## Contents

### Letter from the Editor
Danette Clark, BS, RMA, AAS, CTR ................................................................. 3

### Original Articles
**Indicators of Socioeconomic Status for Individuals, Census Tracts, and Counties: How Well Do Measures Align for Demographic Subgroups?** ......................................................... 4  
Jennifer L. Moss, PhD; Norman J. Johnson; Mandi Yu; Sean F. Altekruse; Kathleen A. Cronin

**Improved Geocoding of Cancer Registry Addresses in Urban and Rural Oklahoma** .......................................................... 13  
Naci Dilekli; Amanda Janitz; Janis Campbell

### How I Do It
**Cutaneous Melanoma Abstracting** .............................................................. 21  
Melissa Riddle, CTR

### Raising the Bar
**Are You Ready for Curated Data and Skill Shifts?** ...................................... 24  
Michele Webb, CTR

### Spring 2020 Continuing Education Quiz .................................................... 26  
Cari Vida, BS, RHIA, CTR; Susanna Mitchell, BS, CTR

### Call for Papers .................................................................................................. 27

### Information for Authors .................................................................................. 28

---

*Journal of Registry Management* is published quarterly by the National Cancer Registrars Association. 
1330 Braddock Place, #520 Alexandria, VA 22314 
(703) 299-6640 (703) 299-6620 FAX 

Address change of address and subscription correspondence to: 
National Cancer Registrars Association 
1330 Braddock Place, #520 Alexandria, VA 22314 

Address all editorial correspondence to: 
Danette Clark, BS, RMA, AAS, CTR, Editor-in-Chief 
Email: JRMEditor@ncra-usa.org 
Telephone: (973) 971-5189

Letters to the Editor
Letters to the Editor must be signed and include address and telephone number; none will be published anonymously. Letters subject to editing.

Editorial Policy
Opinions expressed in articles appearing in *Journal of Registry Management* are those of the authors and not necessarily those of the National Cancer Registrars Association, or the editor of *Journal of Registry Management*.

Subscription to *Journal of Registry Management*
Subscription is a benefit of membership in the National Cancer Registrars Association. Individual subscriptions may be purchased for $40/year U.S. dollars. Single copies may be purchased for $12 U.S., $15 International, prepaid only.

Advertising Sales
Advertising is accepted with the understanding that products and services are professionally related and meet the ethical standards of public health practice. Acceptance by the Journal does not indicate or imply endorsement by the Journal or NCRA. Advertising sales should be directed to Marketing Coordinator, NCRA, 1330 Braddock Place, #520, Alexandria, VA 22314, (703) 299-6640, (703) 299-6620 FAX.

© 2020 National Cancer Registrars Association
The Journal of Registry Management is indexed in the National Library of Medicine's MEDLINE database. Citations from the articles indexed, the indexing terms (key words), and the English abstract printed in JRM are included and searchable using PubMed.

For your convenience, the Journal of Registry Management is indexed in the 4th issue of each year and on the Web (under “Resources” at http://www.ncra-usa.org/jrm). The 4th issue indexes all articles for that particular year. The Web index is a cumulative index of all JRM articles ever published.

Indexing
Dear Colleagues,

No one could have predicted that COVID-19 would impact not only the United States, but the world. As we move forward with “the new normal,” all health care workers have been, and continue to be, impacted by this pandemic. Cancer registries shifted to working remotely overnight. Some of the changes have been challenging, such as managing staff remotely and moving tumor conferences to virtual platforms. We have overcome those obstacles and have become stronger. One of our greatest resources is sharing information. Some of our fellow cancer registries may be struggling and need guidance on how to adapt during these unprecedented circumstances. We need to help each other during this difficult period. Several educational events have been canceled or postponed until next year. NCRA has moved the annual conference to September 20-23. For more information, visit http://www.ncra-usa.org/Conference/2020-Annual-Conference/2020-Conference-Information.

The American College of Surgeons has provided several resources that provide guidance for the COVID-19 pandemic at https://www.facs.org/quality-programs/cancer/events/past/covid19-webinars. They have also established a free, voluntary COVID-19 registry at https://www.facs.org/quality-programs/covid19-registry.

Several resources are available on NCRA’s Center for Cancer Registry Education regarding remote abstracting at http://www.cancerregistryeducation.org/other-education-no-ce-credit. Also, NCRA Education Foundation has telecommuting toolkits available at http://www.ncra-usa.org/Foundation/Links.

This issue of the Journal of Registry Management contains 2 original manuscripts and 2 articles. We start with a manuscript by Jennifer L. Moss, PhD, and colleagues that reviews socioeconomic status variables that may impact data. The next manuscript from Naci Dilekli, Amanda Janitz, and Janis Campbell examines the impact of improving geocoding in urban and rural Oklahoma.

The “How I Do It” section features an article from Melissa Riddle, CTR, that reviews the process of abstracting a site that can be tricky: melanoma.

In the “Raising the Bar” feature, Michele Webb, CTR, explores curated data and the concept of skill shifts.

The last 2 pages of the JRM contain the Call for Papers and Information for Authors. Submissions of manuscripts and articles are accepted at any time. The “How I Do It” section comes from readers who want to share their expertise and ideas on varying topics.

Special Call for Articles

The Commission on Cancer (CoC) has released the 2020 Standards Optimal Resources for Cancer Care. Cancer registries are looking for guidance on implementing these new standards. We are seeking articles on this topic. Please share your expertise with your cancer registrar colleagues.

Regards,
Danette A. Clark, BS, RMA, AAS, CTR
Editor-in-Chief
JRMeditor@NCRA-USA.org
Indicators of Socioeconomic Status for Individuals, Census Tracts, and Counties: How Well Do Measures Align for Demographic Subgroups?

Jennifer L. Moss, PhD a,b; Norman J. Johnson c; Mandi Yu a; Sean F. Altekruse d; Kathleen A. Cronin a

Abstract: Objectives: Researchers often approximate individual-level socioeconomic status (SES) from census tract and county data. However, area-level variables do not serve as accurate proxies for individual-level SES, particularly among some demographic subgroups. The present study aimed to analyze the potential bias introduced by this practice. Methods: Data included (1) individual-level SES from the Mortality Disparities in American Communities study (n = 3,471,000 collected in 2008), and (2) census tract- and county-level SES from the 2006–2010 American Community Survey. Analyses included correlations among SES indicators (eg, median household income, having a high school degree, unemployment) across individual versus census tract and county levels, stratified by sex, age, race/ethnicity, and urbanicity. Finally, generalized estimating equations evaluated demographic differences in whether area-level SES matched or underestimated individual-level SES. Results: Low correlations were observed between individual- and area-level SES (census tract: Spearman’s ρ range = 0.048 for unemployment to 0.232 for median household income; county: ρ range = 0.028 for unemployment to 0.157 for median household income; all P < .0001). SES indicators were more likely to match for males, older participants, and urban groups. Area-level SES indicators were more likely to underestimate individual-level SES for older participants and rural groups, indicating that individuals who are part of these groups may live in systematically lower-SES communities than their own SES might connote. Conclusions: In this population-based study of 3.5 million participants, area-level indicators were poor proxies for individual-level SES, particularly for participants living in rural areas.

Key words: health disparities, measurement error, social epidemiology, sociodemographics, socioeconomic status

Introduction

Higher socioeconomic status (SES) is generally associated with improved health.1-3 SES reflects an individual’s relative position within a social structure,4 often captured with indicators such as income, education, and employment.5,6 However, collecting these data can be difficult.7 To overcome this challenge, researchers may use aggregated, area-level (eg, census tract, county) indicators as proxies for individual-level SES.8,9 Similar to research on individual-level indicators, area-level SES demonstrates positive associations with health for most outcomes.10-12 A general rule of thumb for conducting research with area-level proxies of individual-level SES is to use the smallest ecological unit available, since those data will be least vulnerable to aggregation bias.13,14 As Schuurman and colleagues14 wrote, “Clearly aggregation has a homogenizing effect”; that is, larger ecological units of aggregation mask important variations across space. While some studies have demonstrated that area-level proxies may not serve as valid representations of individual-level SES,15-19 this research practice will likely continue due to data availability and privacy. Further, some researchers use area-level SES indicators not as true proxies for individual-level SES, but as indicators of risk based on socioeconomic characteristics.

Given the interrelationships among SES and other social determinants of health,20 it is important to evaluate whether area-level proxies operate similarly for different demographic groups. For example, racial/ethnic minorities face unique barriers to residential mobility, such as segregation. As a result, even as their income increases, they...
may not move out of low-income neighborhoods. Thus, area-level income may not approximate individual-level income as well for minorities as for non-Hispanic whites. These differences could result in biased estimates of the association between SES and health across demographic subgroups.

In this study, we used a population-based sample (N ≈ 4 million) to evaluate the validity of area-level proxies of individual-level SES across demographic subgroups. To do this, we quantified the concordance for individual- versus census tract- and county-level indicators of SES for the overall population and then by subgroups defined by sex, age, race/ethnicity, and urbanicity. The findings from this analysis can influence researchers’ choices about analyzing area-level SES indicators, especially for certain demographic subgroups.

Methods

Data Source

Data came from the Mortality Disparities in American Communities (MDAC) study, a project by the US Census Bureau, Centers for Disease Control and Prevention, and National Institutes of Health to support research on mortality disparities by social and economic characteristics (https://www.census.gov/mdac). MDAC links data from the 2008 American Community Survey (ACS) to mortality data from the National Death Index and other sources (2008–2015). Currently, the 2008 ACS data is the only cohort of MDAC are available elsewhere. For this paper, analyses were restricted to persons of age ≥18 (n ≈ 3,471,000) linked to a census tract (n ≈ 2,830,000) or county (n ≈ 2,854,000). We linked individual-level MDAC data to SES characteristics for census tracts and counties from the 5-year ACS estimates (2006–2010; ie, centered around 2008). When 5 years of ACS data are combined, they allow robust estimates at block, census tract, and county levels.

Measures

We gathered demographic subgroup indicators and widely-used indicators of SES (ie, income/poverty, education, and employment) across individuals, census tracts, and counties.

Demographic Subgroup Indicators

We gathered individual-level demographic data: sex (male or female), age (18–24, 25–44, 45–64, or ≥65 years), race/ethnicity (non-Hispanic [NH] white, NH black, NH Asian/Pacific Islander [API], NH American Indian/Alaska Native [AIAN], Hispanic, or other), and urbanicity (urban or rural). Participants self-reported sex, age, and race/ethnicity in ACS. We classified county urbanicity using rural–urban continuum codes from the US Department of Agriculture; specifically, participants living in counties with codes 1–3 were classified as urban/metropolitan (“urban”), and participants living in other counties were classified as rural/nonmetropolitan (“rural”).

