Assessing fitness for use of two indicators of the rural-urban environment in the NAACCR data files

Kevin A Henry (Temple University)
Recinda Sherman (NAACCR)
Dave Stinchcomb (Westat)
Introduction

- Understanding how cancer risks vary both socially and geographically is important in efforts to plan and improve cancer prevention.

- Some past studies have shown that residents of rural areas have lower screening rates, higher late-stage diagnosis rates, and differences in cancer treatment patterns than urban residents.

- Study findings by rural/urban residence have been inconsistent across cancer sites likely due to variations in how urban/rural was defined and the geographic unit used (e.g., county vs tract).
The majority of US nationwide rural/urban cancer studies have used county-level measures of rural/urban

Using a smaller geographic area (e.g., tract) to evaluate area-based influences on health is more precise than assigning area-based measures using county
Introduction

• In its November 2014 call for data, NAACCR began collecting 2 census tract-level urban/rural indicators in the CINA dataset (1995-2012 dx years)
  
1. Census Bureau’s Urban Rural Indicator Codes (URIC)

2. US Department of Agriculture’s (USDA) Rural Urban Commuting Area (RUCA) Codes

• Tract-level indicators of rural-urban residence in the CINA dataset will facilitate research in rural-urban disparities at the national level
Objective

• Provide an overview of how the NAACCR rural/urban indicators are derived

• Summarize their “fitness for use”

• Provide basic descriptive statistical summaries of these new measures.
Rural/Urban Indicators

• Census Bureau’s Urban Rural Indicator Codes (URIC)
  – Based is based on % of persons in tract living in urban area.
  – The URIC codes that are included in this NAACCR data item include
    • 1: all urban – the percent of the population in an urban area = 100%
    • 2: mostly urban – the percent of the population in an urban area < 100% and ≥ 50%
    • 3: mostly rural – the percent of the population in a rural area < 100% and > 50%
    • 4: all rural – the percent of the population in an rural area = 100%
    • 9: unknown or not applicable – census tract not available or tract population was zero at the last decennial census
<table>
<thead>
<tr>
<th>URIC Code</th>
<th>Description</th>
<th>Count</th>
<th>Total Population</th>
<th>Percent of Population</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>All Urban</td>
<td>45394</td>
<td>185840944</td>
<td>60.19</td>
</tr>
<tr>
<td>2</td>
<td>Mostly Urban</td>
<td>13035</td>
<td>68276029</td>
<td>22.11</td>
</tr>
<tr>
<td>3</td>
<td>Mostly Rural</td>
<td>5989</td>
<td>27050232</td>
<td>8.76</td>
</tr>
<tr>
<td>4</td>
<td>All Rural</td>
<td>8113</td>
<td>27578333</td>
<td>8.93</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td></td>
<td><strong>72531</strong></td>
<td><strong>308745538</strong></td>
<td></td>
</tr>
</tbody>
</table>
Rural/Urban Indicators

• USDA’s RUCA code
  – RUCA codes classify U.S. census tracts using measures of population density, urbanization, and daily commuting flows.
  – 2000 and 2010 RUCA codes available in CINA
  – The RUCA codes that included in this NAACCR data item include:
    • 1: urban commuting area – RUCA codes 1.0, 1.1, 2.0, 2.1, 3.0, 4.1, 5.1, 7.1, 8.1, and 10.1
    • 2: not an urban commuting area – all other RUCA codes except 99
    • 9: unknown or not applicable – census tract not available or RUCA code = 99
<table>
<thead>
<tr>
<th>RUCA Code</th>
<th>Description</th>
<th>Count</th>
<th>Total Population</th>
<th>Percent of Population</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Urban Commute Area</td>
<td>63112</td>
<td>273603951</td>
<td>88.62</td>
</tr>
<tr>
<td>2</td>
<td>Not Urban Commute Area</td>
<td>9419</td>
<td>35141587</td>
<td>11.38</td>
</tr>
<tr>
<td>Total</td>
<td></td>
<td>72531</td>
<td>308745538</td>
<td></td>
</tr>
</tbody>
</table>
NAACCR rural/urban indicators

• For cancer research involving residential context:
  – URIC measures the rural nature of the environment
    • May be most relevant for studies involving
      – Links with exercise and diet
      – Exposure to pollutants
      – Stress and crime levels
  – RUCAs measure the accessibility to large urban centers
    • May be most relevant for studies involving
      – Access to oncology specialists
      – Access to cancer treatment options
      – Knowledge of medical options, likelihood of seeking care
NAACCR rural/urban indicators

• New for fall 2014 submission (1995-2012 dx years)

• Currently not a NAACCR required field

• All cases, regardless of year of dx, derived for 2000, 2010 (2020, etc.) to support retrospective analysis & cross-sectional studies

