Does the Choice of Rurality Measurement Matter? A Case Study of Rural-Urban Gradients in the Prevalence of Depression

INTRODUCTION

The concept of "rural" in demographic and health research has been examined in depth throughout the past several decades. The environmental, socio-political, cultural, economic, and demographic characteristics of rural America are, in many ways, vastly different than those of urban and suburban areas, presenting a unique set of circumstances that have implications far beyond research and policy.

However, appropriate measurement of rurality remains a challenge in demographic and health research in several ways. One specific challenge is seeking the optimal geographic unit on which to assess rurality. For research in the United States, several choices exist, each their own benefits and drawbacks, including state, county, zip code, and census tract, among others [1-3]. A second challenge is defining rurality itself. While urban-rural gradients in resources, health indicators, and other components of society are well documented, there has been comparatively little inquiry into how rurality is actually defined and measured [4], and there remains a lack of a united consensus across disciplines as to a universal measurement of rurality [5]. **Therefore, the objectives of this study are: (1.) to describe, compare, and contrast five common measures of rurality, highlighting distributional properties of each measure in US counties; and (2.) to demonstrate how the population prevalence of depression varies by rurality and (3.) the inferences about the association between rurality and depression depends upon the rurality measurement used.** In this extended abstract, we will highlight several key findings of this empirical research and its applications for future research in the development and use of rurality measures.

METHODS

Data. To conduct the analysis, data from several sources were first merged to form one large database of county characteristics. Rurality measurements were obtained from the 2010 US Decennial Census and the US Department of Agriculture (USDA). Reported county-level prevalence of depression was estimated from the 2010 Behavioral Risk Factor Surveillance System (BRFSS), after aggregating individual observations to the county level. Five measurements of rurality were used in for this analysis. **Table 1** shows the details of each measure.

Statistical Analysis. The univariate distributions and frequencies were obtained separately for each variable, including both the five rurality measures and self-reported depression, obtained from the BRFSS. Next, each of the five measurements was compared to each other using Spearman's rank correlations. We

then stratified the analysis by Census region—Northeast, South, Midwest, and West—and repeated the analysis. For the second objective, age-adjusted depression prevalence was modeled using each measure using simple bivariate statistical measures appropriate to the rurality construct using graphs and tables.

Table 1: Summary of rurality measured used in analysis

Source	Rurality Measure	Distribution	Description
2010 US Census	Log of population density	Continuous, right-skewed, even after taking logarithm	Natural log of the quotient of county population size divided by county land area
2010 US Census	Log of population density quartile	Ordinal, 4 levels	Counties divided into quartiles based on log of population density
2010 US Census	Percent urban population	Continuous, right-skewed	US Census definition of percent of county population considered "urban"
2003 USDA	Rural-urban continuum codes	Ordinal, 10 levels	Based on proximity of metropolitan statistical area and population size, arranged as a continuum
2003 USDA	Urban influence codes	Ordinal, 6 levels	Based on the estimated economic influence of urban areas on counties and population size

RESULTS

To address the first study objective, the results of Spearman's correlation analysis are shown in **Table 2.** Not surprisingly, there were strong rank correlations among variables from the same source. For example, log of population density was strongly associated with population density quartile (rho = 0.598, $p < 0.001$). Similarly, rural-urban continuum codes were strongly associated with urban influence codes (rho = 0.447, $p < 0.001$). Comparisons across data sources (Census Bureau vs. USDA) were all statistically significant, but generally were smaller in magnitude.

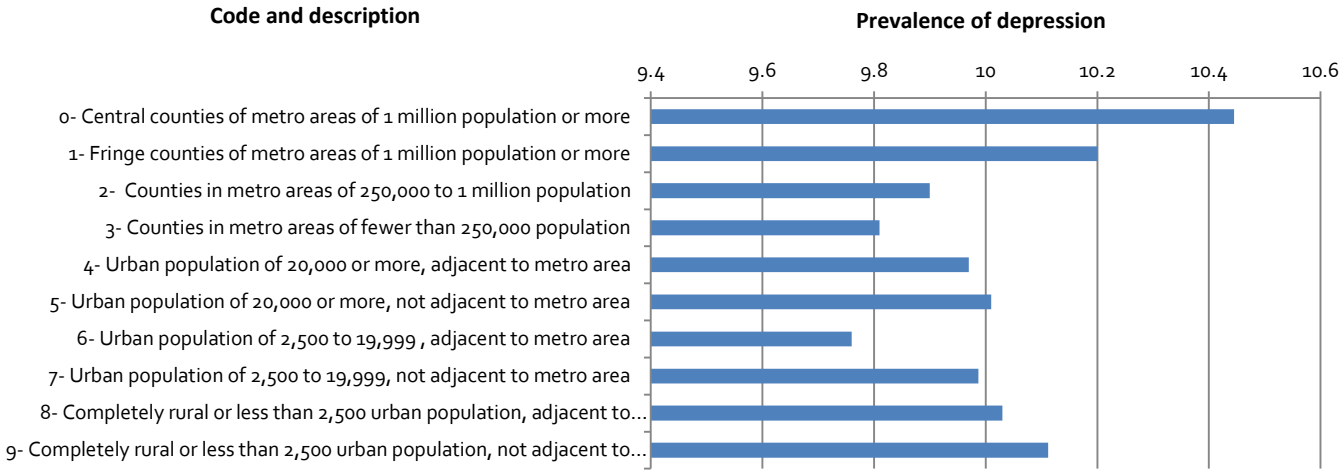
Table 2: Spearman's rank correlation (and p-values) for each combination of rurality variables, including the outcome of reported depression prevalence (N = 3033 counties)
(Note: for rural-urban continuum codes and urban influence codes, each has been reversed for consistent directionality with population density. I.e. Higher numbers indicate more urban areas.)