SES Indicators

Income/poverty. For individuals, we developed indicators for whether their household had an income that was at/above versus below the median across all US households and ≤100% of the federal poverty level (ie, in poverty) versus was not ≤100% of the federal poverty level (ie, not in poverty), based on household income and size. We categorized census tracts and counties as at/above versus below the median household income of US census tracts and counties (respectively) and at/above versus below the median percentage of households in poverty.

Education. For individuals, we developed indicators for whether they had a high school degree and they were ≥25 years and had a 4-year college degree (participants age <25 years were assigned missing values). We categorized census tracts and counties as at/above versus below the median percentage of age-appropriate individuals in that area who had a high school or college degree.

Unemployment. For individuals, we developed an indicator for whether they were in the labor force but unemployed (participants who were not in the labor force were assigned missing values). We categorized census tracts and counties as at/above versus below the median percentage of unemployed individuals in that area (ie, individuals aged ≥16 years who were not “at work”/“with a job but not at work” in the week before completing ACS; actively looking for work in last 4 weeks; and available to accept a job).

Next, we constructed dichotomous variables to reflect (1) whether participants’ individual SES matched or were equal to their census tract and county SES indicators (eg, if, for each variable, they were classified into a category indicating higher SES according to individual-level household income and also classified into a category indicating higher SES according to census tract–level household income or according to county-level household income) and (2) whether participants’ individual SES was underestimated by their area-level SES indicators (eg, if they were classified into a category indicating higher SES according to individual-level household income but classified into a category indicating lower SES according to census tract- or county-level household income). SES indicators for poverty and unemployment were reverse-coded for the underestimation variable; that is, individuals were coded as having their SES underestimated if they were not in poverty but were living in a high-poverty census tract or county, and if they were not unemployed but were living in a high-unemployment census tract or county. Participants whose individual-level indicators were underestimated by area-level indicators had higher SES than most of their neighbors; their higher SES may obscure results of analyses examining the SES-health gradient.

Statistical Analysis

We calculated the proportion and standard deviation (SD) for each dichotomous SES variable by dividing the number of individuals/census tracts/counties with the
relevant characteristic by the total units (Table 1); by design, proportions for variables that reference a median should be approximately 0.50. Next, we estimated the Spearman’s correlations among SES indicators at the individual, census tract, and county levels for the overall sample and stratified by demographic characteristics. Spearman’s correlation statistics reflect the direction and strength of the relationship between 2 dichotomous variables.

Finally, we used generalized estimating equations to examine concordance among SES indicators across subgroups. We examined the association between demographic characteristics and the individual- to census tract- level matched variable for each SES indicator, controlling for county population (in 1,000s) and adjusting for clustering of individuals within census tracts. We repeated this analysis for the individual- to census tract-level underestimated variables, and then for the individual- to county-level matched and underestimated variables (adjusting for clustering of individuals within counties). This approach allowed us to examine whether certain subgroups were more likely to live in areas where (1) their individual-level SES matched the area-level indicators and (2) their individual-level SES was underestimated by the area-level indicators.

Analyses were conducted in SASv9.3. Individual-level observations were weighted to account for nonequal probability of selection into ACS/MDAC. Below, we present unweighted frequencies and weighted percentages. The Office of Management and Budget approved data collection for ACS. The procedures for the current analysis were approved by the Center for Economic Studies at US Census Bureau. Output was reviewed by US Census Bureau staff to maintain confidentiality of patient data.

### Results

Correlations between individual- and area-level SES indicators were weak but statistically significant (all $P < .0001$) (Table 2). For example, for the overall sample, individuals living in households with income above the US median were more likely to live in census tracts (Spearman’s $r = 0.232$) and counties ($r = 0.157$) with incomes that were above the median. Almost all correlations were stronger for the comparison of individual- versus census tract-level SES than for individual- versus county-level SES, indicating stronger relationships for the smaller ecological units.

Across demographic subgroups, correlations for individual- to area-level SES indicators were generally stronger for participants who were male, aged 18–44 years, and urban compared to their counterparts; that is, members of these groups were more likely to live in areas with SES similar to their own. For example, individual- to census tract-level correlations for household income were stronger for males ($r = 0.235$) than females ($r = 0.231$); for 18- to 24-year-old individuals ($r = 0.256$) than other age groups (eg, $r = 0.172$ for ages ≥65 years); and for urban ($r = 0.258$) than rural participants ($r = 0.164$). Results were mixed for correlations across race/ethnicity groups.

#### Area-Level SES Indicators Matching Individual-Level SES Indicators

Subgroup differences for living in areas whose SES matched their individual-level SES are summarized in Table 3. More than half the sample had area-level SES indicators that matched their individual-level SES.

The likelihood of being matched for individual- versus census tract-level SES (Table 3, top panel) varied across demographic characteristics. For example, living in a census tract with a median household income category that matched their individual-level category was less common for females (odds ratio [OR], 0.997), participants aged ≥25 years (OR range, 0.957–0.993), Hispanics (OR, 0.987), and rural groups (OR, 0.971) than for their counterparts (all $P < .01$). In contrast, matching for household income was more common for NH minorities (OR range, 1.028 for NH minorities).

### Table 1. Descriptive Statistics for Individual-, Census Tract-, and County-Level Indicators of Socioeconomic Status, Mortality Disparities in American Communities, 2008

<table>
<thead>
<tr>
<th></th>
<th>%</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Individuals (n=3,470,968)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income/Poverty</td>
<td></td>
<td></td>
</tr>
<tr>
<td>At or above median HH income</td>
<td>47.4</td>
<td>0.50</td>
</tr>
<tr>
<td>At or below 100% FPL</td>
<td>9.8</td>
<td>0.30</td>
</tr>
<tr>
<td>Education</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Had a high school degree</td>
<td>85.9</td>
<td>0.35</td>
</tr>
<tr>
<td>Had a college degree</td>
<td>26.2</td>
<td>0.44</td>
</tr>
<tr>
<td>Employment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Was unemployed</td>
<td>3.5</td>
<td>0.18</td>
</tr>
<tr>
<td><strong>Census tracts (k=73,057, n=2,829,760)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income/Poverty</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥ median HH income</td>
<td>50.7</td>
<td>0.50</td>
</tr>
<tr>
<td>≥ median % at or below 100% FPL</td>
<td>48.6</td>
<td>0.50</td>
</tr>
<tr>
<td>Education</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥ median % with high school degree</td>
<td>50.1</td>
<td>0.50</td>
</tr>
<tr>
<td>≥ median % with college degree</td>
<td>47.4</td>
<td>0.50</td>
</tr>
<tr>
<td>Employment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥ median % of unemployed</td>
<td>48.9</td>
<td>0.50</td>
</tr>
<tr>
<td><strong>Counties (k=3,242, n=2,854,360)</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Income/Poverty</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥ median HH income</td>
<td>49.0</td>
<td>0.50</td>
</tr>
<tr>
<td>≥ median % at or below 100% FPL</td>
<td>46.0</td>
<td>0.50</td>
</tr>
<tr>
<td>Education</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥ median % with high school degree</td>
<td>51.7</td>
<td>0.50</td>
</tr>
<tr>
<td>≥ median % with college degree</td>
<td>47.8</td>
<td>0.50</td>
</tr>
<tr>
<td>Employment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>≥ median % of unemployed</td>
<td>47.5</td>
<td>0.50</td>
</tr>
</tbody>
</table>

FPL, federal poverty level; HH, household.
Table 2. Spearman’s Correlation Coefficients for the Relationships between Individual- and Census Tract-Level (Top Panel) and Individual- and County-Level (Bottom Panel) Socioeconomic Status Indicators

