

Abstract: Some studies have shown disproportionate cancer incidence burden in rural areas which may be attributable partly due to the use of 'rural' as a generic term implying homogeneity of risk/protective factors across wide geographic spans. Counties in SEER 18 registries (years 2001-2011) were classified by their Rural-Urban Continuum Code (RUCC) and aggregated into urban, adjacent rural, and non-adjacent rural and were also aggregated into 3 regions: North, South, and West. Two-way ANCOVA was performed with region and RUCC as factors with adjustment for rates of common risk factors obtained from the County Health Rankings (2013). RUCC has a significant effect on incidence rate in urban areas on breast (P =0.001) and prostate (P =0.009). Colorectal significantly varies by region (P<0.0001), and the effect of rurality significantly varies across regions with North highest (P =0.0005). Lung rates significantly vary across both region and RUCC (P<0.0001 and P =0.0001, respectively). The analysis shows that risk-adjusted cancer incidence varies significantly across regions. However, we also found that rural cancer incidence significantly varied across otherwise-similar rural areas implying that 'rural' is not a homogeneous classification.

Keywords: Rural, cancer, health disparities

Introduction

Disparities in cancer incidence among defined populations have been recognized for over 100 years, and there has been much effort to characterize cancer disparities within populations defined by age, race, and gender [1, 2]. Rural populations (which comprise 19% of the US population) are infrequently studied and may suffer increased risks of some cancers [3, 4]. However, assessments of rural populations' cancer risk are frequently inconsistent in magnitude and direction, likely due in part to locally variable prevalence of behaviors such as smoking, alcohol use, inactivity, and detrimental environmental exposures [5, 6]. Also, the contextual nature of health, and local relationships among people, environment and policies are important factors [7]. One of the most widely used measures of life context is the Rural-Urban Continuum Code (RUCC). RUCC is a county-level assessment of rurality/urbanization and is frequently used as it is imbedded within the Surveillance, Epidemiology, and End Results (SEER) database. Codes range from 1-9, with the most metropolitan counties categorized 1 and the most rural 9, while considering adjacency to a metropolitan area a contributing factor. A weakness of the RUCC assignment is that it may not accurately reflect the local context. For example, Macoupin County in Illinois is assigned a RUCC of 1 (Metro-Counties in metro areas of 1 million population or more; 2013) (http://www.ers.usda.gov/data-products/rural-urban-continuum-codes.aspx) yet the Census Bureau reports that 59% of these residents live in rural areas (outside urban areas of 50,000+ people or urban clusters of 2,500-49,000 people) (http://www.hrsa.gov/ruralhealth/policy/definition_of_rural.html) and it is surrounded by counties of lower RUCC designations. Does the presence of an interstate highway to St. Louis 95 miles away imply that Macoupin County residents experience a more metropolitan lifestyle than their rural context would suggest? The RUCC 1 designation does

Table 1. Correlation between rucc and % R by combined and individual SEER registries, 2000-2011

States	Registries	Counties[#]	RUCCs included	RUCC/% R correlation	*P*-value
Overall/combined	N/A	612	1-9	0.65	<0.0001
CA	Los Angeles	1		0.90	<0.0001
	San Francisco	5	1		
	San Jose-Monterey	4	1-2		
	California excluding	48	1-8		
CT	Connecticut	8	1, 2, 4	0.31	NS
GA	Atlanta Metropolitan	5	1	0.51	<0.0001
	Rural Georgia	10	1, 6-9		
	Greater Georgia	144	1-9		
HI	Hawaii	5	2, 3, 5	0.005	NS
IA	Iowa	99	2-9	0.58	<0.0001
KY	Kentucky	120	1-9	0.60	<0.0001
LA	Louisiana	64	1-9	0.55	<0.0001
MI	Detroit Metropolitan	3	1	ND	ND
NJ	New Jersey	21	1-3	0.21	NS
NM	New Mexico	33	2-7, 9	0.67	P<0.0001
UT	Utah	29	1-4, 6-9	0.77	<0.0001
WA	Seattle (Puget Sound)	13	1-6, 9	0.95	<0.0001

ND = no data. [#]Number of counties.

not accurately reflect the rural nature of this county of <47,000 residents, with substantial farmland, and <6,000 inhabiting its largest town (http://quickfacts.census.gov/qfd/index.html).

