


A comparison between perceived rurality and established geographic rural status among Indiana residents

Oindrila Bhattacharyya, PhD^{a,b,c,*} , Susan M. Rawl, PhD, RN^d, Stephanie L. Dickinson, MS^e, David A. Haggstrom, MD, MAS^{f,g,h}

Abstract

The study assessed the association and concordance of the traditional geography-based Rural-Urban Commuting Area (RUCA) codes to individuals' self-reported rural status per a survey scale. The study included residents from rural and urban Indiana, seen at least once in a statewide health system in the past 12 months. Surveyed self-reported rural status of individuals obtained was measured using 6 items with a 7-point Likert scale. Cronbach's alpha was used to measure the internal consistency between the 6 survey response items, along with exploratory factor analysis to evaluate their construct validity. Perceived rurality was compared with RUCA categorization, which was mapped to residential zip codes. Association and concordance between the 2 measures were calculated using Spearman's rank correlation coefficient and Gwet's Agreement Coefficient (Gwet's AC), respectively. Primary self-reported data were obtained through a cross-sectional, statewide, mail-based survey, administered from January 2018 through February 2018, among a random sample of 7979 individuals aged 18 to 75, stratified by rural status and race. All 970 patients who completed the survey answered questions regarding their perceived rurality. Cronbach's alpha value of 0.907 was obtained indicating high internal consistency among the 6 self-perceived rurality items. Association of RUCA categorization and self-reported geographic status was moderate, ranging from 0.28 to 0.41. Gwet's AC ranged from -0.11 to 0.26, indicating poor to fair agreement between the 2 measures based on the benchmark scale of reliability. Geography-based and self-report methods are complementary in assessing rurality. Individuals living in areas of relatively high population density may still self-identify as rural, or individuals with long commutes may self-identify as urban.

Abbreviations: Gwet AC = Gwet agreement coefficient, RUCA = Rural-Urban Commuting Area.

Keywords: concordance, RUCA codes, rural, survey, urban

1. Introduction

The construct of rurality, often used as proxies to understand individuals' social structure, has been associated with a wide range of health behaviors and health outcomes. Interdisciplinary studies have shown that socio-demographic, cultural, and environmental constructs specific to individuals' area of residence lead to significant health disparities, with rural populations experiencing higher obesity rates, low physical activity levels,^[1-3] higher smoking rates,^[4,5] and greater

alcohol consumption.^[6] These are all health behaviors that serve as risk factors for chronic diseases like cancer, diabetes, and heart disease. Moreover, rural populations, across all racial/ethnic groups, have higher all-cause mortality rates than do their urban counterparts.^[7]

Rurality is commonly defined by Rural Urban Community Area (RUCA) codes from the census tract (rural defined as large rural city/town, small rural town, or isolated small rural town). Criteria are applied to measures of population density, urbanization, and daily commuting to identify urban cores and adjacent

This work was aided by National Cancer Institute supplement to Indiana University Cancer Center Grant (P30 CA082709-17S6), Indiana University Melvin and Bren Simon Cancer Center, Indianapolis, Indiana.

The authors have no conflicts of interest to disclose.

The datasets generated during and/or analyzed during the current study are not publicly available, but are available from the corresponding author on reasonable request.

This study contains human subject. The IUPUI Institutional Review Board approved the study.

Supplemental Digital Content is available for this article.

^a Indiana University Purdue University, Department of Economics, Indianapolis, IN, USA, ^b James Comprehensive Cancer Center, The Ohio State University, Columbus, OH, USA, ^c The William Tierney Center for Health Services Research, Regenstrief Institute Inc, Indianapolis, IN, USA, ^d Indiana University School of Nursing, Indiana University Melvin and Bren Simon Cancer Comprehensive Center, Indianapolis, IN, USA, ^e Department of Epidemiology & Biostatistics, Indiana University School of Public Health-Bloomington, Bloomington, IN, USA, ^f Indianapolis VA HSR&D Center for Health Information and Communication,

Roudebush VA, Indianapolis, IN, USA, ^g Division of General Internal Medicine & Geriatrics, Indiana University School of Medicine, Indianapolis, IN, USA, ^h Center for Health Services Research, Regenstrief Institute, Indianapolis, IN, USA.

*Correspondence: Oindrila Bhattacharyya, Research Specialist II, Center for Health Services Research, Regenstrief Institute, Inc, 1101 W 10th St, Indianapolis, IN-46202, USA (e-mail: oinbhatt@iu.edu).

Copyright © 2023 the Author(s). Published by Wolters Kluwer Health, Inc. This is an open-access article distributed under the terms of the Creative Commons Attribution-Non Commercial License 4.0 (CCBY-NC), where it is permissible to download, share, remix, transform, and buildup the work provided it is properly cited. The work cannot be used commercially without permission from the journal.

How to cite this article: Bhattacharyya O, Rawl SM, Dickinson SL, Haggstrom DA. A comparison between perceived rurality and established geographic rural status among Indiana residents. *Medicine* 2023;102:41(e34692).

Received: 21 November 2022 / Received in final form: 19 July 2023 / Accepted: 20 July 2023

<http://dx.doi.org/10.1097/MD.00000000000034692>

areas with varying levels of economic integration with those cores.^[8] However, the common measurement of rurality, through designated zip-code/census tract/county-based scales, presents challenges in capturing variability in rurality across different geographies, understanding mechanisms among the rural population related to health and healthcare access, and finally, how well geographic location represents the multidimensional nature of rurality, including an individual own perception of rural-urban status.^[9] Self-perception of belonging to a rural community is related to a “sense of place,” a construct that is less rigid than geographic boundaries, and consists of personal attachments and meaning associated with places by residents^[10]; Paskett et al have previously disseminated a set of items intended to measure rural community identity.^[11] Therefore, assessing the internal validity of self-perceived measures of rural status, and their correlation with geographic-based scales of rurality, is important to advance our understanding of the mechanisms, and interpretation of findings, related to rural health disparities.^[9]

Numerous studies in the past have explored various geographic units and definitions for measuring rurality,^[12–15] while acknowledging the sensitivity of key health outcome measures to different measuring techniques for rurality.^[16] Given the disease burden among rural populations and the regular use of area of residence measures like RUCA in health services research, it is imperative to evaluate whether attributing rural-urban context effectively captures factors based on individuals’ perceptions of rurality in various geographic contexts which is crucial in understanding how different measures of rurality may have differential associations with health-related behaviors, access to care, and subsequent health outcomes. Onega et al 2020 study was the first to examine the relationship between individuals’ perceived rural-urban status and their RUCA designation based on zip code in the New England area.^[9] We used a different geographic setting—Industrial Midwest and studied the residents in Indiana to compare traditional geography-based RUCA codes based on their place of residence with self-reported rural status, measured using 6 items through a survey scale to assess their level of association and concordance. Additionally, to ensure robustness of our findings, we evaluated the internal consistency of the self-perceived rurality scale and examined the scale construct validity to, respectively, ensure the items appropriately captured the underlying construct of rurality and supported the interpretation of what these response items reflect. We hypothesized that the study population might perceive their rural-urban context differently from their designated RUCA category, as indicated by measures of association and concordance.