<table>
<thead>
<tr>
<th>Correlation with respective census tract characteristic</th>
<th>Median HH income</th>
<th>Poverty</th>
<th>HS degree</th>
<th>College degree</th>
<th>Unemployment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>0.232</td>
<td>0.122</td>
<td>0.158</td>
<td>0.225</td>
<td>0.048</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>0.235</td>
<td>0.110</td>
<td>0.163</td>
<td>0.241</td>
<td>0.062</td>
</tr>
<tr>
<td>Female</td>
<td>0.231</td>
<td>0.132</td>
<td>0.153</td>
<td>0.209</td>
<td>0.056</td>
</tr>
<tr>
<td>Age (y)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18–24</td>
<td>0.256</td>
<td>0.145</td>
<td>0.112</td>
<td>0.124</td>
<td>0.080</td>
</tr>
<tr>
<td>25–44</td>
<td>0.248</td>
<td>0.133</td>
<td>0.170</td>
<td>0.268</td>
<td>0.070</td>
</tr>
<tr>
<td>45–64</td>
<td>0.235</td>
<td>0.113</td>
<td>0.164</td>
<td>0.235</td>
<td>0.049</td>
</tr>
<tr>
<td>≥65</td>
<td>0.172</td>
<td>0.086</td>
<td>0.162</td>
<td>0.180</td>
<td>0.013</td>
</tr>
<tr>
<td>Race/ethnicity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NH White</td>
<td>0.216</td>
<td>0.096</td>
<td>0.113</td>
<td>0.216</td>
<td>0.044</td>
</tr>
<tr>
<td>NH Black</td>
<td>0.221</td>
<td>0.112</td>
<td>0.111</td>
<td>0.162</td>
<td>0.066</td>
</tr>
<tr>
<td>NH API</td>
<td>0.251</td>
<td>0.150</td>
<td>0.163</td>
<td>0.208</td>
<td>0.053</td>
</tr>
<tr>
<td>NH AIAN</td>
<td>0.194</td>
<td>0.114</td>
<td>0.108</td>
<td>0.158</td>
<td>0.072</td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.200</td>
<td>0.115</td>
<td>0.164</td>
<td>0.177</td>
<td>0.058</td>
</tr>
<tr>
<td>Other</td>
<td>0.119</td>
<td>0.131</td>
<td>0.142</td>
<td>0.184</td>
<td>0.050</td>
</tr>
<tr>
<td>Urbanicity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urban</td>
<td>0.258</td>
<td>0.141</td>
<td>0.179</td>
<td>0.230</td>
<td>0.066</td>
</tr>
<tr>
<td>Rural</td>
<td>0.164</td>
<td>0.078</td>
<td>0.108</td>
<td>0.147</td>
<td>0.040</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Correlation with respective county characteristic</th>
<th>Median HH income</th>
<th>Poverty</th>
<th>HS degree</th>
<th>College degree</th>
<th>Unemployment</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall</td>
<td>0.157</td>
<td>0.067</td>
<td>0.096</td>
<td>0.150</td>
<td>0.028</td>
</tr>
<tr>
<td>Sex</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>0.160</td>
<td>0.059</td>
<td>0.097</td>
<td>0.159</td>
<td>0.037</td>
</tr>
<tr>
<td>Female</td>
<td>0.155</td>
<td>0.074</td>
<td>0.095</td>
<td>0.142</td>
<td>0.034</td>
</tr>
<tr>
<td>Age (y)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18–24</td>
<td>0.153</td>
<td>0.084</td>
<td>0.057</td>
<td>0.090</td>
<td>0.051</td>
</tr>
<tr>
<td>25–44</td>
<td>0.160</td>
<td>0.071</td>
<td>0.106</td>
<td>0.170</td>
<td>0.043</td>
</tr>
<tr>
<td>45–64</td>
<td>0.159</td>
<td>0.063</td>
<td>0.107</td>
<td>0.159</td>
<td>0.032</td>
</tr>
<tr>
<td>≥65</td>
<td>0.133</td>
<td>0.049</td>
<td>0.091</td>
<td>0.124</td>
<td>0.008</td>
</tr>
<tr>
<td>Race/ethnicity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NH White</td>
<td>0.163</td>
<td>0.048</td>
<td>0.061</td>
<td>0.174</td>
<td>0.027</td>
</tr>
<tr>
<td>NH black</td>
<td>0.161</td>
<td>0.071</td>
<td>0.051</td>
<td>0.092</td>
<td>0.030</td>
</tr>
<tr>
<td>NH API</td>
<td>0.133</td>
<td>0.059</td>
<td>0.044</td>
<td>0.095</td>
<td>0.019</td>
</tr>
<tr>
<td>NH AIAN</td>
<td>0.136</td>
<td>0.071</td>
<td>0.058</td>
<td>0.105</td>
<td>0.059</td>
</tr>
<tr>
<td>Hispanic</td>
<td>0.119</td>
<td>0.045</td>
<td>0.077</td>
<td>0.052</td>
<td>0.029</td>
</tr>
<tr>
<td>Other</td>
<td>0.131</td>
<td>0.073</td>
<td>0.059</td>
<td>0.073</td>
<td>0.025</td>
</tr>
<tr>
<td>Urbanicity</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urban</td>
<td>0.161</td>
<td>0.070</td>
<td>0.097</td>
<td>0.127</td>
<td>0.035</td>
</tr>
<tr>
<td>Rural</td>
<td>0.138</td>
<td>0.065</td>
<td>0.093</td>
<td>0.123</td>
<td>0.035</td>
</tr>
</tbody>
</table>

AIAN, American Indian/Alaska Native; API, Asian/Pacific Islander; HH, household; NH, non-Hispanic.
The likelihood of being matched for individual- versus county-level SES (Table 3, bottom panel) also demonstrated differences by demographic characteristics, but these associations tended to be smaller in magnitude than for individual- versus census tract-level comparisons. For example, living in a county with a median household income category that matched their individual-level category was more common for some NH minorities (NH API: OR, 1.041; NH AIAN: OR, 1.035) and less common for Hispanics (OR, 0.952) than for NH whites. For the remaining indicators, matching values for individual and county levels tended to be more common for males, older participants, NH whites, and urban groups.

**Area-Level SES Indicators Underestimating Individual-Level SES Indicators**

Subgroup differences for living in areas whose SES underestimated their individual-level SES are summarized in Table 4. Underestimation ranged from 9% (college degree) to unemployment (47%) for individuals compared to census.
tracts and from 10% (college degree) to 46% (unemployment) for individuals compared to counties.

The likelihood of individual-level SES being underestimated by census tract-level SES (Table 4, top panel) varied across demographic characteristics. For example, living in a census tract with a median household income category that underestimated their individual-level category was less common for females (OR, 0.985) and NH minorities (OR range, 0.964 for NH API to 0.987 for NH black) (all $P < .0001$). In contrast, underestimation of household income was more common for participants aged 25–64 years (eg, 25–44 years: OR, 1.036) and for rural groups (OR, 1.059) (all $P < .0001$). Across SES indicators, underestimation tended to be more common for older participants and rural groups.

The likelihood of individual-level SES being underestimated by county-level SES (Table 4, bottom panel) also demonstrated differences by demographic characteristics. For example, living in a county with a median household income category that underestimated their individual-level category was less common for females (OR, 0.985) and NH minorities (OR range, 0.921 for NH API to 0.944 for NH AIAN) (all $P < .0001$). In contrast, underestimation of household income was more common for participants aged 25–64 years (25–44 years: OR, 1.044; 45–64 years: OR, 1.081) and for rural groups (OR, 1.052) (all $P < .0001$).
Correlations across socioecological levels were consistently low. This suggests that individual- versus area-level SES indicators are poor proxies for individual-level SES, particularly for rural individuals.

Studies on SES and health are challenged by limited data on individual-level SES. The findings from the present study indicate that using area-level proxies is not ideal due to the low correlations among individual- and area-level SES indicators. For instances in which individual-level data are not available, the most optimal area-level proxies may be household income and college degree, since these variables had the greatest individual- to area-level correlations. However, these variables also demonstrated some of the greatest variation in correlations across demographic subgroups. For example, the race/ethnicity-specific $r$’s for individual- versus census tract-level household income ranged from 0.119 to 0.251 (difference, 0.132), compared to the $r$’s for unemployment, which ranged from 0.044 to 0.72 (difference, 0.029). Analyses of demographic differences in the SES-health gradient may need to exercise more caution in selecting and interpreting area-level proxies for individual-level SES. Another suggestion for best practice when using area-level SES proxies, as suggested elsewhere, is to use data at the smallest ecological unit possible. Generally, the correlations with individual-level SES were stronger for census tracts than for counties. However, area-level SES indicators are not always used as proxies for individual-level SES; they may alternatively serve to indicate high- or low-risk groups, or to provide unique, contextual information about the environments in which people live.

The concordance between individual- and area-level SES indicators was particularly sensitive to urbanicity. Correlations across socioecological levels were consistently lower for people living in rural compared to urban areas; for...
example, the individual- to census tract-level correlations in Table 2 are at least 50% greater in magnitude for urban versus rural participants. In addition, participants in rural areas were less likely to live in census tracts or counties that matched their individual-level SES (except for the high school degree indicator) (Table 3) and more likely to live in areas that underestimated their individual-level SES (Table 4). Taken together, this pattern of findings indicates that area-level SES may systematically underestimate the socioeconomic resources held by rural residents; for example, the odds of a person’s SES (measured by household income, poverty, or having a high school degree) being underestimated by area-level indicators were 5%–10% higher for people living in rural versus urban areas. As a result, analyses of SES and health in rural areas may be vulnerable to bias if they leverage area-level indicators as proxies for individual-level SES. The differences observed here could be related to historical patterns in rural/urban migration, but additional research is needed on how individual- and area-level SES indicators differ for rural versus urban areas.

Less consistent patterns developed for SES indicators across the other demographic subgroups. For sex, females generally had lower correlations than males, except for poverty, which could be related to the “feminization” of poverty. Further, females were generally less likely to live in areas that matched their individual-level SES. For age, correlations tended to be strongest for the younger age groups, but matching was more common for the middle age groups. These results could reflect developmental processes in establishing SES (which changes substantially over the life course), particularly among the 18- to 24-year-old group; secular trends in educational attainment for different age cohorts, and the challenge of measuring SES among older individuals who may be retired. For race/ethnicity, results were mixed. Individual- to census tract-level correlations for household income, for example, were strongest for NH API and weakest for “other” racial/ethnic groups, but the individual- to county-level correlations were strongest for NH whites and weakest for Hispanics. Matching for household income was generally more common among NH minorities and less common for Hispanics compared to NH whites; however, underestimation was most common for NH whites. Thus, NH whites are more likely than other groups to live in areas that have SES levels discordant with their individual-level SES, which could reflect racialized economic segregation that results in very homogenous racial/ethnic enclaves but more diverse “white” neighborhoods. Additional research is warranted on these processes, disparities, and implications for public health.

This study has several strengths. The analysis leveraged data from MDAC, an innovative, nationally-representative, high-quality dataset with 3.5 million participants. In particular, this study advances the work of previous analyses with more limited sample size and/or geographic scope. In addition, >80% of participants were successfully linked to their census tract/county of residence. In terms of limitations, SES data were self-reported, which could bias results. In addition, for simplicity, we only examined dichotomous measures of SES, although many studies examine other ways of categorizing SES, eg, quartiles; future studies could expand on the current results to examine concordance across more fine-grained categorizations, which could uncover larger differences in concordance. Finally, although individual-level SES data were obtained for 2008, area-level SES estimates were drawn from 5 years of data (2006–2010) due to concerns about data stability; the differences in these time periods could introduce noise into the analysis.

Public Health Implications

Area-level SES indicators are poor proxies for individual-level SES. However, when research studies must use area-level proxies, indicators of household income and college education may be most applicable to national studies. Alternative indicators, especially those that are compositional or theory-based, may be more appropriate for studies focused on demographic differences in SES and health. In addition, census tract-level indicators performed more consistently than did county-level indicators, supporting other studies suggesting that area-level proxies for individual-level SES should come from the smallest ecological unit available. Area-level indicators matched individual-level SES (ie, performed best as proxies) for males, older participants, and urban groups. Concordance between individual- and area-level indicators of SES should continue to be monitored (eg, with the 2020 census data currently being collected). Researchers conducting epidemiologic and intervention studies on the SES-health gradient should consider the limitations of these variables when designing their analysis.