It is important to accurately capture the rural/urban life context so that proper risk assessments may be made. Cancer risk and interventions are not one-size-fits-all across different populations. Consider, for example, Bronfenbrenner's social ecologic model as a method of examining health bi-directionally within multiple circles of influence [8]. The innermost circle of the individual is located within the larger contexts of family, other personal relationships, local culture and norms and overarching public policies and environment. Researchers and policymakers need to factor how an individual's disposition and resources contribute to their health as well as the effectiveness of an intervention [9]. The Centers for Disease Control and Prevention has slightly modified this theory for their Colorectal Cancer Control Program to incorporate all levels of influence to increase colorectal cancer screening (http://www.cdc.gov/cancer/crccp/sem.htm). While it is intuitive that cancer risk would vary across communities separated by distance, the most frequent assumption for rural/urban studies is that 'rural' is synonymous. These distinct entities are combined into a single group and may be insensitive to important local variations.

We sought to explore three main questions regarding rural classifications and cancer risk. The first asks if RUCC classifications accurately reflect the extent of rural life, i.e. does RUCC correlate well with the actual proportion of a county's residents living in rural areas (% R) as defined by the Census Bureau? The second further explores correlations between cancer incidence and degree of rurality, and if correlational variations within RUCC and % R significantly differ. The third question asks if otherwise-similar rural areas across SEER data have similar cancer incidence after adjusting for known factors (e.g. smoking rates). The hypothesis driving this work is local cultural/contextual influence varies across geographic regions and has direct impact upon cancer incidence.

Methods

Data were obtained from SEER 18 registries for the time period 2001-2011. We included breast, colorectal, lung and prostate (B/C/L/P) cancer as they are the four of the most common cancers and account for 57.6% of all cancer diagnoses in the U.S. as of 2008 [10]. To

Table 2. Intergroup analysis by cancer type by RUCC and region, 2000-2011

Cancer site	Cancer incidence rate			Intergroup analyses
	RUCC-u	RUCC-a	RUCC-na	
breast	114.2	108.4	108.5	RUCC-u vs. RUCC-a: P=0.0018
				RUCC-u vs. RUCC-na: P=0.0055
				RUCC-a vs. RUCC-na: n/s
colorectal	44.8	45.6	45.8	RUCC-u vs. RUCC-a: n/s
				RUCC-u vs. RUCC-na: n/s
				RUCC-a vs. RUCC-na: n/s
lung	72.1	67.6	65.8	RUCC-u vs. RUCC-a: P=0.0001
				RUCC-u vs. RUCC-na: P=0.0001
				RUCC-a vs. RUCC-na: n/s
prostate	143.3	140.1	132.8	RUCC-u vs. RUCC-a: n/s
				RUCC-u vs. RUCC-na: P=0.0097
				RUCC-a vs. RUCC-na: n/s
	North	South	West	
breast	117.9	118.6	94.6	North vs. South: n/s
				North vs. West: P<0.0001
				South vs. West: P<0.0001
colorectal	50.1	45.6	40.6	North vs. South: P=0.0031
				North vs. West: P<0.0001
				South vs. West: P=0.0023
lung	69.3	77.0	59.3	North vs. South: P=0.0013
				North vs. West: P<0.0001
				South vs. West: P<0.0001
prostate	140.3	150.6	125.3	North vs. South: n/s
				North vs. West: n/s
				South vs. West: n/s

obtain representation across multiple levels of rurality, registries representing entire state populations were included in the analysis. For regional variation analysis, state registries were categorized into 3 regions based on geography: North (Iowa, Michigan, Connecticut and New Jersey), South (Georgia, Louisiana, and Kentucky) and West (New Mexico, Utah, California, Washington and Hawaii). As not all state registries include all RUCCs, these regions were chosen in order to allow for all RUCCs to be represented. For rural/urban variation analysis, counties were categorized into urban (RUCC-u; 1-3), adjacent rural (RUCC-a; 4, 6, 8) and non-adjacent rural (RUCC-na; 5, 7, 9). The following covariates were obtained from County Health Rankings (CHR): % smokers (% Sm), % obesity (% Ob), % excessive drinking (% Al), % physical inactivity (% Ia), and proportion of the residents living in rural areas (% R) [11].