2. Methods

We conducted a cross-sectional, mailed survey from January 2018 through February 2018, known as the Hoosier Health Survey, among Indiana residents who were seen at least once in the past 12 months at Indiana University Health, a statewide integrated health system with 19 hospitals in Indiana and 178 clinics with outpatient practices. The survey was conducted to understand better the needs of the population catchment area served by the I.U. Cancer Center.^[17] The Indiana University Purdue University Indianapolis Institutional Review Board approved the study.

2.1. Study cohort

A random, stratified sample of 8,000 adults was generated out of an initial list of 284,062 individuals, seen at least once in the past 12 months in the health system and lived in 1 of 34 Indiana counties with a higher-than-average cancer mortality rate. Geographic location and race were evenly weighed in the sample stratification. The initial goal was to sample 2000 individuals from each of 4 strata (rural White, rural Black, urban

White, urban Black). However, due to the rural Black category having only 524 participants, the remainder of the 2000 allocated was sampled from the rural White strata, ensuring 4000 participants each from rural and urban areas. Following sample stratification, 21 patients were excluded because their primary care provider declined to approve their participation in the survey; hence, a total of 7979 surveys were mailed. A total of 970 adults returned a completed survey, generating a 12% response rate. Based on our inclusion criteria, these 970 adults, aged 18 to 75 years, and seen at least once in the past year at Indiana University Health were included in our study sample (see Fig. 1 for a simplified representation of the sampling process by a flowchart). Non-English-speaking individuals were excluded from our study. Further details of the survey methodology have been published elsewhere.^[17]

The respondents’ zip codes were used to assign them a Rural-Urban Commuting Area (RUCA) code (ranging from 1.0 Metropolitan area to 10.6 Rural area), based on geographically linked measures of population density, urbanization, and daily commute.^[18,19] Categorization C of the RUCA taxonomy, developed at the University of Washington in collaboration with many state and federal agencies, was then used to map the RUCA codes to a binary rural or urban classification for each residence.^[8]

To ensure the quality of the survey, returned paper surveys were tagged with serial numbers and dispositions entered into a tracking database. Quality control tests were done continuously to verify scanning precision and data accuracy. Staff examined completed surveys for illegible marks, corrected them when necessary to enhance data capture, and digitally scanned the forms.

2.2. Survey instrument

Socio-demographic characteristics measured in the survey included age, sex, race, education level, marital status, insurance status, income, home ownership status, and employment status. Perceived rural-urban status was measured by 6 questions, each using a 7-point Likert response (not at all, barely, a little, about half, pretty much, very much, extremely), hereafter called “Rurality Scale.” The measure of self-reported rural status was adapted from a measure of rural and community identity fielded by Paskett et al^[11] Drs Haggstrom and Rawl, together with Dr

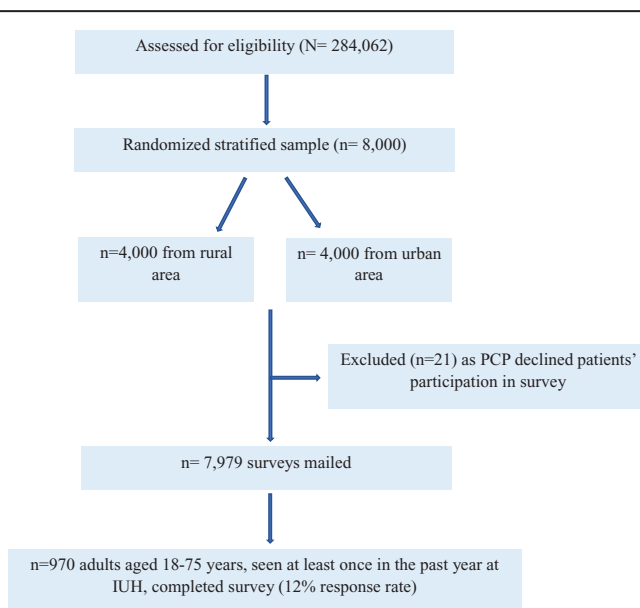


Figure 1. Flowchart to describe the sampling process.

Paskett, were supported by an NCI-led initiative providing supplemental funds to conduct research to better characterize the population catchment areas served by NCI-designated cancer centers (<https://cancercontrol.cancer.gov/brp/hcirb/catchment-areas.html>).

These questions were constructed to ask each participant,

- “How much do you see yourself belonging to a rural community?”
- “How much is being from a rural community a part of who you are?”
- “How much do you identify with people who live in rural communities?”
- “To what extent do you feel your general attitudes and opinions are similar to people who live in rural communities?”
- “To what extent do you feel that you are typical of people who live in rural communities?”
- “To what extent do you consider yourself a ‘city’ person?”

The last question regarding perceived “city person” status was reverse coded to be aligned with the responses in the remaining 5 questions.

2.3. Statistical analysis

Descriptive statistics were performed on individual socio-demographic characteristics, their designated RUCA category, and perceived rural-urban status.

To measure the internal consistency of the Rurality Scale, we computed Cronbach alpha (α), a widely used measure of internal consistency or reliability in psychometric research.^[20–22] α ranges from 0 to 1, with 0 indicating no consistency between the items and close to 1 indicating highly consistent items.^[23,24] To examine the construct validity of the Rurality Scale, we used exploratory factor analysis based on the polychoric correlation matrix. This method determined the principal components (rurality or urbanicity) best accounting for the variability in the data.^[25] Next, we performed varimax rotation to minimize the complexity of the factor loadings obtained from principal component factor analysis; this makes the structure simpler to interpret. The exploratory factor analysis validated 3 underlying constructs of the response items; “rural attitude” (formed by 3 questions), “rural identity” (formed by 2 questions), and “urban identity” (formed by 1 question).

