Estimating and Mapping Health Literacy in the State of Missouri

NICOLE LURIE, LAURIE T. MARTIN, TEAGUE RUDER, JOSE J. ESCARCE, BONNIE GHOSH-DASTIDAR, DAN SHERMAN, CHLOE E. BIRD, ALLEN FREEMONT

WR-735-MFH
December 2009
Prepared for the Missouri Foundation for Health
EXECUTIVE SUMMARY

Health literacy refers to literacy within a health context and is defined as the “degree to which individuals have the capacity to obtain, process, and understand basic health information and services needed to make appropriate health decisions”.\(^1\) Findings from the 2003 National Assessment of Adult Literacy (NAAL)\(^2\) paint a grim picture of health literacy in America. About a third (36%) of adults have limited health literacy\(^3\)-\(^6\) and find it difficult to calculate the appropriate dose of an over-the-counter medication for a child or comprehend a consent form.\(^6\) Of these approximately 80 million adults with limited health literacy, 31 million have difficulty understanding directions for taking medicine.

National, state and local stakeholders seeking to improve health and health care increasingly recognize that low health literacy contributes to poor health and gaps in care. Yet, resources to screen individuals for low health literacy are limited, and individually-tailored interventions may not be feasible or affordable in the long run. We offer an alternative population-based approach to identify geographic areas with a high concentration of individuals with low health literacy (i.e., ‘hot spots’) where interventions aimed at low health literacy populations might be most effective in improving health and health care quality.

To facilitate efforts to identify hot spots of low health literacy, we developed predictive models using data from the 2003 National Assessment of Adult Literacy (NAAL), an in-person assessment of English language literacy among a nationally representative sample of U.S. adults age 18 and over. We estimated mean health literacy scores as well as the probability of being classified by the NAAL as having ‘above basic’ health literacy (i.e., health literacy in the ‘intermediate’ or ‘proficient’ range). Using these models and information about the demographic makeup of census areas (public use microdata areas (PUMAs) and census tracts), we generated estimates of health literacy for each geographic area. We then mapped the results to identify ‘hot spots’ of low health literacy across Missouri.

Individuals living in northeast St. Louis and the western part of Kansas City had the lowest estimated average health literacy scores, with one area in East St. Louis having the lowest health literacy in the state (see Fig 1). Most areas of central Missouri and the bootheel region have populations in which 60% to 70% has ‘above basic’ health literacy (see Figure 2).
The models and maps presented in this report provide a potential tool to assist stakeholders in the efficient identification of communities with low health literacy and advance action on health literacy in the state of Missouri. The maps provide an easy-to-use resource to identify hot spots of low health literacy within the state so that interventions may be effectively targeted in areas in area of greatest need. The maps may also help foster collaboration with local stakeholders by facilitating dialogue about health literacy in their community. Finally, the maps are readily understandable by community groups and can be used by stakeholders to promote buy-in from community members.
CHAPTER 1. INTRODUCTION

Health literacy, or literacy within a health context, is defined as the “degree to which individuals have the capacity to obtain, process, and understand basic health information and services needed to make appropriate health decisions” and is essential to navigating the health care system.\(^1\) However, the complexity of the health care system, the medical jargon used by many providers,\(^7\) and the exposure to novel health concepts (many times while under a great deal of stress), all have the potential to exceed patient’s health literacy skills, even among adults with adequate literacy.\(^8\) Therefore, most experts believe that the prevalence of limited literacy is higher when considered within a health context.

Low health literacy (LHL) remains a formidable barrier to reducing gaps in quality of care and improving health outcomes. LHL is common, affecting approximately a third of the population, with a higher prevalence among the elderly, those with low income and education, and certain minority groups. LHL may contribute to suboptimal care and outcomes through limited understanding of the origin, nature, and likely course of health problems,\(^9-11\) lower participation in screening programs,\(^12-14\) reduced ability to act on and understand the advice of a health professional,\(^15\) and limited ability to access and navigate the health care system.\(^16-18\) Related constructs of linguistic isolation and limited English proficiency are also critically important factors in areas with growing immigrant populations, as language barriers can also compromise the quality of care individuals receive through decreased access to primary and preventive care services, decreased patient adherence, and diminished patient satisfaction.

Health departments, health plans, and other community organizations seeking to improve health and health care recognize that such improvements require them to address gaps in care related to low health literacy. However, there are limited resources to screen individuals for low health literacy, and individually-tailored interventions may not be feasible or cost effective in the long run. Given that many interventions to improve health outcomes for individuals with low health literacy require resources to train health care workers (e.g., pharmacists, physicians, health educators) or make modifications to physical spaces or educational materials, it is not feasible to implement these interventions universally. An alternate approach is a population-based one—to identify and target geographic areas with a high concentration of individuals with low health literacy (i.e., ‘hot spots’) where interventions aimed at low health literacy populations might be most effective in improving health and health care quality.

One challenge, however, is the lack of data on neighborhood, or community-level, health literacy. To help the state of Missouri overcome this barrier, we developed a predictive model of health literacy based on data from the National Assessment of Adult Literacy, which when applied to census-derived variables for a given geographic region, provides an estimate of the mean health literacy scores for area residents or the percent of adults with ‘above basic’ health literacy. These estimates can then be applied to GIS mapping and decision tools to help stakeholders in Missouri quickly visualize where individuals with low health literacy reside and to develop interventions in those areas.

