

Effectiveness of End-of-Life Strategies to Improve Health Outcomes and Reduce Disparities in Rural Appalachia: An analytic codebook

Version 2. October 12, 2022

Funding:

This publication was made possible by Grant Number: R56NR019444 (PI: Lindley) from the National Institute of Nursing Research. Its contents are solely the responsibility of the authors and do not necessarily represent the official views of the National Institute of Nursing Research or National Institutes of Health, University of Tennessee, Knoxville, or the College of Nursing.

Recommended citation:

Svynarenko, R., Huang, G., Profant, T., Lindley, L.C. (2022). Effectiveness of end-of-life strategies to improve pediatric health outcomes and reduce disparities in rural Appalachia: An analytic codebook. DOI:

Table of Contents

<i>Abstract</i>	4
<i>Introduction</i>	5
Purpose of the codebook	5
Data overview	5
Data use restrictions	5
<i>Geographic Information System Data Files</i>	6
Description of Geographic Information System Data files	9
Tutorials on Creating Maps of Appalachia Using ArcGIS Pro 3.0	11
Tutorials on Conducting travel time analysis.....	12
<i>References</i>	16
<i>Appendices</i>	17
Appendix A. Hospices Providing Care for Children in the Appalachian Region	17
Appendix B. Hospices and Counties of Residence of Children enrolled in Hospice Care (Rural vs. Urban)....	18
Appendix C. Driving Times to Hospices for Children and their Families	19

Abstract

Appalachia is one of the most medically underserved areas in the nation. The region has provider shortages and limited healthcare infrastructure. Children and adolescents in this area are in poor health and do not receive the needed quality care. Implementation of section 2302 of the 2010 Patient Protection and Affordable Care Act (ACA) enabled children enrolled in Medicaid/Children's Health Insurance Program with a terminal illness to use hospice care while continuing treatment for their terminal illness. In addition to being more comprehensive than standard hospice care, this relatively new type of care is more culturally congruent with the end-of-life values of rural Appalachian families, who often view standard hospice as hastening death. The overall goal of this project was to investigate access to pediatric concurrent hospice care in Appalachia. Our central hypothesis was that concurrent care reduces rural/urban disparities in access to hospice care. Data from the Centers for Medicare and Medicaid Services (CMS) used in this project was used and included 1,788 children who resided in the Appalachian region— from January 1, 2011, to December 31, 2013. Observations with missing birth dates, death dates, and participants older than 21 years were removed from the final sample. Geographic Information Systems (GIS) databases were created to map the boundaries of the Appalachian region, hospice locations, and driving times to them.

Keywords: hospice care, Medicaid, GIS, geographic information systems.

Introduction

Purpose of the codebook

This codebook provides overview of data that was used in the research project *Effectiveness of End-of-Life Strategies to Improve Health Outcomes and Reduce Disparities in Rural Appalachia*, sponsored by the National Institute of Nursing Research and Office of the Director, Grant Number R01NR017848 (PI: Lindley). For more information on this project please visit the web site at <https://pedeolcare.utk.edu>. Complementary publications from this project can be found at [PubMed](#) and also [here](#).

Data overview

This research project utilizes data from Medicaid Analytical Extract (MAX) database, a person-level, administrative collection of claims prepared by the Centers for Medicare and Medicaid Services (CMS) from data submitted electronically by all 50 states and the District of Columbia.^{1,2}

For this study, the following CMS Medicaid data files were used:

- MAX Personal Summary File (PS) for enrollment and demographic information of hospice patients.
- MAX Other Therapy File (OT) with data on diagnosis and procedures, coded using International Classification of Diseases (9th version ICD-9) and Current Procedural Terminology (CPT) codes.
- MAX Inpatient File (IP) that includes hospitalization data on service dates and procedure codes.
- MAX Prescription Drug File (PD) that contain information on prescription fill date and national drug codes.

