NOSORH MANUAL

Using Data to Define Rural Populations

Introduction

The State Offices of Rural Health (SORH) focus on three core activities: the collection and dissemination of rural health information, coordination of rural health activities within their state, and the provision of technical assistance to rural public and non-profit entities. The work of the SORH is founded upon understanding the needs of rural communities. SORH are called upon to help rural communities identify and understand rural community demographics, health needs and other facets of the community that influence health. SORH who are able to articulate the breadth of rural community characteristics are positioned to successfully accomplish the three core activities.

The 5-year grant cycle requires SORH to conduct an assessment of, and report on, their rural populations. Understanding the data points that characterize rural communities ensures SORH have a solid base for providing technical assistance and helping rural communities. With a baseline of rich community descriptions, SORH can support rural community efforts to seek grants, conduct health needs assessments, plan environmental scans, and strategize community development initiatives with data that supports transparent decision-making and effective allocation of resources.

The best descriptions of rural populations are more than population counts, percent male or female, or even basic race and ethnicity data. They include social, environmental, and economic factors that influence health outcomes and effective planning for health services. Capturing the breadth of information to accurately define rural communities requires an ongoing effort which can seem overwhelming, especially to small staff SORH.

Many SORH do not have access to an epidemiologist for assistance in collecting rural-relevant information. Even SORH who have the advanced skills necessary to develop a comprehensive dataset often struggle with where to focus their initial efforts. This document outlines a simple compilation method and provides reputable sources for state and county-level data. It should be noted that more advanced SORH may opt to find the same, or similar, data from direct sources that may be more timely or relevant to their state.

The intent of this NOSORH manual is to provide easy-to-access resources for SORH to compile rich descriptions of rural populations without a large investment in data and staffing. This document provides a step-by-step guide on how to develop a rural community dataset, links to free resources and existing datasets, and guidance on analyzing data to answer key questions about rural communities. The manual is intended to help develop the baseline data capacity of SORH, recognizing that this is only one initial resource to help SORH with basic community descriptors. SORH should anticipate additional effort will be needed to develop their data capacity.

Getting Started

This manual identifies four overarching steps that a SORH can undertake to develop an initial dataset in defining the rural populations in their state. As outlined here, any rural population dataset should include the complete demographic, social, environmental and economic factors that influence health outcomes.

The steps to defining your rural populations include:

1. Define rural for your state
2. Identify your questions
3. Compile your dataset
4. Answer your questions

1 Define rural for your state

There are a number of ways in which rurality can be defined at the state level. A growing number of states are adopting policies within their Department of Health that define *rural*; if your state has this designation it should be incorporated into the dataset.

The majority of data that you will collect and work with is at the county level, therefore NOSORH suggests SORH begin with the county-level USDA Rural-Urban Continuum Codes (RUCC) definition for rural. The RUCC method provides county-level designations that are most closely aligned with the Rural/Urban Commuting Area (RUCA) codes, which are offered at the census tract level. It should be noted that FORHP grant programs base eligibility on RUCA codes. To assist with the determination of FORHP grant program eligibility, refer to the Rural Health Information Hub's "Am I Rural?" Tool.

2 Identify your questions

Developing any dataset should start with identifying the questions that will be answered with the collected data. The development of rich community descriptions requires SORH to have an understanding of how these descriptions will be used, by what target audience, for what purpose and how often they will be updated. NOSORH recommends that SORH assign staff to plan community descriptions that relate specifically to the current SORH programs or goals of the SORH. The descriptions may be used to provide TA to rural communities or support their grant seeking, community needs assessments and other efforts. NOSORH recommends that SORH have an internal plan to refresh their dataset every other year.

For the purpose of defining the rural populations of the state, stakeholders can begin by considering the following questions:

- What are the basic demographics of rural communities?
- What is the health care and public health infrastructure in rural communities?