References


Improved Geocoding of Cancer Registry Addresses in Urban and Rural Oklahoma

Naci Dilekli a,b,*, Amanda Janitz c, Janis Campbell c

Abstract: Background: Between 1997 and 2013 (the included study years), approximately 23% of addresses in the Oklahoma Central Cancer Registry (OCCR) were not geocoded to the address level. Addresses in rural counties were geocoded with poorer quality, preventing the instructive geographic research that informs policymaking. Methods: To improve the accuracy of the geocodes, we first utilized the United States Postal Service’s LACSLink database to correct addresses; specifically, to convert old rural route-based addresses to modernized Enhanced 911 (E911) addresses. We created custom geocoders using regional E911 reference data sets and used existing national scope geocoders of NAVTEQ and the North American Association of Central Cancer Registries. We attempted to geocode 5,102 addresses, which are either regular street addresses or rural route addresses. In the process, we evaluated and tabulated performances of the address correction. Accordingly, we first tabulated how well each geocoder could geocode original and LACSLink corrected addresses. We then documented the overall performances of geocoders based on pairwise comparisons. Results: We were able to geocode 1,945 addresses out of this data set using 5 distinct geocoders. We observed that the LACSLink correction and E911 data were useful in the specific purpose of geocoding rural addresses, as found in the literature. Conclusions: We conclude that both LACSLink correction and E911 data were useful for improving geocoding of cancer records, many of which were in rural areas. Future directions include further validation of the geocoding and plans to conduct spatial exploratory data analysis to generate hypotheses related to the distribution of cancer in Oklahoma.

Key words: cancer, geocoding, geographic information systems (GIS), Oklahoma, policy, rural

Introduction

Oklahoma had the sixth highest age-adjusted cancer mortality rate in the United States from 2014–2016 (Oklahoma, 196.5 per 100,000; United States, 158.6 per 100,000).1 In order to conduct high-quality studies of environmental exposures and health outcomes, such as cancer, it is essential to have valid geocoding information.2 Spatial epidemiology is the study of geographic variation of diseases or disease risk factors. It is important to understand the underlying spatial factors such as environmental exposures, including tobacco and multiple cancers, asbestos exposure and mesothelioma, aflatoxin and liver cancer, benzene and acute myeloid leukemia, and air pollution and lung cancer.3-6 Other factors such as demographics, socioeconomic status, behavioral factors, and genetic factors are linked to cancer as well,7,8 and need to be considered when evaluating the relationship between the environment and cancer. Other authors have discussed the importance of accurate geocoding in relation to mapping multiple health outcomes,9-16 including cancer.17-20 Key to implementing these studies are accurate geocoding.

When using data collected for surveillance purposes, researchers must often rely on self-reported address information, which frequently includes incomplete addresses, rural routes, highway contract (HC) and post office (PO) boxes. Using rural routes, HC boxes, and PO boxes for geocoding of a record’s residence can result in artificial clustering of residences, since these addresses are often geocoded to the ZIP code or other centroid, increasing the risk of exposure misclassification.21-24 Excluding records due to poor geocoding may also result in selection bias, since these residences are more likely to be in rural areas.25 A simulation study of environmental exposures and health outcomes demonstrated that results were biased towards the null value due to poor quality geocoding.26 Moreover, certain populations such as rural, majority, and indigenous populations can be significantly underrepresented when excluding records geocoded to ZIP centroids.9,13 Oklahoma has a significant rural population (34% rural) and has a large population of American Indian/Alaska Natives compared to the United States overall (9.2% vs 1.3%, respectively),27 which highlights the need to improve geocoding quality in Oklahoma.

The geocoding rate and the positional accuracy of successful matches depend on the underlying data sets and the context.28-31 For example, all administrative records,
including health records, in Finland are geocoded with under 1-meter accuracy. However, in Wales, 68% of addresses in urban areas compared and 48% of addresses in rural areas are successfully geocoded. In Florida, 11.5% of all births and 25% of rural births were not geocoded as of 2009. In Oklahoma, 23% of addresses between 1997 and 2013 (the included study years) were not geocoded to the address level, with generally poorer quality in rural counties than urban/suburban counties. In rural counties, Oklahoma addresses were more frequently geocoded to the ZIP code centroid level (Oklahoma Central Cancer Registry, unpublished data).

One solution to improving geocoding of rural routes is using the Enhanced 911 (E911) program. The E911 programs are county level readdressing projects, which modernize the addressing system so that calls for emergency service (911 can be mapped with sufficient accuracy. This often requires renumbering and/or renaming pre-E911 rural route box–based addresses and creating post-E911 street segments, which allows a post-E911 address to be geocoded. While the implementation of this process varies by county, it has been shown to improve geocoding quality in a previous study.

We aimed to evaluate the proportion of addresses previously geocoded above the street level (ie, ZIP centroids) that were successfully geocoded to the street level using an address correction program and multiple geocoders in a systematic manner. We evaluated the usefulness of different geocoders based on the conventional data sets (such as Census Bureau’s TIGER files and NAVTEQ), and incorporating the E911 data and Locatable Address Conversion System Link database (LACS) correction for pre-E911 addresses.

Methods

Data Sources and Study Population

Data on cancer records were obtained from the Oklahoma Central Cancer Registry (OCCR), a Centers for Disease Control and Prevention (CDC) funded National Program of Cancer Registries (NPCR) registry, which maintains all cancers diagnosed or treated among Oklahoma residents since January 1, 1997. This statewide population-based registry enables both private and public health agencies to study cancer trends and to develop and evaluate cancer prevention and control programs. OCCR follows standards developed by the North American Association of Central Cancer Registries (NAACCR). For this study, we included data for all cancer cases diagnosed in Oklahoma from January 1, 1997 through December 31, 2013, which were the years of available data at the time the study was conducted. We included data from the following councils of governments: (1) Association of Central Oklahoma Governments (ACOG) includes Canadian, Cleveland, Logan, and Oklahoma counties; and (2) South Western Oklahoma Development Authority (SWODA) includes Beckham, Custer, Greer, Harmon, Jackson, Kiowa, Roger Mills, and Washita counties. These locations were chosen to represent both urban/suburban (ACOG) and rural (SWODA) regions of Oklahoma. Data include 10,666 records from ACOG and SWODA counties that were not geocoded to the street level by the OCCR. Most of these entries (52%) had PO boxes or unknown addresses, which are not possible to geocode.

Study Design

Address Standardization and Geocoding. To standardize and ensure all addresses were compatible with the E911 system, we used AccuMail Frameworks software (specifically, the LACSLink database) for (1) standardization, which corrects, parses and standardizes the address information using the latest United States Postal Service (USPS) data; (2) validation, which utilizes delivery point validation and LACSLink databases from USPS to transform addresses; and (3) enhancement, which enriches addresses by adding new information such as carrier route information, line of travel, congressional districts, and county codes. In particular, the LACSLink database was developed by the USPS to correct addresses, including conversion of rural routes to street addresses and correcting street addresses that have been renumbered or renamed.

In order to complete and evaluate the geocoding, we followed the methodology outlined in Figure 1. It involved a series of steps that first started with the original address data provided by OCCR. Those addresses were then corrected using the LACSLink database from AccuMail Frameworks software; and both original and processed address sets were retained to be geocoded in the next stages.

We used 4 geocoders (NAACCR, NAVTEQ, ACOG E911, and SWODA E911) with distinct reference data sets to geocode the address information to street addresses. Specifically, we were motivated to use geocoders with E911
records that were expected to perform better with rural addresses and were a large proportion of our address data set. To account for address formatting changes in geocoders, we used both the original and corrected addresses, and we incorporated a range of reference data sets. The SWODA E911 reference data set did not have ZIP codes; therefore, we added ZIP code information using 2 separate ZIP code geographies from the US Census Bureau and ESRI (Environmental Systems Research Institute), resulting in 2 separate geocoders for SWODA E911. Geocoding was conducted both for the original address file and for the LACSLink corrected file. Due to restricted access to the NAACCR geocoder, OCCR personnel conducted the NAACCR geocoding using only the corrected addresses; thus, we provide results only for the corrected address data set for NAACCR.

Address Correction Evaluation. As the key component of the geocoding methodology, we evaluated and tabulated the usefulness of the LACSLink correction when 4 geocoders (ACOG E911, SWODA E911, NAVTEQ, and NAACCR) were used individually. The idea was to determine if the address correction could be more useful for custom built geocoders, ACOG E911 and SWODA E911. To do this, we tabulated how well each geocoder could geocode original and LACSLink corrected addresses compared to each other, and examined the original and corrected versions of addresses that were geocoded in several aspects according to the set theory:

- \( A_{original\ total} \): Total number of geocodes completed by each geocoder using the original addresses
- \( A_{original\ only} \): Total number of geocodes completed by each geocoder using the original addresses whose corresponding corrected addresses were not geocoded by the same geocoder
- \( A_{corrected\ total} \): Total number of geocodes completed by each geocoder using the corrected addresses
- \( A_{corrected\ only} \): Total number of geocodes completed by each geocoder using the corrected addresses whose corresponding original addresses were not geocoded by the same geocoder
- \( A_{union} \): Aggregate number of geocodes completed by each geocoder using both original and corrected versions of addresses
- \( A_{intersect} \): Aggregate number of geocodes completed by each geocoder using both original and corrected versions of addresses

<table>
<thead>
<tr>
<th>Address Type</th>
<th>Result</th>
<th>SWODA</th>
<th>ESRI ZIP</th>
<th>SWODA ZIP</th>
<th>ACOG</th>
<th>NAVTEQ</th>
<th>NAACCR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original – OSDH</td>
<td>Total</td>
<td>93</td>
<td>116</td>
<td>515</td>
<td>556</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Only</td>
<td>S</td>
<td>S</td>
<td>37</td>
<td>S</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>LACSLink Corrected</td>
<td>Total</td>
<td>441</td>
<td>527</td>
<td>655</td>
<td>710</td>
<td>597</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Only</td>
<td>349</td>
<td>412</td>
<td>177</td>
<td>163</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Combined Set</td>
<td>Union</td>
<td>442</td>
<td>528</td>
<td>692</td>
<td>719</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Intersect</td>
<td>92</td>
<td>115</td>
<td>478</td>
<td>547</td>
<td>NA</td>
<td></td>
</tr>
<tr>
<td>Geocode Rate (%)</td>
<td></td>
<td>9%</td>
<td>10%</td>
<td>14%</td>
<td>14%</td>
<td>12%</td>
<td></td>
</tr>
</tbody>
</table>

Letter S indicates that the number has been suppressed since it was smaller than 10, due to the Oklahoma Central Cancer Registry suppression guidelines to protect patient confidentiality.

ACOG, Association of Central Oklahoma Governments; E911, Enhanced 911; ESRI, Environmental Systems Research Institute; NA, not available; NAACCR, North American Association of Central Cancer Registries; OSDH, Oklahoma State Department of Health; SWODA, South Western Oklahoma Development Authority.