Other data sources included:

- National Death Index (NDI) database maintained by the Centers for Disease Control and Prevention's National Center for Health Statistics (NCHS). It contains the date of death and cause of death for all US residents.
- Publicly available 2010 U.S. Census Bureau American Community Survey data on community income, education, and rurality.
- Publicly available CMS Hospice Provider of Services files and CMS Hospice Utilization and Payment files for information on hospice providers.
- Publicly available list of rural areas created by the Health Resources and Services Administration (HRSA) was used to identify rural communities.³ The database was merged with the MAX files using FIPS codes.

Data use restrictions

Medicaid data for this study contained protected health information (PHI) and personally identifiable information (PII) protected under the HIPAA Privacy Rule.⁴ The data files are NOT available for research activities because of the Data Use Agreement (DUA) between CMS and the research institution, the University of Tennessee, Knoxville.

Geographic Information System Data Files

Data	Folder name	File name	Feature class	Content
1. Appalachian counties with independent cities	ARC_2008 with IC.zip	Subregions-in-Appalachia_2021_Data with IC.csv	Comma-separated values file	Names and FIPS codes of 423 counties and 8 independent cities located in Appalachian region, 431 FIPS codes in total.
2. Appalachian counties' FIPS codes	ARC_2008 with IC.zip	Subregions-in-Appalachia_2021_Data with IC_fips only.csv	Comma-separated values file	FIPS codes of 431 Appalachian counties
3. Appalachian counties' polygons	ARC_2008 with IC.zip	App1.shp	Polygons	Boundaries of 431 Appalachian counties
4. Appalachian ZIP codes areas	Appzips.zip	appzips.shp	Polygons	boundaries of 4,880 postal zip codes of the Appalachian region
5. US hospices	Geocoded_Hospice_PUF.zip	Geocoded_Hospice_PUF.shp	Points	Coordinates of 813 US hospices
6. US States polygons	States_generalized.zip	States_generalized.shp	Polygons	Boundaries of 50 US States and District Columbia
7. US States lines	States_outlines.zip	States_outlines.shp	Lines	Boundaries of 50 US States and District Columbia

8.	ZIP Code Population Weighted Centroids	ZIP_Code_Population_Weighted_Centroids.zip	ZIP_Code_Population_Weighted_Centroids	Points	Data to estimate ZIP code centroids is extracted from administrative USPS address data. The population weighted centroids are based on the number of residential addresses in the component ZIP+4 (also sometimes referred to as 'ZIP9') locations for each ZIP code. The data is based on ZIP+4 centroids, not ZIP Code Tabulation Areas (ZCTAs).
9.	ZIP Code Population Weighted Centroids for Appalachian states	ZIP_Code_Population_Weighted_Centroids_Appla.zip	ZIP_Code_Population_Weighted_Centroids_Appla.shp	Points	Population -weighted centroids of 14,619 of Postal ZIP code Areas of 13 Appalachian States
10	US Cities	cities shape,zip	citiesx010g.shp	Points	38,186 US cities and towns
11	Sample of children		fake kids data_fips.csv	Comma-separated values file	Table with a simulation sample of 50 children residing in Appalachian counties

12 Toolbox for
computing
driving times.

Hospice_pro.tbx

Toolbox

Toolbox that calculates
driving distances. (see
the *Tutorial on
Conducting Travel Time
Analysis* below)