CONTINUED

CONTINUED FROM PREVIOUS PAGE

- What is the health status of rural communities, and how does that compare to urban areas?
- What are the rates of disease and injury in rural communities, and how does that compare to urban areas or the nation?
- What are the death rates of rural communities, and how does that compare to urban areas or the nation?
- What are the comparative challenges to health care access for rural communities?
- What social factors contribute to the health outcomes for rural communities, and how does that compare to urban areas or the nation?
- What environmental factors contribute to the health outcome for rural communities, and how does that compare to urban areas or the nation?

3 Compile your dataset

Compiling a dataset to answer the identified questions requires the identification of datasets, a preview of how the data is structured and can be extracted, and the development of an internal policy and procedure for defining the utilization of the identified data in a consistent manner over time, by multiple reporters and for a variety of end-users. Prior to compiling your complete dataset, some key points to consider include:

Measurement levels

It is important to remember that data can only be split at the level it was collected. For instance, if the data you collect on age is in a range (18-29, 30-50, etc.), it is not possible to break this down further (i.e., how many are 19 years old?). NOSORH suggests that SORH consider the most descriptive measurement level available when developing the dataset, rather than waiting until they are needed.

Political Boundaries

Depending upon the questions that you'd like to ask of your data, you may want to consider a number of different political boundaries as a way of splitting the data. These pre-defined boundaries typically represent similarities between population and geography within a state and are easily identifiable.

- Federal congressional district
- State legislative districts

FIPS Code

The Federal Information Processing Standard (FIPS) code is a county identifier that easily allows for mapping software to link data with the correct location. When developing the initial dataset, it is helpful to include the FIPS codes for easy mapping. Though it may seem cumbersome, visualizing the data can make it more actionable and comprehensible. The addition of the FIPS code is a simple inclusion that can make a major difference in the long run.

The following lists of measures are an array of data that may be helpful for developing rural community descriptions. Most of this data is available within your state's County Health Rankings (CHR) dataset from the County Health Rankings and Roadmaps program. There are some additional measures that are helpful in defining

CONTINUED

CONTINUED FROM PREVIOUS PAGE

rural populations that are not available in the CHR dataset. The measures listed below are linked to where additional state-specific data is available. If the measure is available from the CHR dataset, no link is provided. It is recommended that additional measures be combined with the identified CHR base measures to develop a comprehensive state dataset that answers the pertinent questions. If you are looking for additional data from reputable sources, a list of data sources by state is available from CHR's website.

Defining rural: What are the basic demographics of the rural communities?

- Population size, by county
- % of population rural and urban
- Population density *USDA Economic Research Service*
- % of population under 18
- % of adult population *RHIhub Rural Data Explorer*
- % of population 65+
- % of population 85+ *ACS Demographic and Housing Estimates*
- Dependency ratio
- Racial/ethnic minority population
- Gender
- Family household status (Children in single-parent households)

Rural health care: What is the health care and public health infrastructure in the rural communities?

- # of CAHs *Flex Monitoring Team*
- # of rural hospitals *UNC Sheps Center for Health Services Research*
- # of Rural Health Clinics *Centers for Medicare and Medicaid Services*
- # of FQHCs *HRSA Data Warehouse*

Health status: What is the health status of the rural communities, and how does that compare to urban areas?

- Life expectancy
- Years of potential life lost (premature death)
- Self-reported health status
- % of adults reporting fair or poor health
- Poor physical health days
- Poor mental health days
- Disability status

Morbidity rates: What are the rates of disease and injury in the rural communities, and how does that compare to urban areas or the nation?

- Sexually transmitted infections
- HIV prevalence
- Diabetes prevalence
- Violent crime

CONTINUED

CONTINUED FROM PREVIOUS PAGE

Mortality rates: What are the death rates of the rural communities, and how does that compare to urban areas or the nation?

- Injury deaths (total)
- Premature age-adjusted mortality
- Child mortality
- Infant mortality
- Drug overdose deaths

Health care access: What are the comparative challenges to health care access for the rural communities?