Pearson Correlations were used to examine a) correlations between RUCC level and % R, and b) correlations of RUCC and % R with BCLP incidence. Two-way Analysis of Covariance (ANCOVA) was performed to assess the overall impact of RUCC/% R, CHR-derived covariates, and adjacency to a metropolitan area on BCLP incidence. The main effects included region and RUCC category with an interaction of region and RUCC category. If the interaction was not significant, the main effects model was reported. Least squares means multiple comparison tests were performed when appropriate and least-squares means were calculated. All diagnostic tests for the assumptions of ANCOVA were checked and were not violated.

Results

There were a total of 12 states, representing 612 counties and drawn from 16 registries included in the analysis (**Table 1**). Georgia and California are the only 2 states to include multiple registries. Georgia contains three registries: Atlanta (Metropolitan), Rural Georgia and Greater Georgia. California contains four registries: San Francisco-Oakland SMSA, San Jose-Monterey, Los Angeles County and California excluding SF/SJM/LA. The correlation between RUCC and % R was significant overall ($P<0.0001$) and for 7 of the 12 individual states (each $P<0.0001$). However, high significance does not necessarily imply high correlation, and the combined registries correlation was only 0.65, while the other significantly correlated states ranged from 0.51 to 0.95).

Next, we examined rural/urban B/C/L/P incidence variation, with counties organized by rurality (RUCC-u, RUCC-a, RUCC-na) or region (North, South, West; **Table 2**). Breast and lung cancer incidence were significantly higher in RUCC-u compared to both RUCC-a ($P=0.0018$ and 0.0001, respectively) and RUCC-na ($P=0.0055$ and 0.0001), but insignificantly varies between the rural categories. In prostate

Table 3. Modeling interactions and significant covariates

Cancer	N	R^2	Adj. R^2	RUCC	Region	Interaction	Covariates
Breast	484	0.29	0.27	0.0146	0.0013	NS	smoking, *smoking*region*, obesity, *inactivity*, *inactivity*region*, and drinking
Colorectal	484	0.44	0.42	P<0.0001	0.185	0.0005	*smoking*, *smoking*RUCC*, *obesity*, *obesity*region*, inactivity, drinking, and drinking* RUCC
Lung	484	0.75	0.74	P<0.0001	P<0.0001	P<0.0001	*smoking*, *obesity*, *inactivity*, drinking, and *drinking*region*
Prostate	484	0.26	0.24	0.0745	0.0088	NS	*smoking*, *smoking*region*, obesity, *obesity*region*, *inactivity,* drinking, and *drinking*region*

*Italicized covariates are significant at P<0.05.

cancer RUCC-na is significantly lower than RUCC-u (*P*=0.0097). For colorectal cancer incidence, there was no difference between RUCC designations. Within regional designations, breast cancer was lower in the West compared to North and South, while all three regions were mutually significantly different for colorectal and lung (West the lowest for both cancers). No differences were observed for prostate cancer.

Finally, we modeled the interaction between RUCC and Region on B/C/L/P incidence (**Table 3**). Individually, RUCC was significant for B/C/L (*P*=0.01, *P*<0.0001, *P*<0.0001, respectively) and Region for B/L/P (*P*=0.001, *P*<0.0001, *P*<0.009, respectively) while their interactions were significant only for C/L (*P*=0.0005 and *P*<0.0001). Multiple covariates were significant for the cancer-specific models, and while none were common to all models two (smoking and inactivity) were common in three. Overall adjusted model fit (as measured by adjusted R^2) ranged from 0.24 to 0.74.

Discussion

This work revealed many factors which are important for continuing research into rural cancer. One is the commonly-used measure of rurality, RUCC, only modestly correlates with proportion of residents living in rural areas and varies considerably across SEER registries. This may have a significant impact for studies where the rural life context is considered synonymous across equal RUCC. Second, unadjusted modeling found clear differences in 'urban'/'rural' cancer incidence, with lung and breast cancer higher in urban areas and prostate cancer lower than some rural areas. Importantly, there were no differences in rural areas based upon adjacency, implying RUCCs only differing by adjacency status (e.g. 8 and 9) may be combined to increase sample size without losing sensitivity. Finally, risk factor-adjusted modeling shows that region and rurality are frequently significant predictors of all four cancers, often with interactions, implying regional differences in causative exposures are poor estimators of risk.