Description of Geographic Information System Data files

1. **Appalachian counties with independent cities** is a table of Appalachian counties. It was created using the official list of the 423 Appalachian counties and their corresponding Federal Information Processing System (FIPS) Codes from the Appalachian Regional Commission ([ARC](#)).⁵ Then this list was manually modified to include independent cities of the state of Virginia, which in the ArcGIS files were merged with the surrounding counties. These independent cities included: Covington city, VA (FIPS: 51580), Galax city, VA (FIPS: 51640), Martinsville city, VA (FIPS: 51690), Radford city, VA (FIPS: 51750), Buena Vista, VA (FIPS: 51530), Lexington City, (FIPS: 51678), Wise, VA (FIPS: 51195), and Norton city, VA (FIPS: 51720). The final table included 431 counties of the Appalachian region with respective FIPS codes and geographic names.
2. **Appalachian counties' FIPS codes** is a table of 431 Appalachian counties which includes only respective FIPS codes.
3. **Appalachian counties' polygons** is a shapefile with boundaries of 431 Appalachian counties. It was created by joining the table of the Appalachian counties with the layer of 3,142 US counties, developed by [ESRI](#). Non-Appalachian counties were removed from the final shapefile.
4. **Appalachian ZIP codes** is a shapefile created by spatially joining the map of the Appalachian counties with the map of US ZIP areas of the [US Census Bureau](#).
5. **US hospices** is a shapefile with coordinates of US hospices. It was created using the [Hospice Provider Utilization and Payment Public Use Files of](#) the Centers of Medicare and Medicaid Services, which included addresses of US hospices. Addresses were converted into geographic points.
6. **US States polygons** is a shape developed by [ESRI](#) with boundaries for the 50 US States and the District of Columbia and their corresponding labels with names of each state.
7. **The US States outlines** is a shape file developed by [ESRI](#) with boundaries for the 50 US States and the District of Columbia.
8. **ZIP Code Population Weighted Centroids**. Population weighted centroids are a common tool for spatial analysis, particularly when more granular data is unavailable or researchers lack sophisticated geocoding tools. The ZIP Code Population Weighted Centroids allows researchers and analysts to estimate the center of population in a given geography rather than the geometric center. To learn more about administrative USPS address data, please visit: <https://www.huduser.gov/portal/datasets/usps.html>
9. **ZIP Code Population Weighted Centroids for Appalachian states** is a shapefile with coordinates of population-weighted centroids of each zip code area. This file also includes three variables with data on the distance from the centroid to the closest hospice (Total_Mile), driving times (Total_Minu), and hospice name (Hospice).

10. **US cities** is a shapefile developed by [ESRI](#) that includes coordinates of 38,186 US cities and towns.
11. **Sample of children** is a database with simulated data of 50 children enrolled in hospice care and residing in the Appalachian region. The database includes FIPS codes of their counties of residence and a rural/urban variable (ru_hrsa2) for identifying kids in rural and urban counties.
12. Toolbox is a tool for computing driving times. See the *Tutorial on Conducting Travel Time Analysis* below.

Tutorials on Creating Maps of Appalachia Using ArcGIS Pro 3.0

The source file with all the data necessary to create maps is in the zip file “ArcGIS Enclave OD.zip”. It must be unzipped on a local computer with preinstalled ArcGIS pro. All jobs are accompanied with video tutorials and examples of outcomes. The list of all jobs altogether is also listed below.