• Rates by rural/urban codes requires tract-level denominators which are currently not part of CINA

• Rural/urban codes posted in Call for Data

• Variables/definitions should be in Vol II 2016
National Coverage:

- All but 8 registries submitted this item in 2014
  - 3 states “oversight” (2 have already resubmitted data but not included in this assessment)
  - 2 states do not yet geocode (getting trained on NAACCR Geocoder)
  - 1 state has not yet responded
  - 1 state misunderstood confidentiality issues; working to clarify
  - 1 state will not submit any geographic variables at this time (not even true county)

- Overall 88% of population currently available 1995-2012; 86% for 2010 tracts
National Coverage:

Table 1. Percentage of valid and unassigned URIC and RUCA codes among 2008-2012 cancer cases (86% of US Cancer Cases)

<table>
<thead>
<tr>
<th>Total Cases (1998-2012)</th>
<th>Valid URIC/RUCA code</th>
<th>No URIC/RUCA code assigned</th>
</tr>
</thead>
<tbody>
<tr>
<td>5,848,393</td>
<td>91.6 %</td>
<td>8.4%</td>
</tr>
</tbody>
</table>

Table 2. Distribution of URIC and RUCA codes by census tract certainty (2008-2012 cancer cases (86% of US Cancer Cases))

<table>
<thead>
<tr>
<th>Census Tract Certainty</th>
<th>Valid URIC/RUCA code</th>
<th>No URIC/RUCA code assigned</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>5,358,529</td>
<td>489,864</td>
</tr>
<tr>
<td>%</td>
<td>%</td>
<td>%</td>
</tr>
<tr>
<td>Full street</td>
<td>90.7</td>
<td>2.3</td>
</tr>
<tr>
<td>ZIP Code^</td>
<td>6.1</td>
<td>1.8</td>
</tr>
<tr>
<td>CT Certainty not assigned</td>
<td>3.2</td>
<td>95.9</td>
</tr>
</tbody>
</table>

^ includes ZIP+4, ZIP=2, ZIP only, ZIP of PO-Box, residential city/ZIP where city/ZIP has only one census tract assigned
Table 3. Distribution of URIC codes based on cases geocoded to 2010 census tract URIC Codes, 2008-2012 (86% of US Cancer Cases)*

<table>
<thead>
<tr>
<th>CINA URIC Codes</th>
<th>Total Cancer Cases (2008-2012)</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>1: all urban – the percent of the population in an urban area = 100%</td>
<td>3,227,005</td>
<td>60.2%</td>
</tr>
<tr>
<td>2: mostly urban – the percent of the population in an urban area &lt; 100% and ≥ 50%</td>
<td>1,183,652</td>
<td>22.1%</td>
</tr>
<tr>
<td>3: mostly rural – the percent of the population in a rural area &lt; 100% and &gt; 50%</td>
<td>487,291</td>
<td>9.1%</td>
</tr>
<tr>
<td>4: all rural – the percent of the population in an rural area = 100%</td>
<td>460,134</td>
<td>8.6%</td>
</tr>
<tr>
<td>9. UNKNOWN or not applicable – census tract not available or tract population was zero at the last decadal census</td>
<td>447</td>
<td>0.0%</td>
</tr>
</tbody>
</table>

*non-geocoded cases and cases with unassigned URIC codes are excluded n=489,864
National Coverage:

Table 4. Distribution of cases geocoded to 2010 census tract RUCA Codes, 2008-2012 (86% of US Cancer Cases) *

<table>
<thead>
<tr>
<th>CINA RUCA Codes</th>
<th>Total Cancer Cases (2008-2012)</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Urban commuting area - RUCA codes 1.0,1.1,2.0,2.1,3.0,4.1,5.1,7.1,8.1,10.1</td>
<td>4,790,811</td>
<td>89.4%</td>
</tr>
<tr>
<td>2. Not an urban commuting area - all other RUCA codes except 99</td>
<td>567,553</td>
<td>10.6%</td>
</tr>
<tr>
<td>•9: unknown or not applicable – census tract not available or RUCA code = 99</td>
<td>165</td>
<td>0.003%</td>
</tr>
</tbody>
</table>

*non-geocoded cases and cases with unassigned URIC codes are excluded n=489,864
Recommendations

• Rural/Urban codes ‘fit for use’ but must be used in conjunction with census tract certainty variables to determine appropriateness for research question

• Registries should follow the guidelines to geocode entire database to most current census boundary (all to 2010)

• Use NAACCR Geocoder (NAACCR currently training two registries)
Thank you—Questions?
References

USDA RUCA:

U. Wash. RUCA:
http://depts.washington.edu/uwrura/

U. Wash. RUCA categorizations:
http://depts.washington.edu/uwrura/ruca-uses.php

Census urban/rural:
https://www.census.gov/geo/reference/urban-rural.html