To measure the strength and direction of association between the ranked measures, RUCA-based rural-urban status (binary scale, a special case of ordinal scale) and the self-reported perceived rurality (Likert scale, ordinal in nature), we used the Spearman rank correlation coefficient (ρ). ρ ranges from -1 to $+1$, indicating a perfect negative association and a perfect positive association, respectively, with 0 indicating no association at all.^[26,27] Cohen (1988) guidelines were used to interpret the effect size.^[28]

Lastly, the concordance of the ranked measures was calculated using the Gwet agreement coefficient (Gwet AC), a measure of correlation, defined as the conditional probability of 2 randomly chosen observational measurements to agree, given no agreement by chance. This was calculated following Gwet new framework of chance-corrected inter-rater agreement coefficients weighted ordinally, extending all existing agreement coefficients to include multiple raters, multiple rating categories, any level of measurement, and numerous ratings per subject. Gwet AC is interpreted based on the guidelines developed by Landis and Koch.^[29,30] Violin plots were also generated for each response item to visualize the data distribution and its probability density across zip code based RUCA category (see Appendix Fig. 1, Supplemental Content, <http://links.lww.com/MD/J808> which reports violin plots for survey response items distributed across zip code based RUCA category).

Association and concordance were calculated between the following groups:

- RUCA categorization of geographic status and self-reported perceived rurality (Rurality Scale of all 6 items individually).
- RUCA categorization of geographic status and self-reported perceived rurality (Rurality subscale of 3 items combined under the underlying construct of “rural attitude”).
- RUCA categorization of geographic status and self-reported perceived rurality (Rurality subscale of 2 items combined under the underlying construct of “rural identity”).
- RUCA categorization of geographic status and self-reported perceived rurality (Rurality Scale of all 6 items combined to get a total Rurality Scale score).

Details about the construction of the subscale scores and the total scale score are provided in Appendix Figure 2 (see Appendix Fig. 2, Supplemental Content, <http://links.lww.com/MD/J810> which illustrates Stata code for construction of combined rurality scores). The analyses were performed in Stata version-16.1 (StataCorp LLC, College Station, TX).

3. Results

3.1. Descriptive statistics

All 970 patients who completed the survey responded to their perceived rural status. The participants had a mean age of 56.9 years, were most often male (54%), white (77%), had college or post-graduate education (36%), were partnered (61%), had health insurance (94%), reported an income range of \$50,000 to \$99,999 (32%), owned their own home (68%), and were employed (41%). Participants’ characteristics were also distributed across the RUCA categorization of geographic status (Table 1).

According to the RUCA category, the majority of the participants lived in a rural area (54%), while 46% lived in an urban area. For perceived rural-urban status, approximately 24% neither saw themselves belonging to a rural community nor felt being from a rural community is a part of who they are. However, when it comes to attitude towards rural-urban status, 21% identified with people who lived in rural communities, 22% felt their general attitudes and opinions were similar to people living in rural communities and felt typical of people living in rural communities. Around 13% very much considered themselves a city person (Table 2).

3.2. Internal consistency

In this study, the Cronbach alpha obtained was 0.907 indicating high internal consistency among the 6 self-perceived rurality items. Each item was correlated with the overall Rurality scale ranging from 0.696 to 0.908. If items were deleted, Cronbach α value did not change much; except that it increased from 0.907 to 0.919 when the item for the “city” person was removed. Hence, “City person” did not correlate as strongly with the other items (Table 3). The 2 survey questions comprising the “rural identity” subscale and the 3 questions comprising the “rural attitude” subscale showed high internal consistency with Cronbach α being 0.896 and 0.905, respectively.

3.3. Factor analysis

The polychoric correlation matrix showed high correlations among all 6 items on self-perceived rurality (see Appendix Table 1, Supplemental Content, <http://links.lww.com/MD/J811> which illustrates polychoric correlation matrix of survey response items). Based on this, the first 3 unrotated factor loadings extracted from the principal component factor analysis

Table 1
Distributions of participants' characteristics stratified by zip code-derived RUCA category.

Sample characteristics	Overall (970) n (%)	Rural (523) n (%)	Urban (447) n (%)
Age			
Mean (SD)	56.9 (14.3)	57.6 (14.1)	56.0 (14.4)
Categories	105 (10.8)	50 (47.6)	55 (52.4)
18–34	130 (13.4)	73 (56.2)	57 (43.8)
35–49	380 (39.2)	193 (50.8)	187 (49.2)
50–64	355 (36.6)	207 (58.3)	148 (41.7)
65+			
Sex			
Male	522 (53.8)	278 (53.3)	244 (46.7)
Female	448 (46.2)	245 (54.7)	203 (45.3)
Race			
White	743 (76.6)	465 (62.6)	278 (37.4)
Black	192 (19.8)	43 (22.4)	149 (77.6)
Multi/other	35 (3.6)	15 (42.9)	20 (57.1)
Education			
<High school	71 (7.8)	38 (53.5)	33 (46.5)
High school graduate (or GED)	260 (28.4)	159 (61.2)	101 (38.8)
Post HS/some college	253 (27.6)	142 (56.1)	111 (43.9)
College graduate or higher	332 (36.2)	157 (47.3)	175 (52.7)
Missing	54 (5.6)	27 (5.2)	27 (6.0)
Marital status			
Partnered	571 (60.7)	342 (59.9)	229 (40.1)
Not partnered	369 (39.3)	164 (44.4)	205 (55.6)
Missing	30 (3.1)	17 (3.3)	13 (2.9)
Health insurance			
Yes	888 (94.2)	473 (53.3)	415 (46.7)
No	55 (5.8)	34 (61.8)	21 (38.2)
Missing	27 (2.8)	16 (3.1)	11 (2.5)
Income			
\$0–19,999	177 (20.5)	83 (46.9)	94 (53.1)
\$20,000–49,999	268 (31.0)	164 (61.2)	104 (38.8)
\$50,000–99,999	277 (32.0)	159 (57.4)	118 (42.6)
\$100,000+	143 (16.5)	55 (38.5)	88 (61.5)
Missing	105 (10.8)	62 (11.9)	43 (9.6)
Own home			
Own	638 (68.4)	379 (59.4)	259 (40.6)
Rent/occupy	295 (31.6)	124 (42.0)	171 (58.0)
Missing	37 (3.8)	20 (3.8)	17 (3.8)
Employed			
Yes	369 (41.1)	177 (48.0)	192 (52.0)
No	225 (25.1)	128 (56.9)	97 (43.1)
Retired	303 (33.8)	181 (59.7)	122 (40.3)
Missing	73 (7.5)	37 (7.1)	36 (8.1)

Results are listed as n (%), with n being the sample size unless otherwise indicated. For the overall sample, the column percentages are reported. For rural and urban sample, the row percentages are reported. The bold values are denoting missing values.