- Health insurance coverage
- Preventive clinical service utilization
- Routine primary care utilization — medical and dental
- Self-reported financial barriers to service utilization
- Primary care, mental health and dental shortage area designations *HRSA Data Warehouse*

Social determinants: What social factors contribute to the health outcomes of the rural communities, and how does that compare to urban areas or the nation?

- Poverty rates
- Income inequality
- Unemployment
- Language limitations
- % of public school children eligible for free or reduced-price lunch
- Level of education
- High school graduation
- Some college

Environmental determinants: What environmental factors contribute to the health outcome of the rural communities, and how does that compare to urban areas or the nation?

- Housing overcrowding (Severe housing problems)
- Housing cost burden (Severe housing cost burden)
- Air quality
- Water quality (Drinking water violations)
- Violent crime
- No-vehicle households
- Food insecurity
- Limited access to healthy food

4 Answer your questions

Interpretation of the dataset is one of the most important elements of defining rural populations to avoid misrepresentation of key messages or inappropriate use of data. This step in the development of community descriptions includes the analysis or interpretation of the data and the compilation of a report or other document to "answer the questions".

Rural data often has more extreme values compared to urban data, raising questions of small sample size and heterogeneous populations (Clawar, Thompson, & Pink, 2018). Commonly occurring mistakes in the analysis include using averages as the primary statistic and combining averages to arrive at a total. As rural values tend to be more extreme, consideration must be given to the distribution before determining whether the mean (average) or median is most appropriate. More detailed information on these common errors can be found in Range Matters: Rural Averages Can Conceal Important Information or by joining NOSORH's Rural Health Data Institute.

While there are a number of ways to analyze data and even more software platforms designed to conduct statistical analysis, the most functional option for most SORH is using the Microsoft Excel software. When downloading CHR data, the file will be available as a ".csv" file type, which can be opened or imported with many platforms, including Excel. (Note: save the document as a ".xsl" or ".xslx" file to keep formatting changes). Some current resources for creating and analyzing as dataset in Microsoft Excel can be found at:

- Princeton Data & Statistical Services
- How to Create a Database in Excel
- Analyze your data instantly (for Excel)

Once the data has been analyzed, it is important to consider how this data will be used to answer the initial questions. This should be considered in light of the anticipated target audience and method of dissemination. Simple social media graphics and one-page fact sheets have become common methods of dissemination that make information easy to digest for many target audiences. Some current resources for developing meaningful data products can be found at:

- Health DataViz's Resource Documents
- County Health Rankings and Roadmap's Action Center: Communicate
- Depict Data Studio's Ann Emery Blog
- NOSORH's Rural Health Data Institute.

Conclusion

By accessing and analyzing data, State Offices of Rural Health can develop rural community descriptions, gain an in-depth understanding of rural populations, and lay the groundwork to make data-driven decisions and program planning, which improve the health of rural communities. The creation of rural community descriptions can be a singular event or the first step in the development of a regularly maintained, robust dataset. SORH staff with a goal to build their data capacity are encouraged to participate in the NOSORH virtual Data Institute, join in data discussion groups with other SORH, or contact the NOSORH Technical Assistance Director for support. SORH should recognize that this manual is one of many resources available for building data capacity and that additional training and expertise is needed.

References

Clawar, M., Thompson, K., & Pink, G.H. (January 2018). Range matters: Rural averages can conceal important information. *NC Rural Health Research Program Findings Brief*. Retrieved from https://www.shepscenter.unc.edu/wp-content/uploads/dlm_uploads/2018/01/RuralAveragesConcealInfo.pdf

Disclaimer: *This project is supported by the Health Resources and Services Administration (HRSA) of the U.S. Department of Health and Human Services (HHS) under Cooperative Agreement #U14RH19776 State Offices of Rural Health Coordination and Development Program (SORHCDP) $750,000.*

This information or content and conclusions are those of the author and should not be construed as the official position or policy of, nor should any endorsements be inferred by HRSA, HHS or the U.S. Government