This report summarizes the methodology used to develop the predictive model and maps (Chapter 2), and describes the geographic distribution of health literacy in Missouri (Chapter 3). Finally,
we summarize the study findings and highlight the utility of the model and maps to advance action on health literacy (Chapter 4).
CHAPTER 2. METHODS

This chapter provides an overview of methods for developing the predictive model of health literacy and for mapping the predicted health literacy scores at two geographic levels based on census data (the Public Use Microdata Area (PUMA) and the census tract). We discuss the data, measures, and methods in greater detail in Appendices A-C, respectively.

Data

Health Literacy Data from the National Assessment of Adult Literacy (NAAL). Data from the Health Literacy Component of the 2003 NAAL household sample was used to create predictive models of health literacy. The NAAL is an in-person assessment of English language literacy among a nationally representative sample of US adults age 16 and over.

American Community Survey (ACS). We used Public Use Microdata (PUMS) from the 2007 American Community Survey (ACS) to produce estimates of health literacy for the 41 Public Use Microdata Areas (PUMA) within Missouri. A PUMA has a population of approximately 100,000 individuals and is the smallest sub-state geographic area currently available with individual demographic data in the ACS. Current ACS data are not currently available at the census tract level (but are scheduled for release in 2009).

2000 Census. To enable the prediction of health literacy at the census tract level, we used aggregate demographic data from the 2000 Census Summary File 3 (SF3). Census tracts are small, relatively permanent statistical subdivisions of a county, with approximately 2,500 to 8,000 people residing in each tract. We provide additional information on these data sources in Appendix A.

Study variables

Health Literacy. Health literacy is defined as the “degree to which individuals have the capacity to obtain, process, and understand basic health information and services needed to make appropriate health decisions” [1]. The NAAL measured health literacy on a scale ranging from 0 to 500 points, with a national mean of 245 points. In addition to the continuous scores, we used the four performance levels created by the National Research Council to reflect an individual’s ability to successfully complete tasks of a given difficulty: below basic, basic, intermediate and proficient (Table 1) [3].

Socio-demographic characteristics. To maximize the utility of the model to stakeholders hoping to advance action on health literacy, we restricted the model to sociodemographic characteristics available in the 2000 census and American Community Survey (ACS). We include age, race/ethnicity, education, gender, marital status, poverty status, metropolitan statistical area, language spoken in home, and years residing in the United States. We provide additional information on study variables in Appendix B.
### Table 1. Selected Health Tasks by Health Literacy Score

<table>
<thead>
<tr>
<th>Health Literacy Score</th>
<th>Percent of Population</th>
<th>Activity</th>
</tr>
</thead>
</table>
| **Proficient** (310-500) | 12% | • Calculating one’s personal share of employer health costs using a table  
• Finding definitions for complex medical terms  
• Interpreting legal documents and applying to specific health care situations |
| **Intermediate** (226-309) | 53% | • Determining a healthy weight range for one’s height based on a body mass index (BMI) graph  
• Finding the age range for certain childhood vaccines using a recommendation chart  
• Determining the times at which one can take a medication based on instructions on the prescription label  
• Identifying three substances that may interact negatively with an over-the-counter drug to produce side effects, based on the information provided on the package |
| **Basic** (185-225) | 22% | • Interpreting a clearly-written pamphlet to determine two reasons a person should be tested for a disease even without symptoms  
• Explaining how a chronic disease may be asymptomatic based on a one-page article about the medical condition. |
| **Below Basic** (0-184) | 14% | • Interpreting a clearly-written pamphlet to determine how often one should have a certain medical test  
• Identifying what one can and cannot drink before a medical test based on a short set of instructions  
• Circle the date of a medical appointment on an appointment slip |


### Analysis

We first developed a linear model that predicted the mean (or average) health literacy score. We then used a linear probability model to predict the probability of scoring at the ‘above basic’ level of health literacy (e.g., intermediate or proficient). Of note, the latter model predicts the complement of basic or below basic health literacy.

After developing the regression models, we used information about the demographic makeup of each PUMA or census tract (e.g., 47% female, 53% male), obtained from the ACS or 2000 census, to generate an estimate of both the share of adults at the ‘above basic level’ of health literacy for each geographic area and the average level of health literacy on a 0-500 point scale. We then produced maps to display these estimates. We provide detailed information on our methodology in Appendix C.
CHAPTER 3. MAPS OF HEALTH LITERACY IN MISSOURI

The maps presented below were generated by applying the models predicting health literacy to census data. The first set of maps presents estimates of mean health literacy (Figures 4-6) and the second set presents estimates of the proportion of individuals within a given geographic region with ‘above basic’ health literacy skills (Figures 7-9). We present the results of the models used to predict mean health literacy and ‘above basic’ health literacy in Appendix D.

Maps of Health Literacy in Missouri

To facilitate the interpretation of the health literacy maps and graphs, we include a reference map of the major cities and highways in Missouri in relation to its 41 PUMAs and 1312 census tracts (Figure 3).

Figure 3. Reference Map for the state of Missouri
Mean Health Literacy Levels

The mean health literacy scores for PUMAs and census tracts are presented in Figures 4 - 6. We used five categories of health literacy, each with a 15 point range. Areas shaded in light or dark green have a mean health literacy score above the national average of 245. Yellow areas have a mean health literacy level that is in the ‘intermediate’ range, but that falls slightly below the national average. Red and orange shaded regions have a mean health literacy that falls within the level of ‘basic’ health literacy (range: 185-225). In terms of functional health literacy tasks, the average individual residing in a red or orange shaded region may, at most, be able to say why a person should be tested for a disease, based on information in a pamphlet. However, many adults in these regions have health literacy skills limited to finding the date of an appointment written on a reminder card. Figure 4 suggests areas with the lowest average health literacy are clustered in Kansas City and East St. Louis.