GED = General Educational Development, HS = high school, SD = standard deviation.

explained most of the variability in the data (92%). The percentage of variability explained by factor 1 (three questions) was 73.4%, by factor 2 (two questions) was 11.6% and by factor 3 (one question) was 7.3% (see Appendix Table 2, Supplemental Content, <http://links.lww.com/MD/J813> which reports unrotated principal component factor loadings).

Results showed that 3 items assessing the extent to which individuals “identified with people who lived in rural communities” (0.359), “feel general attitudes and opinions similar to people living in rural communities” (0.699), and “feel typical of people living in rural communities” (0.607) had substantial positive loadings on factor 1, thus representing the underlying construct of “rural attitude.” Two items that assessed the extent to which individuals “see themselves belonging to a rural

community” (0.747), and “being from a rural community is a part of who they are” (0.578) had substantial positive loadings on factor 2, thus representing the underlying construct of “rural identity.” Lastly, a single item that assessed the extent to which individuals “consider themselves a city person” (0.989) had a very large positive loading on factor 3, which was considered to represent the concept of “urban identity” (Table 4). This indicates that the Rurality Scale is multidimensional with 3 distinct dimensions.

3.4. Association and concordance

Spearman rank correlation coefficient showed overall moderate associations among RUCA designated rural-urban category and

Table 2
Distribution of survey rurality scale stratified by zip code-derived RUCA category.

Rurality scales	Overall (970) n (%)	Rural (523) n (%)	Urban (447) n (%)
“How much do you see yourself belonging to a rural community?”	233 (24.9)	52 (22.3)	181 (77.7)
Not at all	90 (9.6)	43 (47.8)	47 (52.2)
Barely	99 (10.6)	48 (48.5)	51 (51.5)
A little	107 (11.4)	62 (57.9)	45 (42.1)
About half	144 (15.4)	105 (72.9)	39 (27.1)
Pretty much	145 (15.5)	110 (75.9)	35 (24.1)
Very much	118 (12.6)	85 (72.0)	33 (28.0)
Extremely	34 (3.5)	18 (52.9)	16 (47.1)
Missing			
“How much is being from a rural community a part of who you are?”	224 (24.0)	46 (20.5)	178 (79.5)
Not at all	90 (9.6)	41 (45.6)	49 (54.4)
Barely	105 (11.2)	57 (54.3)	48 (45.7)
A little	107 (11.5)	67 (62.6)	40 (37.4)
About half	135 (14.5)	93 (68.9)	42 (31.1)
Pretty much	183 (19.6)	132 (72.1)	51 (27.9)
Very much	90 (9.6)	68 (75.6)	22 (24.4)
Extremely	36 (3.7)	19 (52.8)	17 (47.2)
Missing			
“How much do you identify with people who live in rural communities?”	77 (8.2)	15 (19.5)	62 (80.5)
Not at all	104 (11.1)	36 (34.6)	68 (65.4)
Barely	167 (17.9)	70 (41.9)	97 (58.1)
A little	131 (14.0)	75 (57.3)	56 (42.8)
About half	187 (20.0)	122 (65.2)	65 (34.8)
Pretty much	196 (21.0)	139 (70.9)	57 (29.1)
Very much	73 (7.8)	49 (67.1)	24 (32.9)
Extremely	35 (3.6)	17 (48.6)	18 (51.4)
Missing			
“To what extent do you feel your general attitudes and opinions are similar to people who live in rural communities?”	68 (7.3)	15 (22.1)	53 (77.9)
Not at all	95 (10.2)	36 (37.9)	59 (62.1)
Barely	172 (18.5)	72 (41.9)	100 (58.1)
A little	201 (21.6)	114 (56.7)	87 (43.3)
About half	206 (22.2)	136 (66.0)	70 (34.0)
Pretty much	141 (15.2)	97 (68.8)	44 (31.2)
Very much	46 (5.0)	34 (73.9)	12 (26.1)
Extremely	41 (4.2)	19 (46.3)	22 (53.7)
Missing			
“To what extent do you feel that you are typical of people who live in rural communities?”	132 (14.2)	27 (20.5)	105 (79.6)
Not at all	131 (14.1)	64 (48.9)	67 (51.2)
Barely	137 (14.7)	62 (45.3)	75 (54.7)
A little	157 (16.9)	91 (58.0)	66 (42.0)
About half	203 (21.8)	139 (68.5)	64 (31.5)
Pretty much	125 (13.5)	86 (68.8)	39 (31.2)
Very much	45 (4.8)	36 (80.0)	9 (20.0)
Extremely	40 (4.1)	18 (45.0)	22 (55.0)
Missing			
“To what extent do you consider yourself a ‘city’ person?”			
Extremely	212 (22.5)	13 (15.9)	69 (84.2)
Very much	128 (13.6)	34 (27.6)	89 (72.4)
Pretty much	123 (13.1)	44 (31.9)	94 (68.1)
About half	134 (14.3)	82 (61.2)	52 (38.8)
A little	138 (14.7)	79 (64.2)	44 (35.8)
Barely	123 (13.1)	97 (75.8)	31 (24.2)
Not at all	82 (8.7)	159 (75.0)	53 (25.0)
Missing	30 (3.1)	15 (50.0)	15 (50.0)

Results are listed as n (%), with n being the sample size unless otherwise indicated. For the overall sample, the column percentages are reported. For rural and urban sample, the row percentages are reported. Rurality scale for self-perceived responses is Likert in nature (ordinal), the zip-code based RUCA category scale is binary in nature (a special case of ordinal scale).

all of the 6 items on self-perceived rurality (ρ range: 0.28–0.41; $P < .001$). For the 2 individual items under the “rural identity” construct, Gwet AC was 0.19 (95% CI: 0.13–0.26) for individuals’ self-perceived notion of “seeing themselves belonging to a rural community,” and 0.20 (95% CI: 0.14–0.26) for “being from a rural community is a part of who they are” when compared to their corresponding RUCA categorization. This indicates slight agreement between the 2 measures based on Gwet probabilistic benchmarking method to the Landis and Koch reliability scale. Overall, the 2 items comprising the “rural identity” subscale when compared with the RUCA categorization showed moderate association ($\rho = 0.41$; P value $< .001$) and slight concordance (Gwet AC: 0.08 (95% CI: 0.01–0.15)) (Table 5, Panel 1).