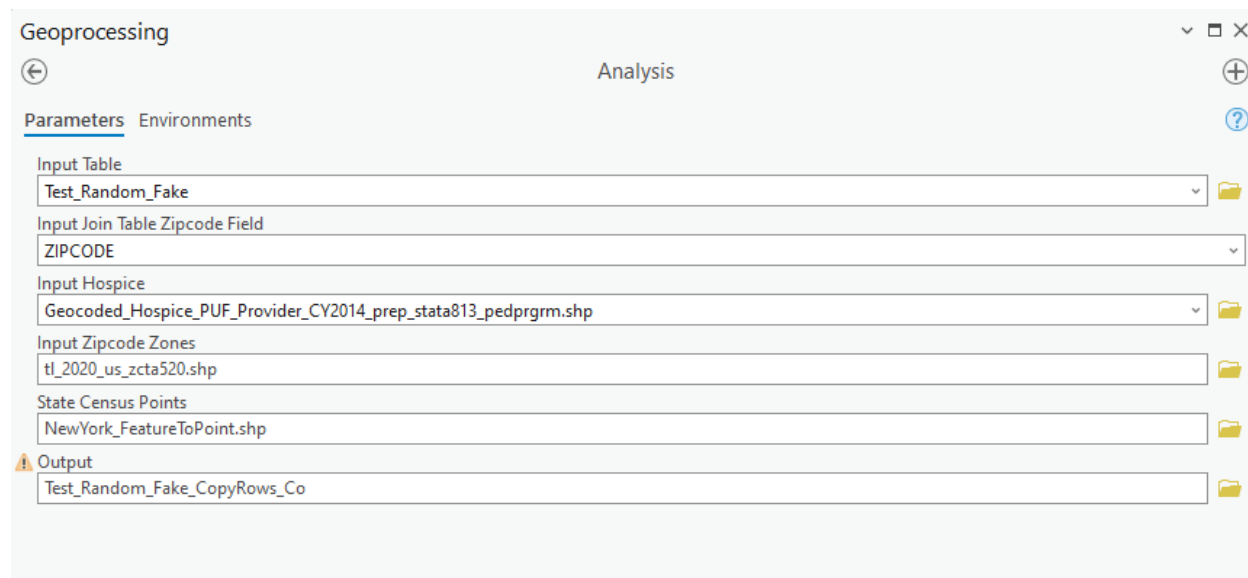
Steps	Video tutorials
1. Creating a map of Appalachian region.	
1. Create a new project on a local drive. Using the Catalog Pane, add Folder Connection to ArcGIS Enclave OD. Then drag and drop shapefiles with Polygons of Appalachian counties (ARC_2008 with IC), Points of US Hospices (Geocoded_Hospice_PUF), Polygons of US States boundaries (States_generalized), Lines of US States boundaries (States_outlines).	create a project (4:29)
2. Create a layout and save it as “jobN1.jpg” (See Appendix A).	export layout (3:03)
2. Adding simulation data to map of Appalachian region.	
1. Import file with a simulation sample of pediatric hospice patients (fake kids data_fips.csv). Then merge CSV file with the Appalachian layer using “joins and relates” menu and FIPST (app1) and FIPS (fake kids data_fips.csv) variables. Both of these variables identify corresponding FIPS code of the county, and are saved as “long” data type variables	merging tables (4:09)
2. Export the final merged table as a “fakefips.csv” file.	export table (1:38)
3. Creating a Heat Map of the Appalachian region.	
1. Use symbology to highlight rural and urban counties.	Heat map (6:24)
2. Convert labels of US States into annotations and adjust them to improve map readability	Labels (2:41)
3. Remove hospices located outside the Appalachian region	Remove hospices (0:58)
4. Same as step 3 but using “spatial join”.	Spatial join (2:29)
5. Export map as a layout “jobN3.jpg” (see Appendix B)	Export heat map (1:09)
4. Creating a Heat Map of driving times in Appalachia	
1. Import ZIP_Code_Population.shp. Create a map of Appalachian postal ZIP codes areas.	ZIP_areas (7:06)
2. Merge the Postal ZIP codes layer with ZIP_Code_Population.shp	Driving times (4:32)
3. Create a layout with driving times (See Appendix C)	Driving layout (5:04)

Tutorial on Conducting Travel Time Analysis

Original more advanced approach (did not take this approach due to technical limitations, see below)

This tool calculates population-weighted average travel time to the nearest hospice by zip code zone. This means that the tool first breaks each zip code zone into census blocks, and calculates the travel time from each centroid of the census block to the nearest hospice. Then by applying to each travel time result a weight which is the percentage of the population size of the block within the entire zip code zone, the tool gets average travel time for the zip code zone. In theory, this is a much better representation of accessibility. But in reality, each routing analysis consumes ArcGIS online credits when ArcGIS Online road network is used, resulting in insufficient credits to complete the entire calculation due to millions of block points are being analyzed. It is possible to use older road network dataset. However, data quality of the road network back in 2010 or 2011 is not good enough to provide reasonable usable output.

- Make sure ArcGIS pro is activated with an account that can access ArcGIS online road network. Typically a university account with site license should be OK, but not a personal account.
- Open or start a new ArcGIS Pro project.
- In Catalog View, navigate the folders and find the Hospice_pro.tbx toolbox. Double click on it and find the Analysis tool. Right click on it to open it up.