Figure 4. Mean Health Literacy Levels for Population within PUMAs
Figure 5. Mean Health Literacy Levels for Population within Census Tracts
In addition to St. Louis and Kansas City, there are isolated pockets of low health literacy in the bootheel region of Missouri. Figure 6 shows that the majority of census tracts in the region have a mean health literacy below the national average, with six having an average health literacy representative of ‘basic’ literacy skills.

Although census tracts with similar levels of health literacy may vary significantly in their demographic compositions, we indicate the average prevalence of selected demographic characteristics for each of the five levels of health literacy in Table 2 below.

Table 2. Select Demographic Characteristics by Average Health Literacy Score

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Percent in Population</th>
<th>Average Health Literacy Score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Red ≤215.0</td>
</tr>
<tr>
<td>Age 65 – 74</td>
<td>10</td>
<td>12</td>
</tr>
<tr>
<td>Age 75+</td>
<td>9</td>
<td>11</td>
</tr>
<tr>
<td>Education &lt; HS</td>
<td>21</td>
<td>38</td>
</tr>
<tr>
<td>College or more</td>
<td>19</td>
<td>7</td>
</tr>
<tr>
<td>Male</td>
<td>48</td>
<td>43</td>
</tr>
<tr>
<td>Married</td>
<td>54</td>
<td>24</td>
</tr>
<tr>
<td>MSA</td>
<td>66</td>
<td>100</td>
</tr>
<tr>
<td>&lt;100% FPL</td>
<td>13</td>
<td>31</td>
</tr>
<tr>
<td>Black</td>
<td>14</td>
<td>93</td>
</tr>
<tr>
<td>Hispanic</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>U.S. Born</td>
<td>97</td>
<td>99</td>
</tr>
</tbody>
</table>
For census tracts with the lowest average health literacy (shaded in red), on average approximately 23% are over the age of 65, 38% have less than a high school diploma, 24% are married, 31% live below the FPL, and 93% are of black race. Among census tracts with the highest average health literacy (shaded in green), 12% were above the age of 65, 7% had less than a high school education, 63% were married, 6% lived below the FPL, and 3% were black. The U-shaped relationship between health literacy and MSA suggests that both highly educated adults and those with the lowest education cluster in metropolitan areas. In contrast, a higher proportion of the suburban and rural population is composed of adults with a high school diploma or some college.

‘Above Basic’ Health Literacy

Maps representing the proportion of adults with ‘above basic’ literacy for PUMAs and census tracts are presented in Figures 7 - 9. We distinguish five levels of health literacy ranging from those in which fewer than half of the residents have health literacy scores at the ‘above basic level’ (red) to areas where over 80% score at the ‘above basic’ level (green). These maps suggest that in areas shaded red, over half of the residents cannot complete tasks such as determining what time a person should take a prescription medication, based on information from the prescription drug label.

Figure 7 indicates that the areas with a higher prevalence of low health literacy are clustered in East St. Louis and Kansas City. Figure 8 indicates that certain census tracts in the southwest corner of the state, the Missouri bootheel, and a limited number of rural areas also have populations where between 40% and 50% have health literacy at the ‘below basic’ level. We provide additional detail and alternative ways to illustrate the geographic distribution of health literacy in Appendix E.
Figure 7. Percent of PUMA Scoring at the ‘Above Basic’ Level of Health Literacy

Missouri

% Above Basic
- <= 50%
- 50.1% - 80%
- 80.1% - 70%
- 70.1% - 80%
- >= 80%

Kansas City

St Louis
Figure 8. Percent of Census Tract Residents Scoring at the ‘Above Basic’ Level of Health Literacy
Several census tracts in the bootheel region of Missouri have a high proportion of individuals with low health literacy. Census tracts highlighted in orange have between 40 and 50% of the resident population has below basic health literacy skills.

The average prevalence of selected demographic characteristics for each of the five levels of 'above basic' health literacy are shown in Table 3 below. Results were similar to those found for mean health literacy categories.

Table 3. Select Demographic Characteristics by Percent ‘Above Basic’ Health Literacy

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Percent in Population</th>
<th>Percent ‘Above Basic’ Health Literacy</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>≤50</td>
<td>50.1 - 60.0</td>
</tr>
<tr>
<td>Age 65 – 74</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>Age 75+</td>
<td>9</td>
<td>9</td>
</tr>
<tr>
<td>Education &lt; HS</td>
<td>21</td>
<td>36</td>
</tr>
<tr>
<td>College or more</td>
<td>19</td>
<td>7</td>
</tr>
<tr>
<td>Male</td>
<td>48</td>
<td>44</td>
</tr>
<tr>
<td>Married</td>
<td>54</td>
<td>26</td>
</tr>
<tr>
<td>MSA</td>
<td>66</td>
<td>99</td>
</tr>
<tr>
<td>&lt;100% FPL</td>
<td>13</td>
<td>30</td>
</tr>
<tr>
<td>Black</td>
<td>14</td>
<td>84</td>
</tr>
<tr>
<td>Hispanic</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>U.S. Born</td>
<td>97</td>
<td>97</td>
</tr>
</tbody>
</table>
For census tracts where less than 50% of the population is ‘above basic’, on average approximately 19% are over the age of 65, 36% have less than a high school diploma, 26% are married, 30% live below the FPL, and 84% are of black race. Among census tracts where over 80% of the population is ‘above basic’, 9% were above the age of 65, 5% had less than a high school education, 61% were married, 8% lived below the FPL, and 4% were black. The U-shaped relationship between health literacy and MSA suggests that both highly educated adults and those with the lowest education cluster in metropolitan areas. In contrast, a higher proportion of the suburban and rural population is composed of adults with a high school diploma or some college.