For the 3 individual items under the “rural attitude” construct, Gwet AC between RUCA codes of geographic area and individuals’ self-perceived notion of “identifying with people who lived in rural communities,” “feeling general attitudes and opinions similar to people living in rural communities,” and “feeling typical of people living in rural communities” were 0.03 (95% CI: –0.03, 0.09), 0.13 (95% CI: 0.07–0.18) and 0.26 (95% CI: 0.21–0.32), respectively, all indicating slight to a fair agreement between the 2 measures. Overall, the 3 items comprising the “rural attitude” subscale when compared with the RUCA categorization showed moderate association ($\rho = 0.33$; P value $< .001$) and poor concordance (Gwet AC: –0.06 (95% CI: –0.12, 0.001)) (Table 5, Panel 2).

Lastly, the Gwet AC between RUCA codes and the single item under the “urban identity” construct assessing participants’ perception of themselves as “being a city person” was –0.11 (95% CI: –0.18, –0.04), indicating poor agreement between the 2 measures (Table 5, Panel 3).

Upon computing a Rurality Scale total score based on responses to all 6 survey items, we still found a moderate association ($\rho = 0.427$; P value $< .001$) and poor concordance with the RUCA-based codes (Gwet AC: –0.12 (95% CI: –0.19, –0.06)). Thus overall, the concordance between the RUCA codes and self-report of geographic status ranged from –0.11 to 0.26, showing poor to a fair agreement, with a moderate association between the 2 measures ranging from 0.28 to 0.41 (Table 5, Panel 4).

4. Discussion

The present study assessed the level of association and concordance between traditional geography-based RUCA codes and patient-reported perceived rural status. Our findings were discovered among a Midwestern population with access to health-care. Our results provide strong evidence of internal consistency of the Rurality Scale with the Cronbach alpha as 0.907. This finding suggests that the response items deployed in the survey capture the underlying construct of rural status adequately. Additionally, good construct validity was demonstrated by the exploratory factor analysis based on the polychoric correlation matrix, which showed that the Rurality Scale was multidimensional, composed of 3 factors/dimensions (“rural attitude,” “rural identity,” and “urban identity”) that explained 92% of the variability in the data. However, there was a lack of strong association and concordance between RUCA codes and every self-responded survey item, when tested separately, when combined overall and when looked within each constructed subscale.

Historically there has been a lack of literature on the relationship between self-perceived rural-urban status and geography-based RUCA designation. Our study provides empirical evidence that patients’ views about their rural-urban identity or attitude do not necessarily align with their residential-based rural-urban designations. Similar evidence of a lack of strong agreement between geographic and self-reported methods of ascertaining rurality was also found in the paper by Onega et

al, 2020.^[9] Based on RUCA designation, both our study population and the population in Onega et al 2020 were substantially rural (54% and 50%, respectively). However, our study has some unique features. While the Onega et al sample covered New Hampshire and Vermont in the New England region which is nearly 50% rural, our sample is based in the industrial Midwest (Indiana) which is 14% rural^[31] and thus more closely matches the rural distribution of the wider U.S. rural population, which was estimated to be 19.3% according to the 2010 Census.^[32] The Onega et al study collected data via an internet survey using the online platform Amazon Mechanical Turk, which introduces certain limitations in representativeness, participation, generalization, and data quality.^[33] Since our data were collected through a paper and pencil survey sent via U.S. mail, our method avoided the digital divide faced disproportionately by older people. This difference is reflected in the mean age of our study populations (56.9 vs 18–34 years old). Methodologically, we have employed a rigorous approach of psychometric analysis to firstly, evaluate the relevance of the 6 response items related to the Rurality scale and their consistency with each other and secondly, examine the underlying multidimensional nature of the scale, thus generating more robust and reliable results. Overall, despite the differences in survey populations and methodology, both studies found a lack of strong concordance between geographic-based RUCA codes and self-report of rurality, reinforcing the robustness of these findings.

In our study, the factor loadings helped us in understanding the underlying constructs of “rural identity,” “rural attitude,” and “urban identity.” Identity, defined as “who you are,” is often viewed as a social construct. The definition amalgamates many social elements such as race, ethnicity, religion, education, background, past experiences, and future aspirations that help in shaping a human being. It depicts an individual “feeling of sense” and how they exude this feeling to their surroundings. “Attitude” is more of a cognitive construct that allows one to make mental decisions about how to behave rather than to focus on how the mental decisions affect behavior.^[34] Based on this conceptualization of these 2 concepts, we found the response items where individuals “see themselves belonging to a rural community” and “being from a rural community is a part of who they are” to fit appropriately with the construct of “rural identity” because these 2 items reflect a sense as to “who you are.” Similarly, the response items where an individual “identified with people who lived in rural communities,” “feel general attitudes and opinions similar to people living in rural communities,” and “feel typical of people living in rural communities” fit well with the underlying construct of “rural attitude” because these 3 items reflect how one may make decisions or behave in the community. Based on the same argument for “rural identity,” the extent to which the individuals “consider themselves a city person” also fits well with the underlying construct of “urban identity” because this item again reflects a sense of “who you are.” Given the multidimensional nature of the Rurality Scale, caution is needed when interpreting the total Rurality Scale scores or when using it in analyses. Using the 3 subscale scores identified in this study may benefit future research where perceived rurality is of interest.