- In “Input Table”, choose patient table which is formatted use the exact schema as the Test_Random_Fake table.
- In “Input Join Table Zip code Field”, choose field “ZIP CODE” which is a field come with the previous table.
- In “Input Hospice”, choose the geocoded hospice data.
- In “Input Zip code Zones”, choose US census Zip code zone by state shapefile which can be downloaded from the US census website. Sample shapefiles are provided.

- In “State Census Points”, choose the corresponding state’s census block point data. This data can be create by using the “Feature to point” tool from census block polygon data which can be downloaded from the US census website. Sample shapefiles are provided.
- The specify the output file name and location. To save the table in a folder, remember to add “.csv” or “.txt” to specify the output format. This tool is to create a new table that includes all the original fields from the patient data, plus the population-weighted travel time field for each zip code zone in the patient data.
- Click on Run.

However, in NY alone, there are 350,000+ blocks. In the tool, it computes travel time from each block centroid to several nearby hospices. The routing analysis not only consumes a lot of ESRI ArcGIS online credits, but also reaches the limit of 5000 points per batch.

Updated new approach:

- Make sure ArcGIS pro is activated with an account that can access ArcGIS online road network. Typically a university account with site license should be OK, but not a personal account.
 - Open a map in a new ArcGIS pro project.
 - Add ZIP code Population Weighted Centroids from the Department of Housing and Urban Development(<https://hudgis-hud.opendata.arcgis.com/maps/zip-code-population-weighted-centroids>).
- There is a limit of 5,000 points per batch in routing analysis so we need to filter out state by state.
- Click on “select by attributes”, and build an expression “USPS_ZIP_1 is equal to NY”. Apply the filter to select NY centroids. Right click on the layer and use Export Data to save the selected points into a new shapefile.
 - Add geocoded hospice shapefile.
 - Go to menu, choose “Analysis”, and under Network Analysis choose closest facilities analysis, which will open Closest Facility Layer on the menu.
 - Select Facilities layer, and choose Import facilities to import hospices.
 - Select Incidence later, and choose import incidents from saved centroids .

- Click on Run. Closest travel time and travel distance is calculated in the Output Routes layer.

Geoprocessing v # x

Find Closest Facilities +

i This tool has been deprecated. x

i This tool consumes ArcGIS credits. x

Parameters **Environments** ?

Incidents
 ZIP_Code_Population_Weighted_Centroids_Appla1 v #

i The input has a filter. Records to be processed: 5,000 ↺

Facilities
 Geocoded_Pediatric_Hospices_Appalachian v #

Measurement Units
 Minutes v

Number of Facilities to Find

Cutoff

Travel Mode
 Custom

> Advanced Analysis

> Custom Travel Mode

> Barriers

▼ Output

Route Shape
 True Shape v

Populate Directions

Directions Language
 en

Directions Distance Units
 Miles v

Directions Style Name
 Network Analyst Desktop v

Save Output Network Analysis Layer

Save Route Data

Output Format
 Feature Set v

> Network Locations

- In the next step, join the output table to the Zip code centroid table, so that travel time and distance can be matched to the Zip code.
- In Geoprocessing Toolbox, choose Add Join tool.
- Specify the tool as shown below. Note that the first input is the centroids layer by state.

The screenshot shows the 'Add Join' tool interface in a geoprocessing environment. The window title is 'Geoprocessing' and the tool name is 'Add Join'. There are navigation icons (back, forward, search, close) and a help icon. The interface is divided into 'Parameters' and 'Environments' tabs. The 'Parameters' section includes:

- Input Table:** A dropdown menu showing 'ZIP_Code_Population_Weighted_Centroids_Appla3'.
- Input Join Field:** A dropdown menu showing 'FID'.
- Join Table:** A dropdown menu showing 'Output Routes'.
- Join Table Field:** A dropdown menu showing 'IncidentOID'.
- Keep All Target Features
- Index Joined Fields

At the bottom, there is a 'Validate Join' button.

- In the final step, right click on the centroids by state layer, open the attribute table and make sure travel time and distance are now added to the centroids layer by state. Then choose to export the table into a new file.