**Differences between the model and individual predictors of health literacy**

Characteristics such as education and income are often thought of as markers for health literacy. Maps of the percentage of individuals without a high school diploma (Figure 10a), and with an income below 100% of the Federal Poverty Limit (Figure 10b) are presented below. Orange areas indicate that between a quarter and a third of the population has less than a high school diploma or lives below the poverty line, and red areas indicate that between a third and two thirds of the population has those characteristics. While both education and income suggest, as does the model, that parts of St. Louis and Kansas City have a high proportion of individuals with low health literacy, they both also suggest a high concentration of individuals with low health literacy in the south, and particularly the southeastern part of the state. However, demographic characteristics such as education underestimate the health literacy of those in the lowest categories (e.g., less than HS) when compared to the predictive models (see Appendix D). Thus, both for estimating individual and area-level health literacy, our model predicts health literacy better than any single indicator. Demographic maps of the Missouri bootheel can be found in Appendix F.

**Figure 10a. Percent Without HS Diploma**  **Figure 10b. Percent with Income < FPL**
Examining Health Literacy in Rural Areas.

As can be seen in Figures 4 - 9, relatively few census tracts in rural areas are identified as ‘hot spots’ of low health literacy (shown in red). The absence of hot spots is not an indication of high health literacy. Rather, it indicates that rural areas do not have the lowest health literacy in the state, contrary to what some might expect. However, the majority of rural areas (shaded in yellow) still have roughly 1 in 3 individuals with health literacy at the ‘below basic’ level, or 60%-70% above basic health literacy (Figure 8).

Although the models of health literacy identify characteristics such as race as significant predictors of health literacy, being over 65 years of age is an even stronger predictor. Given the prevalence of older individuals with less than a high school diploma in rural areas (Figure 11), it is understandable that stakeholders have identified low health literacy as a problem in these communities. However, only about 19% of the population is over the age of 65 (see Table 3), and older populations are fairly spread out across rural areas and are not concentrated in a single census tract or PUMA. In other words, the population is not segregated by age. As a result, there is not a high enough prevalence of older individuals in any geographic region to significantly lower the average health literacy locally. The diffusion of this population does not diminish, however, the need for health literacy interventions among this population. Rather, it suggests that in place of a broad community level intervention more applicable in the inner cities of St. Louis and Kansas City, a more refined intervention targeted to those ages 65 and older may be warranted in rural areas.

Using the Maps

The maps we have produced are a tool that will be most effectively used in combination with local stakeholder knowledge of the community of interest, and the specific health literacy needs of the community. We present three examples below (Figure 12), providing suggestions for how these data and maps may be used to advance action on health literacy, with the caveat that these approaches could be refined in concert with the insights of local stakeholders, clinicians and other service providers.

Census tract A, in a rural, southwestern part of the state contains a high proportion of Hispanic immigrants, individuals living below the poverty line, and individuals with low educational attainment.
Interventions and services for individuals with low health literacy in this area may focus on the use of community health workers for assistance navigating the system, ensuring health materials are written at a low grade level and appropriately translated, working with the media to get culturally sensitive health messages out and in a format accessible to this population, expanding English as a Second Language (ESL) classes in the area, and working with pharmacies and pharmacists to develop tools and materials to ensure full comprehension of medication directions, and contraindications. Given the high prevalence of recent immigrants to the United State and this area of Missouri in particular, stakeholders interested in addressing challenges related to low health literacy in this area may use these maps to engage prominent citizens and local leaders who have a stronger understanding of the barriers faced by this population.

Figure 12. Examples of Three Census Tracts with Low Health Literacy

Census Tract A
Mean HL = 230.5
Percent above basic = 56.0%
Education less than HS: 42%
Race/ethnicity: 30% Hispanic, 64% white
Below FPL: 22%
Age 65 or older: 13%
English spoken in home: 72%

Census Tract B
Mean HL = 232.4
Percent above basic = 60.0%
Education less than HS: 42%
Race/ethnicity: 89% white, 10% Hispanic
Below FPL: 24%
Age 65 or older: 28%
English spoken in home: 90%

Census Tract C
Mean HL = 207.7
Percent above basic = 37.0%
Education less than HS: 49%
Race/ethnicity: 100% Black
Below FPL: 42%
Age 65 or older: 22%
English spoken in home: 96%

Census tract B in the Missouri bootheel contains a high percent of low income, low education, elderly individuals. Given that elderly individuals often manage several medication regimens at one time, health literacy interventions in this geographic region may consider focusing on prescription drug labeling, pharmacy reminder systems and medication management tools. Additional health literacy work may be conducted around helping members understand and enroll in Medicare, a notoriously challenging process necessary to obtaining care and prescription drug coverage. Here, in addition to local stakeholders,
support may be obtained from national organizations such as the AARP or other organizations dedicated to improving the lives of older individuals.

Finally, census tract C in inner city St. Louis, contains a high percent of low income, low education, single, black women. While the average health literacy may be similar to the other census tracts described above, the health needs, health literacy interventions and potential stakeholders may be quite different. Given that this population is younger, and likely to have young children, potential stakeholders may include pediatricians, WIC offices, and family planning clinics. Health literacy interventions may be similar to those described above, particularly with respect to pharmacy and medication interventions and ensuring health materials are written at a low grade level, but may be tailored differently to reflect the culture and demographics of this community. Outreach workers to facilitate enrollment in children’s health insurance programs may also be warranted.