Our findings are essential to the health sciences, given the highly prevalent use of RUCA categories, commonly based upon residential zip code, in the fields of epidemiologic, social/behavioral, and health services research. These rural-urban geographic designations are often used as a proxy to help understand individuals’ livelihood and social conditions that affect domains such as health behaviors, access to health care services, and socioeconomic needs. However, individuals who live in areas of relatively high population density may still self-identify as rural. Those living in

rural locations may self-identify as urban, even while facing longer travel distances to healthcare facilities associated with rural residence. These unique measures of rurality may differentially confound the effect pathways between social, political, or religious beliefs and particular health-related behaviors and subsequent health outcomes. Individuals with rural or urban identities or attitudes may have different perspectives when it comes to social, political, or religious issues, with increasing rurality being associated with more conservative views, such as COVID-19 vaccine hesitancy^[35]; subsequently, lack of receipt of the COVID-19 vaccine has been associated with worse outcomes in terms of hospitalization,^[36] and mortality.^[37] Similarly, there are observed variations in health-related behaviors and outcomes across the geographically defined rural-urban continuum; rural areas tend to have higher rates of risky behaviors such as smoking and low physical activity, which lead to higher morbidity and mortality rates such as for cancer.^[38–41] Hence, understanding the similarities and differences between geographic and perceived rural-urban status is essential to interpreting health-related studies.

Designing health education programs, tailored to the unique needs and perspectives of individuals regarding rurality, might also be a useful approach in facilitating adoption of health-promoting behaviors. For example, implementing policies that incentivize the development of walking and biking paths and community gardens in geographically rural areas, while at the same time utilizing culturally appropriate messaging and approaches might be a combined approach in facilitating adoption of health-promoting behaviors.

From the methodological aspect, as the measure of concordance, we used Gwet AC instead of Cohen/Conger Kappa (κ). There are some statistical issues with Cohen Kappa; unlike Gwet AC, it assumes that raters are independent, generating much of the agreement due to chance, which is not entirely accurate. In addition, Kappa values suffer from the “Kappa

paradox”; they tend to change a lot as prevalence changes, that is, the values become higher and closer to the percentage agreement when prevalence is high (the prevalence problem). Kappa values are also impacted by the degree of disagreement between observers (the bias problem).^[42,43] As Gwet AC tends to minimize the Kappa limitations, we consider this coefficient to be more stable as an inter-rater reliability coefficient than Cohen/Conger Kappa in our study, following other recent studies that preferred Gwet AC over Cohen/Conger Kappa as a more stable coefficient.^[44–46]

5. Limitations

Some study limitations must be considered when interpreting these results. First, our survey response rate was relatively low, at 12%, despite using established methods for survey research. The survey sample included younger and working adults who are less likely to respond to mailings.^[47] Due to our expectation of a low response rate to a mailed survey, enough surveys were sent to ensure a relatively large absolute number of surveys among rural individuals, one of the key groups among whom the stratified sample was targeted. Other data collection methods, such as in-person or telephone interviews, might have improved our participation rate. Nonetheless, we received completed surveys from every surveyed county.^[48] Second, our sample was drawn from a large academic health care system in the state of Indiana. Therefore, the results of our study on the relation between self-identified rurality and geographic location should be interpreted in light of statewide research; these cannot be generalized to the population. However, it is important to understand regional variation as it is more likely to capture underlying social constructs than a national sample.^[49] Moreover, the study findings can be interpreted as an addition to a growing body of research focused on individual states, which may be used to complement aggregate results from national surveys.

Table 3
Internal consistency of survey response items.

Survey response item	Item-total correlation	Cronbach alpha if item deleted
“How much do you see yourself belonging to a rural community?”	0.852	0.888
“How much is being from a rural community a part of who you are?”	0.908	0.874
“How much do you identify with people who live in rural communities?”	0.880	0.881
“To what extent do you feel your general attitudes and opinions are similar to people who live in rural communities?”	0.805	0.894
“To what extent do you feel that you are typical of people who live in rural communities?”	0.851	0.886
“To what extent do you consider yourself a ‘city’ person?”	0.696	0.919
Total perceived rurality scale		0.907

Rurality scale for self-perceived responses is Likert in nature (ordinal). Refer to Appendix Figure 2 for details of total perceived Rurality scale construction.

Table 4
Rotated Principal component factor loadings.

Principal components/ correlation	Variance	Proportion	Cumulative percentage
Component			
Comp1	2.354	0.392	0.392
Comp2	2.158	0.360	0.752
Comp3	1.028	0.171	0.923
Varimax rotated components			
Items	Rural attitude	Rural identity	Urban attitude
“How much do you see yourself belonging to a rural community?”	–0.098	0.747	–0.035
“How much is being from a rural community a part of who you are?”	0.055	0.578	0.072
“How much do you identify with people who live in rural communities?”	0.359	0.319	–0.048
“To what extent do you feel your general attitudes and opinions are similar to people who live in rural communities?”	0.699	–0.076	–0.064
“To what extent do you feel that you are typical of people who live in rural communities?”	0.607	–0.033	0.095
“To what extent do you consider yourself a ‘city’ person?”	–0.003	–0.002	0.989

Rurality scale for self-perceived responses is Likert in nature (ordinal). Component/factor 1 = rural attitude, component/factor 2 = rural identity, component/factor 3 = urban identity.

Table 5

Association and Concordance between zip code-derived RUCA category and self-reported geographic status.

	Spearman rank correlation		Gwet agreement coefficient	
	Coefficient (95% CI)	Scale	Coefficient (95% CI)	Benchmark scale for reliability (Gwet AC)
Rural identity scales (from SR)	0.412*** (0.358, 0.467)	Moderate	0.080** (0.009, 0.151)	Slight
"How much do you see yourself belonging to a rural community?"	0.392*** (0.333, 0.450)	Moderate	0.192*** (0.126, 0.259)	Slight
"How much is being from a rural community a part of who you are?"	0.391*** (0.334, 0.448)	Moderate	0.200*** (0.136, 0.265)	Slight
Rural attitude scales (from SR)	0.328*** (0.269, 0.385)	Moderate	-0.060** (-0.119, 0.001)	Poor
"How much do you identify with people who live in rural communities?"	0.309*** (0.248, 0.369)	Moderate	0.031 (-0.032, 0.093)	Slight
"To what extent do you feel your general attitudes and opinions are similar to people who live in rural communities?"	0.284*** (0.222, 0.345)	Moderate	0.129*** (0.073, 0.184)	Slight
"To what extent do you feel that you are typical of people who live in rural communities?"	0.315*** (0.254, 0.376)	Moderate	0.265*** (0.209, 0.319)	Fair
Urban attitude scales (from SR)				
"To what extent do you consider yourself a 'city' person?"	0.413*** (0.359, 0.468)	Moderate	-0.108*** (-0.179, -0.037)	Poor
Combined scale (all 6 items)	0.427*** (0.372, 0.483)	Moderate	-0.125*** (-0.187, -0.063)	Poor