All these tabulations were conducted using each of the 4 geocoders (Table 1). In simpler terms, we made a series of comparisons to see if the address correction was helpful or not. For example, we aimed to determine how many original (uncorrected) address records a specific geocoder (ie, NAVTEQ) can geocode successfully that cannot be geocoded using the same geocoder (ie, NAVTEQ) when those addresses are corrected (\( A_{original\ only} \)). If the same geocoder can geocode an original address at a higher quality than the corrected address, that indicates when the correction process was counterproductive, meaning that the corrected address range is not in the reference data set of the geocoder, which uses an older address range. Similarly, we measured the usefulness of the correction by looking at number of addresses that could be geocoded only when they were corrected (\( A_{corrected\ only} \)).

Pairwise Comparison of Geocoders. As shown in Table 2, we evaluated and tabulated performances using pairwise comparisons of all geocoding reference data sets based on matched confidence and precision (eg, street, city, ZIP code, county level), similar to previous literature. We only kept results that were geocoded at street level with at least 90% confidence. We measured the performance of address correction, and then we assessed the relative performance of each geocoder compared to all the rest of geocoders. For example, when we compared the ACOG E911 and NAVTEQ geocoders, we calculated how many records were geocoded only by the ACOG E911 geocoder, only by the NAVTEQ
geocoder, and by both. Large number of records that could only be geocoded by 1 geocoder indicate a relatively higher performance of this geocoder compared to the rest of geocoders.

We documented the performances of geocoders based on pairwise comparisons by examining the number of addresses that were geocoded in several aspects according to the set theory. The pairwise comparison was conducted in the following manner:

- $G_i$ total: Total number of geocodes completed by geocoder $G_i$
- $G_i$ only: Total number of geocodes completed by geocoder $G_i$, and not by geocoder $G_j$
- $G_i$ total: Total number of geocodes completed by geocoder $G_j$
- $G_i$ only: Total number of geocodes completed by geocoder $G_j$, and not by geocoder $G_i$
- $G_i$ union: Aggregate number of geocodes completed by geocoders $G_i$ and $G_j$
- $G_i$ intersect: Total number of geocoding completed by only both geocoders $G_i$ and $G_j$

Where $i$ and $j$ were the indices for the geocoders of ACOG E911, SWODA E911, NAVTEQ, and NAACCR compared in each pairwise comparison. All these comparisons were evaluated for the aggregate of the original and corrected addresses by all the geocoders. Only street addresses and rural routes were included in this analysis, since PO boxes could not be further geocoded to the street level without obtaining additional information from the cancer case, which was not feasible.

**Results**

Approximately half of included records ($n = 5,102$) had either regular street addresses or rural route addresses, with the remaining ones being PO box or unknown records ($n = 5,564$) (Table 3). At least 8% of nongeocoded addresses in Beckham, Custer, Harmon, and Kiowa counties were rural route based. However, more than half of nongeocoded addresses in Canadian, Cleveland, Jackson, Logan, and Washita counties are either PO boxes or unknown addresses. We successfully geocoded 1,945 addresses using any of the 5 geocoders used in this study. This represents 38% of the addresses (street or rural route–based addresses) that could be geocoded.

The LACSLink database was used to standardize addresses according to the USPS, resulting in 357 corrected addresses (Table 4). LACSLink corrected rural routes by converting older rural route addresses (eg, RR 101) into regular street addresses (eg, 10680 Street) and corrected misspellings/nonstandardized rural routes ($n = 2,298$).
Table 1 demonstrates a comparison of the geocoding performance of original and LACSLink corrected data sets. For example, the ACOG E911 geocoder was able to geocode 515 of the records from the original address data set and 655 records from the LACSLink corrected data set. Of the original records, 37 could be geocoded using the ACOG E911 geocoder, and their corrected version/counterparts could not be geocoded. Only 478 original addresses and their corresponding corrected addresses could be geocoded by the ACOG E911 geocoder.

Table 2 demonstrates the performances of geocoders based on pairwise comparisons. We compared the number of addresses that were geocoded by each geocoder in sets of 2. This allowed us to evaluate which geocoders geocoded only a subset of addresses and evaluate both the union and intersect between geocoders. Largest sizes of unions (over 1,200 addresses) occurred between the pairs of ACOG E911–SWODA E911, SWODA E911–NAVTEQ, and ACOG E911–NAACCR. The smallest size of unions occurred within the SWODA region, using 2 variations of geocoders.
based on different ZIP code geometries. The largest sizes of intersections occurred between the pairs of ACOG E911–NAVTEQ and within the SWODA E911 geocoders. The smallest sizes of intersections occurred in the pair of ACOG E911–SWODA E911 (since they have no common geography) and SWODA–NAVTEQ. When comparing ACOG E911 and SWODA E911 geocoders (Pairwise Comparison 3), which originate from different regions of Oklahoma, each geocoder can only geocode addresses within the respective regions (note that the total and only results for each geocoder are the same). Furthermore, the union of these 2 geocoders (geocoded with either geocoder) was the sum of each independent geocoder, but none of them were geocoded by both geocoders.

Discussion

The purpose of this study was to enable more accurate geographic research in epidemiology through the reduction of geocoding bias.40-42 The quality of geocoding varies by geographic location, with some countries geocoding records from birth to death for all individuals, and others, such as the United States, only geocoding records that are data set–specific. Varying quality of geocoding can make comparability of results between localities difficult, highlighting the need for improved geocoding in areas such as Oklahoma.

We were able to geocode 1,945 addresses out of 5,102 using 4 different geocoders. This was labor-intensive work, since we had to preprocess and build the georeference data sets. We observed that the LACSSLink correction and E911 data were useful in the specific purpose of geocoding rural addresses, similar to previous literature.2,43,44 This is based on a few observations, including the 560 addresses out of 1,077 rural route and street-based addresses that were geocoded in the SWODA region through this project, but not previously geocoded. Similarly, when we inspected the pairwise comparisons in Table 2, we observed that an important number of geocodes were completed by the custom geocoders (SWODA E911 and ACOG E911) rather than the generic geocoders (NAVTEQ and NAACCR). We observed that intersections and unions between 2 geocoders can reveal insights about the strengths and characteristics of each geocoder. For example, a smaller intersection between 2 geocoders indicates presence or absence of overlap between the geocoders either in terms of the geographical scope (eg, the ACOG and SWODA E911 geocoders) or in terms of data resolution (eg, original street addresses vs E911 addresses). In contrast, a large intersection between distinct geocoders implies redundancy, meaning that the same address has been geocoded by multiple geocoders. Based on our analyses, the custom and generic geocoders are often complementary with a high number of unions, but low number of intersects. Nevertheless, we determined that NAVTEQ and NAACCR both provided extensive geographical coverage and performed similarly in the studied areas since they could individually geocode addresses in both study regions. This provides confirmation that use of NAVTEQ and NAACCR are useful geocoders when customized geocoders are not available.

Rural route–based addresses constituted a larger proportion of total addresses in the SWODA region, compared to the ACOG region. SWODA is a predominantly rural area and our in-depth geocoding resulted in an increase in the number of geocoded records in the SWODA region (as seen in Table 2). A lower percentage of geocoded addresses improved for the ACOG region, suggesting a lower proportion of rural routes in this more urbanized region of Oklahoma. Finally, the NAVTEQ geocoder had the highest geocoding performance among all geocoders based on the initial results since it has the highest number of geocodes (n = 719) as displayed in Table 1. This is expected, since both NAVTEQ and NAACCR geocoders have the same spatial extent. The SWODA and ACOG E911 geocoders have the reference data sets for their respective geographies; therefore, they would not be useful outside of these regions, but provide a guide for geocoding small regions with a high proportion or rural addresses. However, as part of the E911 address conversion, rural route–based addresses have been converted to street addresses, which should allow for improvement of geocoding these addresses over time, though is still problematic for pre-E911 addresses.45

In 2007, Zimmerman et al wrote a broad and comprehensive work on geocoding in cancer research, introducing the requirements, beneficial uses, and challenges such as positional errors, nonstandardization, and confidentiality concerns.46 Several works (including ours) have added tools to this important topic.47-50 For example, Jacquez suggested a research agenda to improve geocoding, stressing the importance of positional accuracy.48 Similarly, Goldberg suggested a framework for evaluating geocoding that can assist large organizations, such as state health departments, to integrate their geocoding. This framework involves a Web-based interactive system to research and re-geocode data and to verify the accuracy of previously geocoded records.31 By implementing this framework, organizations could save costs and provide higher quality geocodes.32

Based on the unique strengths of individual geocoders, we recommend that multiple geocoders be implemented by governmental and research institutions so that individual researchers do not have to conduct the laborious work for each research study conducted. Second, in the presence of commercial and research interest, it is possible to build software tools that would allow for address correction, geocoding, and performance evaluation in a programmatic manner using software services and API (application program interface) technologies.

Results of this study can be further strengthened using ground truthing information from global positioning system (GPS) measurements and redundant geocodes to compare accuracy of each geocoder and, hence, the underlying data set as well as the patterns of accuracy across the study geography. For example, one might expect similar estimates with higher accuracies in urban environments.53

One limitation is that we were unable to geocode PO boxes since no physical address exists within the OCCR for these records.21 This continues to be a limitation in studies using geospatial information on existing records since these records will either be excluded, resulting in
potential selection bias, or geocoded to the ZIP code centroid, resulting in potential misclassification. Because existing data, including cancer registry data, are frequently used in geospatial analyses, it is important to collect physical address data from all records in order to have accurate geocoding information. As a next step, we propose to employ geoinputation methods\textsuperscript{53} to increase accuracy of spatial location when the address information cannot be readily usable by geocoders.

In the future, cancer registries can consider 2 methods for obtaining missing: address correction or obtain the street-level address from the patient. In case of an existing address, manual investigation/correction can be used.\textsuperscript{54} To obtain the street-level address from the patients, we recommend that hospitals and other reporting facilities prioritize collection of address data and follow-up directly with patients when a PO box or rural route address is reported in the medical record. Either method is time consuming and relatively costly, but are necessary if the health services would like to maintain an accurate distribution of the patients.

**Conclusion**

Because each reference data set has strengths and limitations, we chose to use multiple data sets to take advantage of these differing strengths. Thus, we compared the performance of each of these data sets in terms of geocoding. As an end result, we developed a comprehensive geocoding database that can be applied as cancer records are added to the OCCR to more efficiently geocode Oklahoma cancer records.