The maps described in this report provide an important step towards advancing action on health literacy in the state of Missouri. However, to be maximally effective, they must be used in concert with an understanding of the demographic distribution of the area and how low health literacy may affect their health and quality of care.
CHAPTER 5. CONCLUSIONS

Maps of health literacy in the state of Missouri indicated that communities residing in parts of St. Louis, Kansas City, and selected parts of the bootheel region had the lowest health literacy, on average.

Limitations

Several limitations to our findings should be kept in mind. First, we have not validated these estimates with population level screening. Second, we used data from the 2000 census; in areas with large demographic shifts, they may not accurately reflect the current population within each census tract particularly in regions that have experienced an increase in their Hispanic population such as the southwest corner of the state, parts of Kansas City, and north St. Louis. Given that Hispanics have lower health literacy, on average, census tracts within these regions may have lower health literacy than predicted. In addition, in using the 2007 ACS data, we were limited to how the census aggregated variables. For example, tract level poverty measures were only available as above or below the federal poverty limit. The census plans to release more detailed tract estimates based on the ACS next year, at which point it will be possible to further refine these maps.

An additional consideration in applying the predictive model to secondary data sources such as the ACS and Census data are that the questions used to collect information on a given construct vary slightly between the NAAL and the Census. While many of the characteristics included in the predictive models were selected in part because they correspond quite well across datasets (i.e., age, gender, race, marital status, years in United States and education), the construct of rurality, here defined as MSA/non-MSA was more of a challenge as there is not a corresponding MSA variable in the ACS or Census for the census tract or PUMA levels. Using other measures of rurality such as an urban/suburban/rural classification was not possible as they are not available in the NAAL which was the basis of this model.

Implications for Advancing Action on Health Literacy and Next Steps

The predictive model and health literacy maps advance action on health literacy in several ways. First, the maps provide an easy to use resource to identify areas of low health literacy within the state, where interventions may be effectively targeted in areas in greatest need. Second, the maps may help foster stakeholder collaboration by facilitating dialogue around health literacy. The mapping tools can be used to help potential partners identify which stakeholders have an interest in a given community, and may strengthen the business case for implementing interventions as they may be targeted more effectively. Third, the maps are easily understandable by community groups and can be used by stakeholders to promote buy-in from community members.

An additional benefit of these maps is that individual and community-level data related to quality of care, preventive health practices, and access to care can be added. By mapping health literacy and risk of low quality care simultaneously and examining the overlap (or not) of hotspots, key stakeholders
can utilize rich detailed data to obtain a more empirically based understanding of the geographic
distribution of these problems in and around their communities, and select appropriate interventions that
are targeted effectively.

Reducing gaps in quality and improving outcomes of care will require researchers, clinicians,
public health workers and policy makers to address multiple challenges related to low health literacy.
Maximizing the impact of limited resources available to address those challenges, however, requires
effectively targeting those resources towards individuals and communities at greatest need. The
predictive models and maps presented in this report provide a potential tool to assist stakeholders in the
efficient identification of communities with low health literacy and advance action on health literacy in the
state of Missouri.
REFERENCES


25. Adults with Basic and Below Basic Literacy Levels: Findings from NAAL and Implications for Practice. Adults with Basic and Below Basic Literacy Levels: Findings from NAAL and Implications for Practice; 2006; webcast.


APPENDIX A. WHAT DATA SOURCES WERE USED?

Three sources of data were used to generate estimates of health literacy at the PUMA and census tract level. The National Assessment of Adult Literacy (NAAL) was used to create the predictive model, and data from the American Community Survey (ACS) and the Census 2000 were used to provide information about the demographic characteristics of Public Use Microdata Areas (PUMA) and census tracts respectively. Each is described in more detail below.

Health Literacy Data from the National Assessment of Adult Literacy (NAAL)

The 2003 NAAL household sample is an in-person assessment of English language literacy among a nationally representative sample of US adults (age 18 and over) funded by the U.S. Dept of Education.\textsuperscript{3} The primary goal of the NAAL was to measure various components of literacy by assessing the extent to which individuals could understand and use written materials encountered in everyday activities (e.g., bus schedule, newspaper editorial). In addition to assessing general literacy, the NAAL included a component specific to health literacy, assessing the ability to effectively use health-related materials (e.g., medication label, written directions from doctor, consent form), and was the first large-scale assessment of health literacy in the United States. In addition to participating in the national household sample, six states chose to be over-sampled so that they could obtain precise estimates of health literacy in their state: Kentucky, Maryland, Massachusetts, Missouri, New York and Oklahoma.

The target population for the national and state household samples included adults 16 years or older who lived in a household at the time of interview. The final NAAL household sample consisted of 18,541 respondents, with an overall weighted response rate of 60.1 percent.\textsuperscript{3} We included individuals in our analyses if they were 18 years of age or older, and answered enough items to compute a health literacy score\textsuperscript{1} (n=17,466).