Correlation coefficients are interpreted as: <0.1: "insubstantial correlation"; 0.1 to 0.3: "small correlation"; 0.3 to 0.5: "moderate correlation"; >0.5: "large correlation" as per Cohen (1988).^[21] Gwet probabilistic benchmarking method according to the Landis and Koch scale of reliability is: -1.0 to 0.0 (Poor), 0.0-0.2 (Slight), 0.2-0.4 (Fair), 0.4-0.6 (Moderate), 0.6-0.8 (Substantial), 0.8-1.0 (Almost Perfect) (Gwet, 2014).^[22] ***, **, * denote statistical significance at the 1%, 5%, and 10% levels.

AC = agreement coefficient, CI = confidence interval, Gwet AC = Gwet agreement coefficient, SR = survey response.

6. Conclusion

Concordance between RUCA codes and self-report of rural status ranged between poor and fair agreement, suggesting that they may have independent meaning and value in measuring rurality. Future studies should consider these differences when assessing health behavior and outcomes among rural populations, recognizing that geographic-based and self-report methods may provide complementary information. This knowledge may facilitate targeted interventions and policies to address specific challenges and barriers to support rural communities.

Author contributions

Conceptualization: Oindrila Bhattacharyya, Susan M. Rawl, David A. Haggstrom.

Data curation: Susan M. Rawl, David A. Haggstrom.

Funding acquisition: Susan M. Rawl, David A. Haggstrom.

Formal analysis: Oindrila Bhattacharyya, Stephanie L. Dickinson, David A. Haggstrom.

Investigation: Oindrila Bhattacharyya, David A. Haggstrom.

Methodology: Oindrila Bhattacharyya, Susan M. Rawl, Stephanie L. Dickinson, David A. Haggstrom.

Project administration: Oindrila Bhattacharyya, David A. Haggstrom.

Resources: Oindrila Bhattacharyya, David A. Haggstrom.

Software: Oindrila Bhattacharyya, David A. Haggstrom.

Supervision: Oindrila Bhattacharyya, Susan M. Rawl, David A. Haggstrom.

Validation: Oindrila Bhattacharyya, David A. Haggstrom.

Visualization: Oindrila Bhattacharyya, David A. Haggstrom.

Writing – original draft: Oindrila Bhattacharyya, David A. Haggstrom.

Writing – review & editing: Oindrila Bhattacharyya, Susan M. Rawl, Stephanie L. Dickinson, David A. Haggstrom.

References

- [1] Sharkey JR, Johnson CM, Dean WR. Less-healthy eating behaviors have a greater association with a high level of sugar-sweetened beverage consumption among rural adults than among urban adults. *Food Nutr Res.* 2011;55:5819.
- [2] Parks SE, Housemann RA, Brownson RC. Differential correlates of physical activity in urban and rural adults of various socioeconomic backgrounds in the United States. *J Epidemiol Community Health.* 2003;57:29–35.
- [3] Joens-Matre RR, Welk GJ, Calabro MA, et al. Rural–Urban differences in physical activity, physical fitness, and overweight prevalence of children. *J Rural Health.* 2008;24:49–54.
- [4] Nighbor TD, Doogan NJ, Roberts ME, et al. Smoking prevalence and trends among a U.S. national sample of women of reproductive age in rural versus urban settings. Cassidy RN, ed. *PLoS One.* 2018;13:e0207818.
- [5] Doescher MP, Jackson JE, Jerant A, et al. Prevalence and Trends in Smoking: a National Rural Study. *J Rural Health.* 2006;22:112–8.
- [6] Borders TF, Booth BM. Rural, suburban, and Urban variations in alcohol consumption in the United States: findings from the National Epidemiologic Survey on Alcohol and Related Conditions. *J Rural Health.* 2007;23:314–21.
- [7] Probst JC, Zahnd WE, Hung P, et al. Rural-Urban mortality disparities: variations across causes of death and race/ethnicity, 2013–2017. *Am J Public Health.* 2020;110:1325–7.
- [8] Rural Urban Commuting Area Codes Data. Accessed February 19, 2022. <http://depts.washington.edu/uwruca/ruca-uses.php>
- [9] Onega T, Weiss JE, Alford-Teaster J, et al. Concordance of Rural-Urban self-identity and ZIP code-derived Rural-Urban Commuting Area (RUCA) Designation. *J Rural Health.* 2020;36:274–80.
- [10] Najafi M. The Concept of Place and Sense of Place in Architectural Studies. Published online 2011.
- [11] Paskett ED, Young GS, Bernardo BM, et al. Correlates of rural, appalachian, and Community Identity in the CITIES Cohort. *J Rural Health.* 2019;35:167–75.
- [12] Waymack JR, Markwell S, Milbrandt JC, et al. Comparison of rates of emergency department procedures and critical diagnoses in metropolitan and rural hospitals. *Rural Remote Health.* 2015;15:3298.
- [13] Inagami S, Gao S, Karimi H, et al. Adapting the Index of Relative Rurality (IRR) to estimate rurality at the ZIP code level: a rural classification system in Health Services Research. *J Rural Health.* 2016;32:219–27.
- [14] Khairat S, Haitcoat T, Liu S, et al. Advancing health equity and access using telemedicine: a geospatial assessment. *J Am Med Inform Assoc.* 2019;26:796–805.
- [15] Schroeder SM, Peterson ML. Identifying variability in patient characteristics and prevalence of Emergency Department Utilization for Mental Health Diagnoses in Rural and Urban Communities. *J Rural Health.* 2018;34:369–76.
- [16] Zahnd WE, Del Vecchio N, Askelson N, et al. Definition and categorization of rural and assessment of realized access to care. *Health Serv Res.* 2022;57:693–702.