We conclude that improvements in standardizing and converting addresses to E911 addresses were useful for improving geocoding of cancer records, many of which were in rural areas. We observed the highest level of improvement in the rural areas using the SWODA and ACOG E911 geocoders (Table 2), which were developed using E911 data sets from their corresponding administrations. Therefore, we strongly recommend obtaining a local E911 based data set, correcting the addresses with LACSLink if possible, and developing a custom geocoder for the rural addresses. While the resources to develop and maintain this type of geocoder may be difficult to identify, it can be argued that for individual states or the US Census Bureau there is sufficient public benefit from a regularly updated (like TIGER) and consolidated E911 based reference data set that can be used for geocoders. Future directions include further validation of the geocoding and exploratory analysis to generate hypotheses related to the distribution of cancer in Oklahoma. We plan to work with tribal partners to enhance geocoding in tribal areas for program planning and evaluation of behavioral and environmental factors related to cancer and other health outcomes.

**References**


37. USPS. LACSLink® Product Distribution Licensee Performance Requirements. 2014.


**How I Do It**

**Cutaneous Melanoma Abstracting**

*Melissa Riddle, CTR*

---

**Introduction**

Melanoma abstracting can be a love/hate relationship for the registrar. This may appear to be an easy site with little issues, straightforward, and rarely a difficult treatment timeline. Unfortunately, that isn’t how melanoma abstracting feels when you get into the dynamics and attempt to complete the case.

Let’s begin with a few fun facts about melanoma. Cutaneous melanoma arises from the melanocytes within the basal cell layer of the epidermis. Melanocytes can appear in other areas of the body like the mucous membranes of the gastrointestinal or genitourinary tracts and within the eye. The most common and likely area for a melanoma to arise is from the epidermis of the skin and the focus for this discussion will be on cutaneous melanoma.

The primary function of a melanocytes is to give color or pigmentation and help protect the deeper levels of the skin. Increased sun exposure and ultraviolet (UV) light are the major risk factors for developing cutaneous melanoma by causing damage to the DNA in skin cells, which affect the genes that control how skin cells grow and divide. One of those specific genes is the \textit{BRAF} gene. The \textit{BRAF} oncogene is not inherited and occurs during the development of the melanoma. The \textit{BRAF} gene contains the instructions for making the \textit{BRAF} protein and their main job is cell growth. This mutation jams the switch to “on,” and the cell is unable to turn it off when there’s a lack of external signals. This causes the cell to keep dividing and growing. About half of all cutaneous melanomas have a mutation in the \textit{BRAF} gene.

Now, let’s get on to the good stuff—how to abstract these cases.

**Primary Site, Histology, Solid Tumor Rules**

Most of the melanoma cases you come across are cutaneous primary sites (C44._). As many registrars should know by now, the cutaneous melanoma section was not updated for 2018 in the \textit{Solid Tumor Rules} (STR). This means registrars will continue to use the previous rules from the \textit{Multiple Primary and Histology} (MPH) manual to determine the number of primary sites and histology. It is expected that an update for cutaneous melanoma section will be available in the 2021 version of the STR.

For cases that are found to only be “metastatic melanoma” and no primary site is located, you are to code this to skin, NOS (C44.9). Example: Patient has a brain mass biopsy revealing metastatic melanoma. During workup and on history, there is no previous melanoma, nor a primary site identified. This should be coded as \textit{skin, NOS} (C44.9), not \textit{unknown primary} (C80.9), because most melanomas arise from a cutaneous site.

There are 5 main types of cutaneous melanoma. These are based on their location, shape, and the type growth (outward or downward). An acral melanoma is one that occurs on the palms of the hand, soles of the feet, or nail beds. The desmoplastic melanoma is rare and is marked by nonpigmented lesions on sun-exposed areas of the body. Lentigo maligna type usually occurs on the faces of elderly patients. Superficial spreading (flat melanoma) grows in an outward fashion to form an irregular pattern on the skin with an uneven color. The nodular melanomas are lumpy and often blue-black in color and may grow faster with a downward spread. Early or evolving melanomas (behavior is irrelevant) are not reportable for cases diagnosed on or after January 1, 2018 (SEER, Appendix C – Melanoma).

**Site Specific Data Items (SSDI)**

The SSDI’s for cutaneous melanoma are Breslow’s thickness, ulceration, mitotic rate melanoma, lactate dehydrogenase (LDH) pretreatment laboratory value, LDH upper limits of normal, and LDH pretreatment level. Each of these are important when it comes to understanding the prognosis for the patient, how the melanoma will behave, and for staging purposes.

Breslow’s thickness is the measurement of how deeply the melanoma has grown into the skin. This is the tumor thickness from the top of the tumor to the deepest tumor cells. Thicker tumors are linked with lower survival rates. This data item is used to help determine stage. Code the measurement as described on the pathology report from the greatest measured thickness from any procedure (biopsy or excision). Cases with no surgical procedure of the primary site and those with surgical procedures of the primary site after neoadjuvant therapy are coded as XX.9. Record the actual tumor depth in tenths of millimeters. When recording the measurement, include the decimal in this 4-digit field. Example: tumor described as 0.8mm thick – code as 0.8.

Ulceration is the formation of a break on the skin or on the surface of an organ. The presence of absence of ulceration must be confirmed from the pathology report (microscopic exam). Do not assume that ulceration is not present if there is no mention of it on the pathology report. You must have a statement of no ulceration present in order to code 0 (no ulceration). This data item also affects staging.

The mitotic rate is a way to describe the potential aggressiveness of a tumor. This is measured by the pathologist recording the number of cells actively dividing. This may also be referred to as the mitotic index or mitotic activity. The mitotic rate will be found on the pathology report and if more than 1 pathology for the same melanoma is documented with different counts, record the highest mitotic count. This is a prognostic factor for cutaneous melanoma.
When cells are damaged or destroyed, they release LDH into the bloodstream. The LDH pretreatment laboratory value is a predictor for treatment response, progression-free survival, and overall survival for patients with stage IV cutaneous melanoma. LDH is only considered in staging when there is distant metastasis. The LDH must be taken prior to any systemic treatment, radiation, or surgery to a metastatic site. Record the highest serum LDH results documented either before or after surgical resection of the primary tumor with or without regional lymph node dissection. This information may be available from the physician’s office. The registrar will code the actual laboratory value in unit per liter (U/L), which is how enzymes are measured. Example: LDH laboratory report shows 193U/L – code as 193.0.

The LDH upper limits of normal is based on the same laboratory results as the LDH pretreatment laboratory value data item. The normal reference range laboratory value varies by laboratory, patient age, and the units of measure. On the laboratory report, there will be a reference range given to know whether the patient’s value is within that span of values (normal, low, or high). For this data record, the highest number for the normal reference range. Example: LDH laboratory report shows 193U/L and reference range is 111-220 – code 220.

The LDH pretreatment level is based on the same laboratory results as the LDH pretreatment laboratory value and upper limits of normal. This data item records whether the laboratory value is normal (low) or above normal (high). Other values include when the test is known to be ordered but is not available in the medical record or unknown or not documented as being ordered. Example: LDH laboratory report shows 193U/L, range: 111-220, LDH falls in this range making it normal – code 0.

**AJCC TNM Staging**

Always remember the staging windows when it comes to AJCC TNM staging. For clinical classification purposes, include information from the time of diagnosis up until the time treatment begins. For cutaneous melanoma, this includes the first biopsy/excision, as this is typically the first diagnosis. This is known as microstaging. It also includes any clinical assessment or biopsy of regional lymph nodes and imaging to detect metastatic sites. Cutaneous melanoma in situ, T1, and T2 do not require imaging in order to stage the case.

Pathologic classification includes everything from time of diagnosis until the beginning of adjuvant therapy. In cutaneous melanoma, it includes excision after diagnosis of the primary site, which is typically a wide excision or re-excision. Pathologic melanoma in situ and T1 do not require the pathologic evaluation of regional lymph nodes to complete the pathologic classification you can use a cN to assign the pathologic prognostic stage group.

The tumor category (T) is determined by the Breslow’s thickness recorded in tenths of millimeters. If a melanoma is stated to be completely regressed, record as a T0. The mitotic rate was removed from the T1 category in the eighth edition. Primary tumor ulceration is based on the histopathologic exam and can be present in the initial biopsy or on re-excision. If ulceration is mentioned, then it should be recorded as such in the primary tumor category.

The lymph node category (N) is based on presence or absence of disease in the regional lymph nodes. A clinically occult lymph node is when the scans are negative for regional lymph node involvement, but there are microscopic confirmed regional lymph nodes (N1a, N2a, N3a). Those that are clinically detected by scans or physical exam only record based on number of clinically positive lymph nodes (N1b, N2b, N3b). When there is non-nodal locoregional metastasis like satellite metastasis, microsatellite metastasis, and in-transit metastasis record this in the lymph node category. Satellite metastasis is grossly visible (seen with the naked eye) cutaneous/subcutaneous metastasis within 2cm of the primary tumor. Microsatellite metastasis, as the name implies, is cutaneous/subcutaneous metastasis seen during the pathologic exam adjacent or deep to the primary melanoma. In-transit metastasis is clinically evident dermal and/or subcutaneous metastasis greater than 2 cm from the primary tumor, typically appearing between the primary and first echelon regional lymph nodes.

The distant metastasis category (M) is based on the presence or absence of disease distant to the primary tumor. The typical distant metastatic sites are subcutaneous tissue, muscle, or distant lymph node (M1a); lung (M1b); other visceral sites (M1c); central nervous system (M1d). LDH level is a subcategory for the distant metastasis category. The registrar would modify each subcategory based on the LDH values – 0 not elevated or 1 elevated.

**STORE Data Items**

The Standards for Oncology Registry Entry (STORE) manual has a few data items worth discussing. The recording of information on sentinel lymph node (SLN) biopsies for cutaneous melanoma was added for cases diagnosed January 1, 2018 and after. This information includes the number of sentinel lymph nodes pathologically examined. The registrar will record the total number of lymph nodes removed during a SLN biopsy/procedure and examined by a pathologist in the data item Sentinel Lymph Nodes Examined. Include this number in the data item Regional Lymph Nodes Examined as well because they are all regional to the primary tumor. The registrar will record the number of SLN found to contain metastasis during the SLN biopsy/procedure in the data item Sentinel Lymph Nodes Positive. For a cutaneous melanoma case, the registrar will count any isolated tumor cells (ITC) as positive SLN. The date that the patient undergoes SLN biopsy/procedure is recorded in the data item Date of Sentinel Lymph Node Procedure. If the SLN biopsy/procedure is the first or only surgery performed, also record as the data item Date First Surgical Procedure. When the SLN biopsy/procedure and regional lymph node dissection are performed on the same dates, record both dates in each data item. If the SLN biopsy/procedure and regional lymph node dissection are performed on different dates, then record the date that the SLN biopsy/procedure was performed in the data item Date of Sentinel Lymph Node Procedure.
Recording surgeries and biopsies for cutaneous melanoma it can get a bit confusing. Why? Because the first biopsy is typically one that excises the melanoma, not just a core biopsy. The question then remains: What do I code this first surgery as—a Surgical Diagnostic and Staging Procedure (SDSP) or as Surgical Procedure of Primary Site? Well, based on the guidelines in the STORE manual, it all comes down to margin status. If you have an excisional biopsy that has clear or only microscopic margins, then you are to code the procedure as Surgical Procedure of Primary Site. When the margins on an excisional biopsy are grossly positive or macroscopically involved, meaning the physician knows that the tumor was not entirely removed nor was it intended, then record the procedure as a SDSP.