Like many other assessments, The NAAL had to contend with the paradoxical requirement that the test include a broad range of items to create a valid measure of literacy yet short enough to ensure that the test can be completed in a reasonable length of time. NAAL fulfilled both of these competing objectives by using an incomplete-block test design. Under this design, test items are grouped into large number of “blocks,” and only a small number of the blocks appeared in each test booklet so that any one individual only responds to a subset of questions asked across the full population. Data from the different booklets are tied to a common proficiency scale through statistical methods based on item response theory (IRT).\textsuperscript{23,24} More information about the NAAL, its sampling, and scoring procedures can be found in the 2006 National Center for Education Statistics report.\textsuperscript{3}

\textsuperscript{1} Individuals who did not speak English or who had a mental disability or other such barrier were not included in the NAAL sample.
American Community Survey (ACS)

The American Community Survey (ACS), conducted by the U.S. Census Bureau, was initiated in 2005 to phase out the decennial census long form survey and uses a series of monthly samples to produce annually updated data. We used ACS Public Use Microdata (PUMS) from 2007 to produce estimates of health literacy at the PUMA level. A PUMA has a population of approximately 100,000 and is the smallest sub-state geographic area currently available with individual demographic data in the ACS. Five years of ACS samples are required to produce statistics at the census tract level; thus, we cannot produce sub-PUMA estimates at this time using ACS data.

2000 Census

To enable the prediction of health literacy at the census tract level, we used aggregate demographic data from the 2000 Census Summary File 3 (SF3). Census tracts are small, relatively permanent statistical subdivisions of a county, with approximately 2,500 to 8,000 people residing in each tract. This is the smallest area for which reliable demographic counts were available.

One challenge with using 2000 census data to predict health literacy at the census tract level, however, is that the data are somewhat dated. Given the demographic shift in Missouri, particularly among the Hispanic Population in the southwest corner of the state, in north St. Louis and in Kansas City, where there had been a 1 – 6% increase (see Fig. A1; red shading), we opted to include in this report estimates of health literacy both at the census tract level, using older data from the Census SF3 tables, and at the PUMA level, which is a larger geographic area, but utilizes recent demographic data.

Figure A1. Percent Change in Hispanic Population 2000-2006

![Map showing percent change in Hispanic population from 2000 to 2006. The map highlights Missouri with a focus on the southwest corner, showing a 1-6% increase in red shading.](image-url)
APPENDIX B. HOW WERE STUDY VARIABLES DEFINED?

Health Literacy

Unlike school-based definitions of literacy, which focus on learning to read or reading to learn, adult literacy is focused on the ability of adults to use everyday materials (e.g., newspapers, schedules, forms) to accomplish specific tasks. Health literacy is a component of literacy, focusing on the use of materials to accomplish health-related tasks.

Health literacy is defined as the “degree to which individuals have the capacity to obtain, process, and understand basic health information and services needed to make appropriate health decisions.” The NAAL measured health literacy on a scale ranging from 0 to 500 points, with a national mean of 245 points. In addition to the continuous scores, four performance levels were created by the National Research Council to reflect an individual’s ability to successfully complete tasks of a given difficulty: below basic, basic, intermediate and proficient. About 14% of the US population has below basic health literacy skills (scores 0-184). Health-related tasks at this difficulty level include circling the date on an appointment slip in response to a question. About 22% of individuals have basic health literacy skills necessary to perform simple and everyday literacy activities, such as giving two reasons why a person should be tested for a specific disease, based on information in a clearly written pamphlet (score 185-225). 53% of the population has intermediate health literacy skills. These individuals can perform moderately challenging health literacy activities including determining what time a person can take a prescription medication, based on information on the prescription drug label (scores 226-309). Only 12% of the population has proficient health literacy scores with skills necessary to perform more complex and challenging literacy activities including finding information needed to define a medical term using a complex document (310-500).

Socio-demographic characteristics.

In order to maximize the utility of the model to stakeholders hoping to advance action on health literacy, we restricted the model to socio-demographic characteristics available in the 2000 census and American Community Survey (ACS). These variables included age, race/ethnicity, education, gender, marital status, poverty status, metropolitan statistical area, language spoken in home, and years residing in the United States.

Age was grouped into six categories (18-24, 25-39, 40-49, 50-64, 65-74, and 75 and older). Race/ethnicity was categorized as non-Hispanic white, non-Hispanic black, Hispanic, Asian/Pacific Islander, American Indian, and multiracial, and five categories of educational attainment were included (still in high school, less than high school, high school graduate or equivalent (GED), some college, and bachelors degree or higher). Marital status was grouped into three categories: married/living as married, never married, and widowed/divorced/separated. Given that the NAAL, ACS, and 2000 census collected data at different years, and to maximize the utility of the model in years to come, we represented income
as a percent of the federal poverty limit (FPL) (<100% of FPL, 100% of FPL or higher). Metropolitan statistical area was categorized as MSA/non-MSA. Estimates of MSA status for PUMAs and census tracts were based on the MSA status of the block groups within each geographic region: if the majority of block groups within a geographic region were classified as MSA, that geographic region was classified as MSA. Language spoken in home was classified as English or not English and length of residence in the United States was classified as born in the U.S., in U.S. for 1-5 years, in U.S. for 6-10 years, and in U.S. for more than 10 years.
Appendix C: How Were the Models and Maps Created?

Methods to Produce Summary Statistics using NAAL Data

To ensure that the NAAL could be completed in a reasonable length of time, each individual was given a subset of NAAL items. NAAL staff then generated five imputed values of each person’s health literacy proficiency (called ‘plausible values’) using methodology utilized by the National Assessment of Educational Progress (NAEP).26-29 These values represent best estimates of the health literacy score an individual would have been expected to attain had they taken all, rather than a subset, of items on the test. We used Rubin’s combining rules30 to produce summary inferences (e.g., percentages, means, regression coefficients) by taking advantage of all five plausible value scores for each person’s health literacy proficiency. Study results (e.g., regression coefficients, means) were obtained by conducting each complete-data statistical analysis five times, one with each set of imputed values, and averaging the five sets of results. Each statistical analysis included probability weights. The total standard error used in confidence interval calculations is a weighted sum of the within-imputation and between-imputation variances.