- [17] Haggstrom DA, Lee JL, Dickinson SL, et al. Rural and Urban differences in the adoption of New Health Information and Medical Technologies. *J Rural Health*. 2019;35:144–54.
- [18] Hart LG, Larson EH, Lishner DM. Rural Definitions for Health Policy and Research. *Am J Public Health*. 2005;95:1149–55.
- [19] USDA ERS - Rural-Urban Commuting Area Codes. Accessed February 19, 2022. <https://www.ers.usda.gov/data-products/rural-urban-commuting-area-codes/>
- [20] Tang Z, Li X, Zhang Y, et al. Psychometric analysis of a Chinese version of the Sleep Hygiene Index in nursing students in China: a cross-sectional study. *Sleep Med*. 2021;81:253–60.
- [21] Ghanbari S, Ramezankhani A, Montazeri A, et al. Health literacy measure for Adolescents (HELMA): development and psychometric properties. Brucki S, ed. *PLoS One*. 2016;11:e0149202.
- [22] Wortmann JH, Jordan AH, Weathers FW, et al. Psychometric analysis of the PTSD Checklist-5 (PCL-5) among treatment-seeking military service members. *Psychol Assess*. 2016;28:1392–403.
- [23] Cronbach LJ. Coefficient alpha and the internal structure of tests. *Psychometrika*. 1951;16:297–334
- [24] Tavakol M, Dennick R. Making sense of Cronbach's alpha. *Int J Med Educ*. 2011;2:53–5.
- [25] Finch AP, Brazier JE, Mukuria C, et al. An exploratory study on using principal-component analysis and confirmatory factor analysis to identify bolt-on dimensions: the EQ-5D case study. *Value Health*. 2017;20:1362–75.
- [26] Kendall MG. Rank Correlation Methods. Griffin; 1948.
- [27] Lange T, Kopkow C, Lütznier J, et al. Comparison of different rating scales for the use in Delphi studies: different scales lead to different consensus and show different test-retest reliability. *BMC Med Res Methodol*. 2020;20:28.
- [28] Cohen J. *Statistical Power Analysis for the Behavioral Sciences*. 2nd ed. Routledge; 1988. doi:10.4324/9780203771587
- [29] Gwet KL. *Handbook of Inter-Rater Reliability*, 4th Edition: The Definitive Guide to Measuring the Extent of Agreement Among Raters. Advanced Analytics, LLC; 2014.
- [30] Klein D. Implementing a General Framework for Assessing Interrater Agreement in Stata. *Stata J*. 2018;18:871–901.
- [31] STATS Indiana. Accessed February 19, 2022. <http://www.stats.indiana.edu/>
- [32] Bureau UC. *Census.gov*. Accessed February 19, 2022. <https://www.census.gov/en.html>
- [33] Stritch JM, Pedersen MJ, Taggart G. The opportunities and limitations of using Mechanical Turk (MTURK) in public administration and management scholarship. *Int Public Manag J*. 2017;20:489–511.
- [34] Hallajow N. Identity and attitude: eternal conflict or harmonious coexistence. *J Soc Sci*. 2018;14:43–54.
- [35] Muñana C. 2021. Vaccine Hesitancy in Rural America. KFF. Published January 7, 2021. Accessed February 19, 2022. <https://www.kff.org/coronavirus-covid-19/poll-finding/vaccine-hesitancy-in-rural-america/>
- [36] Bozio CH, Grannis SJ, Naleway AL, et al. Laboratory-Confirmed COVID-19 Among Adults Hospitalized with COVID-19-Like Illness with Infection-Induced or mRNA Vaccine-Induced SARS-CoV-2 Immunity — Nine States, January–September 2021. *Morb Mortal Wkly Rep*. 2021;70:1539–44.
- [37] Xu S, Huang R, Sy LS, et al. COVID-19 Vaccination and Non-COVID-19 Mortality Risk — Seven Integrated Health Care Organizations, United States, December 14, 2020–July 31, 2021. *Morb Mortal Wkly Rep*. 2021;70:1520–4.
- [38] Henley SJ, Anderson RN, Thomas CC, et al. Invasive Cancer Incidence, 2004–2013, and Deaths, 2006–2015, in Nonmetropolitan and Metropolitan Counties — United States. *MMWR Surveill Summ*. 2017;66:1–13.
- [39] Matthews KA, Croft JB, Liu Y, et al. Health-related behaviors by urban-rural county classification — United States, 2013. *MMWR Surveill Summ*. 2017;66:1–8.
- [40] Zahnd WE, James AS, Jenkins WD, et al. Rural–urban differences in cancer incidence and trends in the United States. *Cancer Epidemiol Biomarkers Prev*. 2018;27:1265–74.
- [41] Blake KD, Moss JL, Gaysynsky A, et al. Making the case for investment in rural cancer control: an analysis of rural cancer incidence, mortality, and funding trends. *Cancer Epidemiol Biomarkers Prev*. 2017;26:992–7.
- [42] Feinstein AR, Cicchetti DV. High agreement but low Kappa: I. the problems of two paradoxes. *J Clin Epidemiol*. 1990;43:543–9.
- [43] Dettori JR, Norvell DC. Kappa and Beyond: is there agreement? *Glob Spine J*. 2020;10:499–501.
- [44] Wongpakaran N, Wongpakaran T, Wedding D, et al. A comparison of Cohen's Kappa and Gwet's AC1 when calculating inter-rater reliability coefficients: a study conducted with personality disorder samples. *BMC Med Res Methodol*. 2013;13:61.
- [45] Jimenez AM, Zepeda SJ. A comparison of Gwet's AC1 and Kappa when calculating inter-rater reliability coefficients in a teacher evaluation context. *J Educ Hum Resour*. 2020;38:290–300.
- [46] Cibulka MT, Strube MJ. The conundrum of Kappa and why some musculoskeletal tests appear unreliable despite high agreement: a comparison of Cohen Kappa and Gwet AC to assess observer agreement when using nominal and ordinal data. *Phys Ther*. 2021;101:pzab150.
- [47] Fowler F. *Survey Research Methods*. Sage Publications; 2013.
- [48] Rawl SM, Dickinson S, Lee JL, et al. Racial and socioeconomic disparities in cancer-related knowledge, beliefs, and behaviors in Indiana. *Cancer Epidemiol Biomarkers Prev*. 2019;28:462–70.
- [49] Portnoy B, Lee SJC, Kincheloe J, et al. Independent state health surveys: responding to the need for local population health data. *J Public Health Manag Pract*. 2014;20:E21–33.