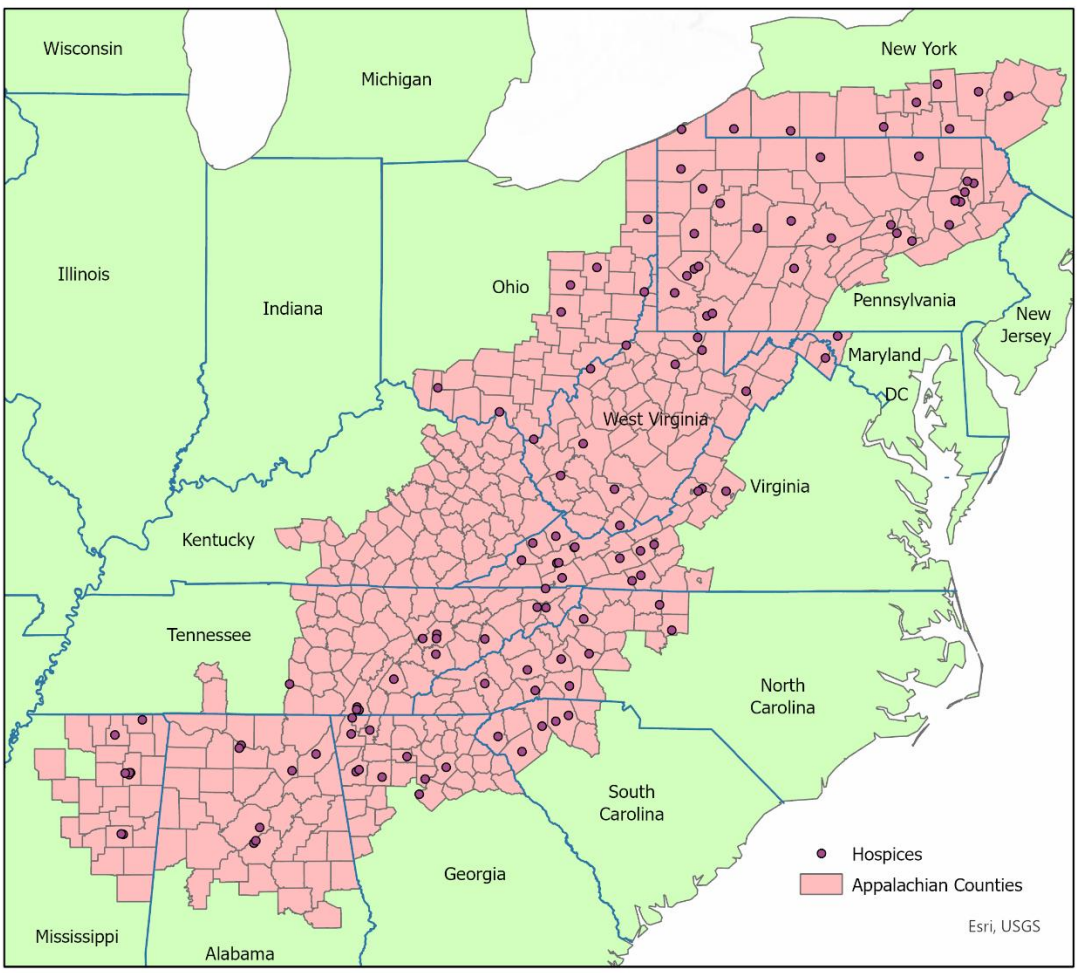
Now each zip code centroid has travel time and distance calculated in the table by state. These new tables can be used to pass travel time and distance to patient table by matching zip code.