For the data item Surgical Procedure of Primary Site, the registrar again comes down to margin status on knowing the correct code as found in the STORE manual, Appendix B. Use the codes found here to accurately reflect the surgeries performed on the primary site. An excisional biopsy that removes the entire tumor or only leaves microscopic margins involved is a code 27. The initial biopsy often removes all the recognizable melanoma. A re-excision, often referred to as a wide-local excision, is usually recommended to further reduce the risk of the melanoma recurring. This involves removing the melanocytes adjacent to the melanoma that may be susceptible to turning into melanoma themselves. A wide local excision removes a 1-2 cm margin around the site of the original melanoma in an elliptical fashion. However, the margin size varies depending on the depth of the tumor and how far it has spread into lower layers of the skin. Most wide local excisions are closed with stitches. Larger excisions can require skin grafts or skin flaps.

When it comes to recording the primary site, surgery consider what you know and the reason for the surgery. The codes 30–33 are those that are less than a wide local excision, margins less than 1 cm, or margin status unknown. When known that a shave biopsy was performed first followed by a re-excision (not meeting wide local excision guidelines), use code 31—shave biopsy followed by excision. The codes 45–47 are for wide local excisions meeting the guidelines of more than 1-cm margin as determined by the pathologist and noted on the pathology report. In the instance that the patient undergoes a shave biopsy followed by the wide local excision, and the operative report states a 1.5-cm margin was taken around previous melanoma site, and the pathology report on the wide local excision has no residual melanoma, you can code this surgery as 46—greater than 1-cm but less than 2-cm margin. This is the only situation where you can use the operative report margins to determine surgery code. If there is residual melanoma on the wide local excision pathology report, then you use the recorded margin by the pathologist. If not stated, you will down code the surgery to 30-33 range for whichever is applicable. As stated before, most if not all cutaneous melanoma primaries will undergo resection.

Systemic therapy is generally used for metastatic cutaneous melanoma. The use of immunotherapy for cutaneous melanoma is on the rise and showing great results. A couple of the top immunotherapy agents include Yervoy or Opdivo for unresectable, metastatic, or for disease progression in cutaneous melanoma. The more common chemotherapy agents include Trametinib for metastatic or unresectable cutaneous melanoma and Fotesmustine for disseminated cutaneous melanoma. There are others, such as Interferon, and melanoma peptide vaccine is also widely used.

**Conclusion**

In summary, the abstracting of cutaneous melanoma cases can feel overwhelming and daunting at times. To make it just a bit easier, it is best to take each part one at a time. Read the manuals and instructions for each data item that you are recording, and do not rely solely on the notes found in your oncology software or your memory. There are many items that have remained the same for cutaneous melanoma and that should give us some relief. Unfortunately, there is still confusion around surgery coding, and some of the new data items. This is expected during times of change, but it is our responsibility as registrars to educate ourselves on these new items, ask questions, and seek understanding, for when this occurs, our abstracting is solid and high quality, which I am sure we all are seeking to achieve.
Raising the Bar

Are You Ready for Curated Data and Skill Shifts?

Michele Webb, CTR

Cancer registrars have always endeavored to speed up manual tasks. From the card file systems to early computers, home-grown databases, networked and multifacility systems, and now, cloud-based systems, we have seen improvement. But is it enough? And with each advancement, our peers have voiced their fear about one day being replaced by a machine.

Currently, science and technology are accelerating a shift in workforce requirements, indicating that the skill sets registrars have used over the past 15 years will need to change. Data entry, input, and traditional spreadsheet or presentation skills are still needed. However, advanced critical thinking, social, and technical skills are increasingly in demand.

As organizations rethink how the work will be organized and address skill shifts needed to stay competitive, so should cancer registrars. There will be a new emphasis on retraining and the deployment of continuous learning opportunities as the health care workforce moves to cross-functional and team-based workflows.

Retraining should raise the skill capacity for all cancer registrars, whether they are currently employed, entering the profession as new hires, or students who will be seeking employment in the future. By strategically preparing for current and future needs, the industry-specific knowledge and understanding of oncology culture can be preserved while everyone acquires new, higher level critical skills.

If you have not already read the National Cancer Registrars Association (NCRA)’s 2019-2023 Strategic Management Plan, I encourage you to do so. The vision of curated data, advancement of outcomes, and NCRA’s mission to empower and advance the registrar’s skill sets points directly to the critical skills discussed here today. The plan’s vision of curated data and advancement of outcomes suggests that registrars need to develop or enhance their critical thinking skills to meet employers’ expectations in the future.

Health care thought leaders and oncology workforce trends categorize advanced critical thinking into analytical skills, communication, creativity, open-mindedness, and problem-solving. Let’s take a closer look:

• **Analytical skills**: Registrars with analytical skills can examine data from multiple sources, determine if there are any problems with its validity or accuracy, discuss its use and relevance to the health care delivery system, draw inferences from the data sets, and appropriately explain its implications and best practice application.

  • **Communication**: It is important that registrars share their ideas and observations effectively. Engaging in brainstorming, think-tank, and strategic planning sessions to identify reasonable solutions to complex problems will demonstrate value and contribution to the bottom line. Active listening, assessment, collaboration, teamwork, and exceptional verbal and written skills are essential.

  • **Creativity**: Using data to effectively tell a story requires creativity and novel thinking. Registrars must be able to insert registry data with other sources of information to propose solutions and facilitate work by cross-functional teams. A flexible mindset, ability to conceptualize the data, and the imagination to visually connect the points is necessary.

  • **Open-Mindedness**: Registrars should put aside assumptions or judgments to analyze the information used from any source. It is imperative to remain objective and evaluate the data and possibilities for use with fairness, humility, and objectivity.

  • **Problem-Solving**: Organizations need registrars who use critical thinking skills to suggest practical solutions. Effective problem-solving requires attention to detail, identification of patterns and trends, and the offering of reasonable, innovative solutions aligned with the clinical decision-making process.

  The cancer registrar should consider developing critical thinking skills such as adaptability, flexibility, inductive and deductive reasoning, strategic planning, quantitative and qualitative data management, and statistical analysis. Additional skills such as business and emotional intelligence and understanding of consumer behaviors are also increasingly in demand.

Next, the cancer registrar should take a personal inventory of the skills they currently have and those they need for the future based on employer’s expectations and how the registrar’s work or desired career path will change.

Ultimately, the goal of retraining and building new skills should focus on the logic and mechanics of how registrars can transform data into real-time information.
that can be immediately used by health care providers and administrators to improve the delivery process and effective outcomes. Achieving this goal will require an open mind, willingness to discard outdated or inefficient processes, and quick replacement with new workflows based on curated data derived from multiple sources.

The focus and goals of the learning activities used by the registrar to build skills must also change. The best learning solutions will incorporate practical how-to instruction with real-time examples using case studies, journal articles, tool kits, media presentations, e-learning, and traditional classroom activities. Learning activities must demonstrate relatable examples of how the high-value skills are used to transform routine data into an effective tool that tells a story and offers simple solutions to solve business and programmatic problems.

A unified call to action addresses the fear of being replaced by a machine and equips the cancer registry workforce with in-demand, high value, critical thinking skills. There are limitless opportunities to participate in retraining to ensure all registrars have a seat at the table now and in the future. All that remains is to see how each cancer registrar will contribute and participate in the process.

References

Michele is a cancer registry speaker and author who works with SCL Healthcare in Colorado and Montana. She is committed to helping others grow and expand their influence as oncology health care leaders and mentors. Outside of work, Michele enjoys diamond painting, getting lost in a good book, and her fur babies, Dolly and Cooper. Your feedback and comments are welcomed by email at michele@michelewebb.com.
IMPROVED GEOCODING OF CANCER REGISTRY ADDRESSES IN URBAN AND RURAL OKLAHOMA

After reading this article and taking the quiz, the participants will be able to:
• Discuss the importance of valid geocoding in rural counties in order to conduct high quality studies.
• Understand methods to improve the accuracy of geocodes.
• Understand the strength and limitations of each reference data set within the study.

1. The study utilized which method FIRST to improve the accuracy of the geocodes?
   a) Enhanced 911 (E911)
   b) North American Association of Central Cancer Registries (NAACCR)
   c) NAVTEQ
   d) United States Postal Service LACSLink database

2. What term describes the study of geographic variation of diseases or disease risk factors?
   a) Environmental epidemiology
   b) Forensic epidemiology
   c) Spatial epidemiology
   d) Occupational epidemiology

3. Oklahoma has a large population of which group?
   a) Asian
   b) American Indian/Alaska Natives
   c) Japanese American
   d) None of the above

4. The AccuMail Frameworks software—specifically, the LACSLink database—was used for what purpose?
   a) Standardization, validation, and enhancement
   b) Correction, evaluation, and comparison
   c) Standardization, comparison, and enrichment
   d) Correction, validation, and comparison

5. According to Table 3, which county had the highest percentage of post office boxes?
   a) Cleveland
   b) Oklahoma
   c) Logan
   d) Jackson

6. All are considered 1 of the 4 geocoders used to geocode the address information to street addresses EXCEPT
   a) LACSLink.
   b) NAACCR.
   c) NAVTEQ.
   d) South Western Oklahoma Development Authority (SWODA) E911.