	Population density quartile	Percent urban population	10 - Rural-urban continuum code	6 - Urban influence code	Prevalence of depression
Log population density	0.598 (< 0.001)	0.349 (< 0.001)	0.211 (< 0.001)	0.192 (< 0.001)	0.023 (0.177)
Population density quartile		0.475 (< 0.001)	0.202 (< 0.001)	0.155 (0.001)	0.010 (0.491)
Percent "urban" population			0.447 (< 0.001)	0.296 (< 0.001)	0.025 (0.152)
10 - Rural-urban continuum code				0.507 (< 0.001)	0.029 (0.114)
6 - Urban influence code					0.011 (0.432)

The analysis by region showed somewhat disparate results compared to the combined analysis shown above. The associations among the rurality measurements were generally highest in the Northeast and West regions (rho values ranged from 0.105 to 0.627), and lowest in magnitude in the Midwest and South regions (rho values ranged from 0.102 to 0.511).

For the second and third objectives, none of the rurality measures were significantly associated with age-adjusted depression prevalence, although each showed a slight positive associated with increased urbanicity and decreased rurality. More detailed associations between depression and each of the five measures were illustrated through graphs. One of the most telling graphs is shown in **Figure 1** for the association between rural-urban continuum code and prevalence of depression. The results suggest a potential non-monotonic association between prevalence of depression and rurality. Depression was greatest in the most urban and the most rural areas. Similar results were observed for other measures.

Figure 1: Weighted, age-adjusted population prevalence of depression by rural-urban continuum code



DISCUSSION

We found moderately strong associations among the five measures of rurality, yet none of the associations were close to perfect correlation. These associations among rurality measures varied somewhat by region. None were significantly associated with depression prevalence. However, there is the potential for the association between depression prevalence and rurality to be non-monotonic, as evidenced by the analysis of individual rurality measures and depression that rank correlation may mask.

There are a number of caveats to this analysis. While the depression prevalence was age-adjusted, other factors were not controlled for that may help explain some of the variability in depression prevalence

among counties. Only five representative measures were used, while myriad others exist. Also, there have been substantial demographic changes in the past several decades. Nearly one-quarter of counties, for example, experienced changes in rural-urban continuum between 1993 and 2003, some changed up to 7 ranks, essentially from a very rural to very urban county in just 10 years' time. Lastly, self-reported depression may not truly reflect the actual prevalence of physician-diagnosed depression, so these measurements may be imprecise.

CONCLUSION

This is among the first empirical analyses comparing and contrasting rurality measures used in demographic, sociological, and health research. Although there are moderately strong associations among the measures, few show bivariate consistency. The heterogeneity in these measures observed in the regional analysis suggests that rurality may have a different meaning with respect to other, unmeasured factors, depending upon the setting. These findings can help initiate future research designed to seek more precise and descriptive measures of rurality, or appropriate composite measures of rurality that reflect the conditions on the ground in each geographic unit. The importance of understanding and quantifying rurality cannot be undervalued when assessing the needs of communities and developing culturally-sensitive interventions designed to address those social, economic, or health care needs in the US population.

REFERENCES

1. Butler DC, Petterson S, Phillips RL, Bazemore AW. Measures of social deprivation that predict health care access and need within a rational area of primary care service delivery. Health Serv Res. 2012 Jul 20. Epublished ahead of print.
2. Krieger N, Chen JT, Waterman PD, Soobader MJ, Subramanian SV, Carson R. Geocoding and monitoring of US socioeconomic inequalities in mortality and cancer incidence: does the choice of area-based measure and geographic level matter?: the Public Health Disparities Geocoding Project. Am J Epidemiol. 2002 Sep 1;156(5):471-82.
3. Mehrotra R, Story K, Guest S, Fedunyszyn M. Neighborhood location, rurality, geography and outcomes of peritoneal dialysis in the United States. Perioneal Dialysis Int 2012;32(3):322-31.
4. Matthews AM. Variations in the conceptualization and measurement of rurality- conflicting findings on the elderly widowed. J Rural Studies 1988;4(2):141-50.
5. Ricketts TC, Johnson-Webb KD. What is "rural" and how to measure "rurality": A focus on health care delivery and health policy. Technical Issues Paper prepared for the Federal Office of Rural Health Policy, 1996 Dec.