References

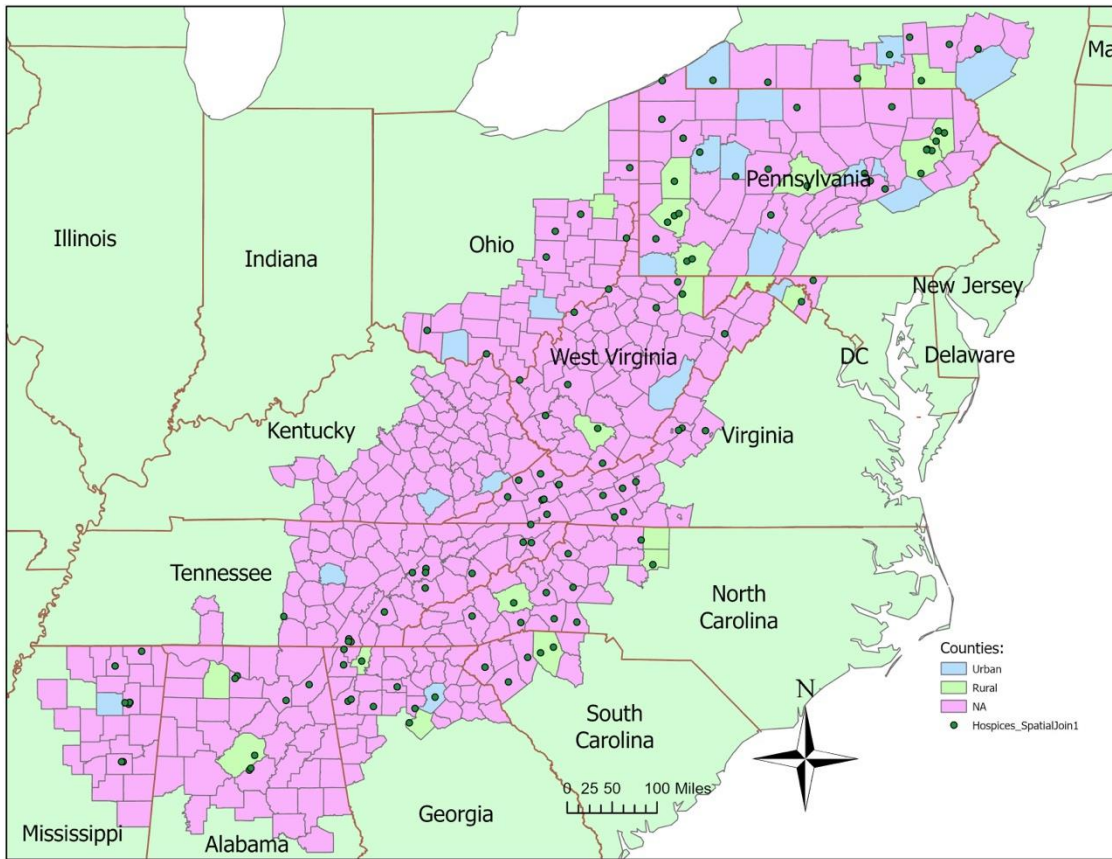
1. Ruttner L, Borck R, Nysenbaum J, Williams S. Guide to MAX data. *Mathematica Policy Research*. 2015;21.
2. Borck R, Ruttner L, Byrd V, Wagnerman K. *Medicaid Analytic Extract 2010 Chartbook*. Mathematica Policy Research. Centers for Medicare & Medicaid services; 2010. Published 2010. Accessed October 21, 2021. https://www.cms.gov/Research-Statistics-Data-and-Systems/Computer-Data-and-Systems/MedicaidDataSourcesGenInfo/MAX_Chartbooks
3. Ssynarenko R, Lindley L. Defining rurality in end-of-life research: Evaluation of common measures. *The Journal of Health Care for the Poor and Underserved*. 2021;32(4):2167-2180. doi:10.1353/hpu.2021.0189
4. U.S. Department of Health and Human Services. Summary of the HIPAA Privacy Rule. <https://www.hhs.gov/hipaa/for-professionals/privacy/laws-regulations/index.html>
5. Appalachian Regional Commission. *Appalachian Counties Served by ARC*. Available at: <https://www.arc.gov/appalachian-counties-served-by-arc/>. Accessed February 24, 2022.

Appendices

Appendix A. Hospices Providing Care for Children in the Appalachian Region



Appendix B. Hospices and Counties of Residence of Children enrolled in Hospice Care (Rural vs. Urban)



Appendix C. Driving Times to Hospices for Children and their Families