Since our results demonstrate there is a component of cancer risk associated with rural areas but it is not consistent across areas of similar rurality, one may utilize the social ecological theory to design personalized interventions for individuals or communities. For example, obesity was a significant factor in C/L/P cancers. While the individualized behavior change approach to diet is specific to the individual needs, it is resource intensive. Further, an individual's diet is influenced by local culture that may significantly vary across regions (e.g. many fried foods in the South). Thus, community-level interventions designed to adjust social norms towards healthier dietary options would complement individual-level efforts. A second example is addressing inactivity, significant for B/L/P cancers. Both these factors have been successfully addressed in Georgia where community-based participatory research methodology was employed to develop diet and exercise plans specific for rural residents [12]. Importantly, rural is not urban, and rural may not be synonymous. Given the different context for life and health, the data and intuition indicate that 'rural' may too frequently be a generic term of insufficient sensitivity for adequate characterization.

There are some limitations in this work. Both the cancer incidence and risk factor data were aggregated to the county level, limiting the strength of identified associations. Also, we combined multiple registries within a single state (e.g. GA) into a single 'state' registry with the assumption that this represented the entire state population. Further, this study examined incidence rates. The inclusion of mortality rates may be needed for a more complete picture of the rural/urban difference. The data also highlight the potential pitfall of *P*-value reliance for estimating significance. As elegantly described by Nuzzo, very low *p*-values do not necessarily translate into equally high practical relevance [13]. The variables from CHR are from different years than the cancer incidence data in SEER as we were using available data gathered from different sources that do not allow for access to their weighting system. However, the likelihood of significant changes in the percentages of smoking, obesity, physical activity, and alcohol abuse within 1-2 years is expected to be small.

Residents of rural areas comprise nearly 20% of the US population, yet their context for life and risk of cancer (specific to this work) is not well understood. The literature often groups 'rural' areas, no matter how great of geospatial extent, into a single category despite knowing that diet, habits, and culture vary significantly across the country. We have shown rural is not the same everywhere, RUCC may not be an adequate measure of rurality for cancer research, and significant risk factors vary by both rurality and regions of the country. To better understand the cancer risk rural residents face, and to design more specific and effective risk reduction interventions, studies of rural areas need to become more focused and explicit regarding the factors important for their local populations.

Disclosure of conflict of interest

None.

Address correspondence to: Amanda J Fogleman, Center for Clinical Research, Population Health Science Program, Southern Illinois University School of Medicine, 801 N. Rutledge St., PO Box 19664, Springfield, IL USA 62794-9664. Tel: 217-545-6987; Fax: 217-545-0799; E-mail: afogleman@siumed.edu

References

[1] Smuckler EA. Chemicals, cancer and cancer biology. West J Med 1983; 139: 55-74.
[2] Gehlert S, Colditz GA. Cancer disparities: unmet challenges in the elimination of disparities. Cancer Epidemiol Biomarkers Prev 2011; 20: 1809-1814.
[3] McLafferty S, Wang F. Rural reversal? Rural-urban disparities in late-stage cancer risk in Illinois. Cancer 2009; 115: 2755-2764.
[4] Singh GK. Rural-urban trends and patterns in cervical cancer mortality, incidence, stage, and survival in the United States, 1950-2008. J Community Health 2012; 37: 217-223.
[5] Hock JM, Nealley A, Morrison D, Farah C, Hosgood HD, Zahm S. Unmet challenges in cancer disparities--letter. Cancer Epidemiol Biomarkers Prev 2012; 21: 248-249.
[6] Weaver KE, Palmer N, Lu L, Case LD, Geiger AM. Rural-urban differences in health behaviors and implications for health status among US cancer survivors. Cancer Causes Control 2013; 24: 1481-1490.
[7] Cummins S, Curtis S, Diez-Roux AV, Macintyre S. Understanding and representing ‘place’ in health research: a relational approach. Soc Sci Med 2007; 65: 1825-1838.
[8] Bronfenbrenner U. Ecology of the family as a context for human development: Research perspectives. Dev Psychol 1986; 22: 723-42.
[9] Grzywacz JG, Fuqua J. The social ecology of health: leverage points and linkages. Behav Med 2000; 26: 101-115.
[10] Howlader N, Noone AM, Krapcho, M. SEER cancer statistics review, 1975-2008. Bethesda, MD: National Cancer Institute; 2011.
[11] 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; 24: 228-235.
[12] Kegler MC, Alcantara I, Veluswamy JK, Haardörfer R, Hotz JA, Glanz K. Results from an intervention to improve rural home food and physical activity environments. Prog Community Health Partnersh 2012; 6: 265-77.
[13] Nuzzo R. Scientific method: statistical errors. Nature 2014; 506: 150-152.