Analysis Procedures

Two predictive models of health literacy were developed. The first was a linear model that predicted the mean (or average) health literacy score. The second was a linear probability model that predicted the percent of the population scoring at the ‘above basic’ level of health literacy. Of note, the latter model is the opposite of predicting basic or below basic health literacy. We present findings for the model predicting above basic health literacy, however, to maintain consistency in the direction of model coefficients across the two predictive models.

Once the regression models were developed, we used information about the demographic makeup of each PUMA or census tract (e.g., 47% female, 53% male), obtained from the ACS or 2000 census, to generate an estimate of health literacy for that geographic area. Given that the ACS utilized individual level data, probability weights were used to compute weighted summary statistics representative of the population in each PUMA. Maps were then generated using the ESRI ArcMap software.

The standard error of the health literacy estimates was comprised of three parts. First, we addressed the sampling error in the NAAL by using the NAAL replicate weights into account when producing the within-imputation variance. Second, we accounted for uncertainty in individual scores arising from the use of plausible, or imputed, values by including the between-imputation variance in the standard error formulation. Finally, we addressed the sampling and non-sampling error in the secondary source of data (e.g., ACS, 2000 census) through the use of design effect adjustments in the Census SF3 files and replicate weights in the ACS.
APPENDIX D. PREDICTIVE MODELS OF HEALTH LITERACY

Table D1 presents the results of the models predicting mean health literacy scores (column 2) and the percent with 'above basic' health literacy (column 3). All variables, with the exception of living in a Metropolitan Statistical Area (MSA) and language spoken at home, contributed significantly to the model. As expected, educational attainment was strongly associated with health literacy. Compared to those with a college degree, individuals with less than a high school diploma had a health literacy score that was, on average, 43 points lower, and they were about 27% less likely to score in the intermediate or proficient range of health literacy. Health literacy was also lower, on average, for older adults, Blacks and Hispanics (compared to whites), individuals living below the federal poverty line, individuals who were not married or living with a partner (compared to those who were married), and those who were recent immigrants to the United States.

The adjusted $R^2$ of the models suggests that they account for about 30% and 21% of the variance in health literacy scores respectively, compared to 15% and 9.5% explained by education alone, the single best predictor of health literacy in the NAAL. Further, using education as a proxy for health literacy underestimates the health literacy of those with the lowest education. Compared to those with a college degree or more, estimates of health literacy scores comparing individuals with some college differ by less than one point between a model with only educational attainment ($\beta = -14.6$; Table D2) and the health literacy model described in Table D1 below ($\beta = -13.8$). However, for individuals with less than a high school diploma, a model with only educational attainment estimates the health literacy of individuals to be over 20 points lower ($\beta = -64.4$) than that estimated by the health literacy model ($\beta = -43.1$). Similar patterns are seen for income, which accounts for 8% of the variance. Thus, both for estimating individual and area-level health literacy, our model predicts health literacy better than any single indicator.
## Table D1. Predictive Models of Health Literacy

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Mean Health Literacy $\beta$ (SE)</th>
<th>Percent Above Basic $\beta$ (SE)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Constant</strong></td>
<td>299.8 (2.31) **</td>
<td>1.058 (0.023) **</td>
</tr>
<tr>
<td><strong>Age</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18-24</td>
<td>Reference</td>
<td>Reference</td>
</tr>
<tr>
<td>25-39</td>
<td>-3.8 (2.00)</td>
<td>-0.051 (0.019) *</td>
</tr>
<tr>
<td>40-49</td>
<td>-13.8 (2.23) **</td>
<td>-0.110 (0.021) **</td>
</tr>
<tr>
<td>50-64</td>
<td>-19.2 (2.38) **</td>
<td>-0.175 (0.024) **</td>
</tr>
<tr>
<td>65-74</td>
<td>-41.1 (2.33) **</td>
<td>-0.330 (0.024) **</td>
</tr>
<tr>
<td>75+</td>
<td>-51.9 (2.71) **</td>
<td>-0.439 (0.029) **</td>
</tr>
<tr>
<td><strong>Education</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Still in HS</td>
<td>-35.2 (6.41) **</td>
<td>-0.185 (0.059) **</td>
</tr>
<tr>
<td>&lt;HS</td>
<td>-43.1 (2.42) **</td>
<td>-0.270 (0.017) **</td>
</tr>
<tr>
<td>HS/GED</td>
<td>-32.8 (1.58) **</td>
<td>-0.204 (0.015) **</td>
</tr>
<tr>
<td>Some College</td>
<td>-13.8 (2.19) **</td>
<td>-0.075 (0.013) **</td>
</tr>
<tr>
<td>College degree or more</td>
<td>Reference</td>
<td>Reference</td>
</tr>
<tr>
<td><strong>Sex</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>-7.4 (1.08) **</td>
<td>-0.052 (0.010) **</td>
</tr>
<tr>
<td>Female</td>
<td>Reference</td>
<td>Reference</td>
</tr>
<tr>
<td><strong>Language Spoken at Home</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>English</td>
<td>Reference</td>
<td>Reference</td>
</tr>
<tr>
<td>Other Language</td>
<td>1.8 (2.39)</td>
<td>0.007 (0.023)</td>
</tr>
<tr>
<td><strong>Marital Status</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Married, Living together</td>
<td>Reference</td>
<td>Reference</td>
</tr>
<tr>
<td>Never Married</td>
<td>-4.1 (1.54) *</td>
<td>-0.028 (0.014) *</td>
</tr>
<tr>
<td>Divorced, Separated, Widowed</td>
<td>-4.1 (1.56) *</td>
<td>-0.020 (0.016)</td>
</tr>
<tr>
<td><strong>MSA</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>MSA</td>
<td>Reference</td>
<td>Reference</td>
</tr>
<tr>
<td>non-MSA</td>
<td>-2.4 (1.68)</td>
<td>-0.024 (0.016)</td>
</tr>
<tr>
<td><strong>Income</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No Income Data</td>
<td>-3.0 (1.83)</td>
<td>-0.025 (0.015)</td>
</tr>
<tr>
<td>&lt;100% FPL</td>
<td>-8.7 (1.58) **</td>
<td>-0.074 (0.016) **</td>
</tr>
<tr>
<td>≥100% FPL</td>
<td>Reference</td>
<td>Reference</td>
</tr>
<tr>
<td><strong>Race / Ethnicity</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>Reference</td>
<td>Reference</td>
</tr>
<tr>
<td>Black</td>
<td>-33.4 (1.96) **</td>
<td>-0.270 (0.016) **</td>
</tr>
<tr>
<td>Hispanic</td>
<td>-33.7 (2.46) **</td>
<td>-0.253 (0.032) **</td>
</tr>
<tr>
<td>Asian/Pacific Islander</td>
<td>-18.7 (5.12) **</td>
<td>-0.132 (0.053) *</td>
</tr>
<tr>
<td>American Indian</td>
<td>-22.3 (10.27) *</td>
<td>-0.192 (0.114)</td>
</tr>
<tr>
<td>Multiracial</td>
<td>-21.9 (5.12) **</td>
<td>-0.175 (0.045) **</td>
</tr>
<tr>
<td><strong>Time in United States</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;5 years</td>
<td>-21.0 (4.05) **</td>
<td>-0.158 (0.045) **</td>
</tr>
<tr>
<td>6-10 years</td>
<td>-19.5 (4.47) **</td>
<td>-0.137 (0.035) **</td>
</tr>
<tr>
<td>&gt;10 years</td>
<td>-14.5 (2.25) **</td>
<td>-0.108 (0.026) **</td>
</tr>
<tr>
<td>Born in U.S.</td>
<td>Reference</td>
<td>Reference</td>
</tr>
</tbody>
</table>
Table D2. Educational Attainment Predicting Mean Health Literacy