7. What are the 2 methods cancer registries can consider for obtaining missing street level address information?
   a) ZIP code searches / yearly follow-up letters to patient
   b) Internet searches / patient phone calls
   c) Address correction / follow-up directly with a patient
d) Investigation / follow-up with patient contact

8. What was the purpose of the study?
   a) Improve accuracy of cancer registry data
   b) Enable accurate geographic research in epidemiology
   c) Increase yearly follow-up response
   d) Decrease the geocoding bias

9. Using the set theory, what represents the number of addresses that could be geocoded only when they were corrected?
   a) $A_{\text{corrected}}$ total
   b) $A_{\text{corrected}}$ only
   c) $A_{\text{original}}$ only
   d) $A_{\text{original}}$ total

10. All of the following are accurate statements based on the pairwise comparisons of geocoders EXCEPT
    a) The study kept results that were geocoded at street level with at least 90% confidence.
    b) The study compared the number of addresses that were geocoded by each geocoder in sets of 3.
    c) The performance of the address correction was measured and then assessed the relative performance of each geocoder.
    d) The smallest size of unions occurred within the SWODA region.

---

Purchase Quiz to Earn CE:
1. Go to http://www.cancerregistryeducation.org/jrm-quizzes
2. Select quiz and “Add to Cart” (You may be prompted to login using your NCRA login).
3. Continue through the checkout process.
4. Once purchase is complete, the quiz will load automatically into “My Learning Activities” page.
The *Journal of Registry Management*, official journal of the National Cancer Registrars Association (NCRA), announces a call for original manuscripts on registry methodology or research findings related to the 7 subjects listed below and related topics.

**Topics:**
1. Birth Defects Registries
2. Cancer Registries
   a. AJCC TNM Stage
   b. Cancer and Socioeconomic Status
   c. Cancer and Health Disparities
3. Trauma Registries
4. Recruitment, Training, and Retention
5. Public Relations
6. Quality Review
7. Registry Management

Contributed manuscripts are peer-reviewed prior to publication. Manuscripts of the following types may be submitted for publication:

1. **Methodology Articles** addressing topics of broad interest and appeal to the readership, including methodological aspects of registry organization and operation.
2. **Research articles** reporting findings of original, reviewed, data-based research.
3. **Primers** providing basic and comprehensive tutorials on relevant subjects.
4. **“How I Do It” Articles** describe tips, techniques, or procedures for an aspect of registry operations that the author does particularly well. The “How I Do It” feature in the *Journal* provides registrars with an informal forum for sharing strategies with colleagues in all types of registries.
5. **Opinion papers/editorials** including position papers, commentaries, essays, and interviews that analyze current or controversial issues and provide creative, reflective treatments of topics related to registry management.
6. **Bibliographies** which are specifically targeted and of significant interest will be considered.
7. **Letters to the Editor** are also invited.

Address all manuscripts to: Danette Clark, BS, RMA, AAS, CTR, Editor-in-Chief, *Journal of Registry Management*, (973) 971-5189, JRMEditor@ncra-usa.org.

Manuscript submission requirements are given in “Information for Authors” found near the back of each *Journal* and on the NCRA website at [http://www.ncra-usa.org/jrm](http://www.ncra-usa.org/jrm).
Journal of Registry Management

INFORMATION FOR AUTHORS

Journal of Registry Management (JRM), the official journal of the National Cancer Registrars Association, invites submission of original manuscripts on topics related to management of disease registries and the collection, management, and use of cancer, trauma, AIDS, and other disease registry data. Reprinting of previously published material will be considered for publication only when it is of special and immediate interest to the readership. JRM encourages authorship by Certified Tumor Registrars (CTRs); special value is placed on manuscripts with CTR collaboration and publication of articles or texts related to the registry profession. CTR continuing education (CE) credits are awarded; a published chapter or full textbook article equals 5 CE hours. Other published articles or documents equal CE hours. All correspondence and manuscripts should be addressed to the Danette Clark, BS, RMA, AAS, CTR, Editor-in-Chief at JRMEditor@ncra-usa.org or (973) 971-5189.

Manuscripts may be submitted for publication in the following categories: Articles addressing topics of broad interest and appeal to the readership, including Methodology papers about registry organization and operation; Research papers reporting findings of original, reviewed, data-based research; Primers providing tutorials on relevant subjects; and How I Do It papers are also solicited. Opinion papers/editorials including position papers, commentaries, and essays that analyze current or controversial issues and provide creative, reflective treatments of topics related to registry management; Letters to the Editor; and specifically-targeted Bibliographies of significant interest are invited.

The following guidelines are provided to assist prospective authors in preparing manuscripts for the Journal, and to facilitate technical processing of submissions. Failure to follow the guidelines may delay consideration of your manuscript. Authors who are unfamiliar with preparation and submission of manuscripts for publication are encouraged to contact the Editor for clarification or additional assistance.

Submission Requirements

Manuscripts. The terms manuscripts, articles, and papers are used synonymously herein. Email only submission of manuscripts is encouraged. If not feasible, submit the original manuscript and 4 copies to the Editor. Manuscripts should be double-spaced on white 8-1/2” x 11” paper, with margins of at least 1 inch. Use only letter-quality printers; poor quality copies will not be considered. Number the manuscript pages consecutively with the (first) title page as page one, followed by the abstract, text, references, and visuals. The accompanying cover letter should include the name, mailing address, email address, and telephone number of the corresponding author. For electronic submission, files should be IBM-compatible format in Corel WordPerfect™, Microsoft® Word for Windows®, or converted to ASCII code.

Manuscripts (Research Articles). Articles should follow the standard format for research reporting (Introduction, Methods, Results, Discussion, References), and the submission instructions outlined above. The introduction will normally include background information, and a rationale/justification as to why the subject matter is of interest. The discussion often includes a conclusion subsection. Comprehensive references are encouraged, as are an appropriate combination of tables and figures (graphs).

Manuscripts (Methodology/Process Papers). Methodology papers should follow the standard format for research reporting (Introduction, Methods, Results, Discussion), and for explanatory papers not reporting results (Introduction, Methods, Discussion), as well as the submission instructions outlined above.

Manuscripts (“How I Do It” articles). The “How I Do It” feature in the Journal provides registrars with a forum for sharing strategies with colleagues in all types of registries. These articles describe tips, techniques, or procedures for an aspect of registry operations that the author does particularly well. When shared, these innovations can help registry professionals improve their skills, enhance registry operations, or increase efficiency. “How I Do It” articles should be 1,500 words or less (excluding references) and can contain up to 2 tables or figures. To the extent possible, the standard headings (Introduction, Methods, Results, Discussion) should be used. If results are not presented, that section may be omitted. Authors should describe the problem or issue, their solution, advantages (and disadvantages) to the suggested approach, and their conclusion. All submitted “How I Do It” articles will have the benefit of peer/editorial review.

Authors. Each author’s name, degrees, certifications, title, professional affiliation, and email address must be noted on the title page exactly as it is to appear in publication. The corresponding author should be noted, with mailing address included. Joint authors should be listed in the order of their contribution to the work. Generally, a maximum of 6 authors for each article will be listed.

Title. Authors are urged to choose a title that accurately and concisely describes the content of the manuscript. Every effort will be made to use the title as submitted; however, the Journal of Registry Management reserves the right to select a title that is consistent with editorial and production requirements.

Abstract. A brief abstract must accompany each article or explanatory paper. The abstract should summarize the main point(s) and quickly give the reader an understanding of the manuscript’s content. It should be placed on a page by itself, immediately following the title page.

Length. Authors are invited to contact the Editor regarding submission of markedly longer manuscripts.


Visuals. Use visuals selectively to supplement the text. Visual elements—charts, graphs, tables, diagrams, and figures—will be reproduced exactly as received. Copies must be clear and properly identified, and preferably emailed. Each visual must have a brief, self-explanatory title. Submit each visual on a separately numbered page at the end of the manuscript, following the references.

Attribution. Authors are to provide appropriate acknowledgment of products, activities, and support especially for those articles based on, or utilizing, registry data (including acknowledgment of hospital and central registrars). Appropriate attribution is also to be provided to acknowledge federal funding sources of registries from which the data are obtained.

References. References should be carefully selected, and relevant. References must be numbered in order of their appearance in the text. At the end of the manuscript, list the references as they are cited; do not list references alphabetically. Journal citations should include author, title, journal, year, volume, issue, and pages. Book citations should include author, title, city, publisher, year, and pages. Authors are responsible for the accuracy of all references. Examples:


Key words. Authors are requested to provide up to 5, alphabetized key words or phrases which will be used in compiling the Annual Subject Index.

Affirmations

Copyright. Authors submitting a manuscript do so on the understanding that if it is accepted for publication, copyright in the article, including the right to reproduce the article in all forms and media, shall be assigned exclusively to NCRA. NCRA will not refuse any reasonable requests by the author(s) for permission to reproduce any of his or her contributions to the Journal. Further, the manuscript’s accompanying cover letter, signed by all authors, must include the following statement: “We, the undersigned, transfer to the National Cancer Registrars Association, the copyright for this manuscript in the event that it is published in Journal of Registry Management.” Failure to provide the statement will delay consideration of the manuscript. It is the author’s responsibility to obtain necessary permission when using material (including graphs, charts, pictures, etc) that has appeared in other published works.

Originality. Articles are reviewed for publication assuming that they have not been accepted or published previously and are not under simultaneous consideration for publication elsewhere. If the article has been previously published or significantly distributed, this should be noted in the submission for consideration.

Editing. Journal of Registry Management reserves the right to edit all contributions for clarity and length. Minor changes (punctuation, spelling, grammar, syntax) will be made at the discretion of the editorial staff. Substantive changes will be verified with the author(s) prior to publication.

Peer Review. Contributed manuscripts are peer-reviewed prior to publication, generally by 3 reviewers. The Journal Editor makes the final decision regarding acceptance of manuscripts. Receipt of manuscripts will be acknowledged promptly, and corresponding authors will be advised of the status of their submission as soon as possible.

Reprints. Authors receive 5 complimentary copies of the Journal in which their manuscript appears. Additional copies of reprints may be purchased from the NCRA Executive Office.
FINDING YOUR DREAM JOB IS JUST A CLICK AWAY

Visit NCRA’s Job Bank at www.ncra-usa.org/job!

**VIEW JOBS**
Access the newest jobs available to cancer registrars seeking employment.

**POST AN ANONYMOUS RESUME**
Whether you are actively seeking work or curious about what is available, your online resume is your ticket to great job offers!

**DESIGN A PERSONAL JOB ALERT**
Create job alerts and never let a matching job opportunity pass you by! New jobs that match your search criteria will be e-mailed directly to you.

**CREATE A FREE JOB SEEKER ACCOUNT**
Your personal job seeker account will allow you to find jobs, manage your resume, and set-up job alerts. Get started today!

NCRA’S JOB BANK www.ncra-usa.org/job
Printed on SFI fiber sourcing paper with non-petroleum, vegetable based inks and manufactured with renewable electricity