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Mean Health Literacy β (SE)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Education</strong></td>
<td></td>
</tr>
<tr>
<td>Still in HS</td>
<td>-47.8 (9.45)**</td>
</tr>
<tr>
<td>&lt;HS</td>
<td>-64.4 (2.51)**</td>
</tr>
<tr>
<td>HS/GED</td>
<td>-38.2 (1.73)**</td>
</tr>
<tr>
<td>Some College</td>
<td>-14.6 (2.27)**</td>
</tr>
<tr>
<td>College degree or more</td>
<td>Reference</td>
</tr>
</tbody>
</table>
APPENDIX E: WHAT IS THE DISTRIBUTION OF HEALTH LITERACY WITHIN MISSOURI?

Figures 3-8 provide a geographical representation of health literacy within the state of Missouri. Census tracts and PUMAs were shaded according to the mean health literacy or percent of the population above basic, but they do not display the distribution of health literacy within the geographic regions of interest. Figures E1 to E4 below, provide a graphical representation of the standard errors around the estimates and the distribution of health literacy scores overall within the state of Missouri and within PUMAs.

Figure E1, for example, presents the percent of the population scoring above basic with corresponding standard error bars for each PUMA. Figure E2 illustrates this in a different manner by plotting the percent of individuals scoring ‘above basic’ within each census tract, grouped by PUMA. As can be seen from Figure E2, PUMAs in the St. Louis region (01803 — 01601; red border) and in the Kansas City region (01100 – 01001; blue border) have a greater range of health literacy, suggesting that more targeted interventions may be necessary in those areas, compared to many rural areas where the level of health literacy is more consistent. The columns, ranging from red to green, correspond to the respective map coloring in Figures 3-8. For example, the red shaded regions in Figures E3 and E4 represent communities where less than half the population has a health literacy score at the ‘above basic’ level.
Figure E1. Mean Health Literacy Score and Standard Errors by PUMA

- St. Louis
- Kansas City

PUMA map with health literacy scores and frequency distribution. The scores are divided into two categories: Basic (185-225) and Intermediate (226-309).
Figure E2. Distribution of Mean Health Literacy Scores within PUMAs

- **St. Louis**
- **Kansas City**

### Distribution of Mean Health Literacy Scores

- **Basic**: 185-225
- **Intermediate**: 226-309

**PUMA Boundaries**

- St. Louis
- Kansas City

**Frequency**

- 0
- 20
- 40
- 60
- 80
- 100

**Mean Health Literacy Score**

- 200
- 215
- 230
- 245
- 260
- 275
Figure E3. Percent Scoring ‘Above Basic’ and Standard Errors by PUMA

[Graph showing data for St. Louis and Kansas City]
Figure E4. Distribution of Percent of Census Tract Scoring “Above Basic” within PUMA

- St. Louis
- Kansas City
Appendix F. Demographic Maps of the Bootheel Region of Missouri

The figures below provide closer views of the distribution of education less than HS (Figure F1), income at or below the federal poverty level (Figure F2), and age 65 or older in the bootheel region of Missouri (Figure F3).

Figure F1. Percent without a HS diploma

Figure F2. Percent with income ≤ FPL

Figure F3. Percent 65 years and over