TITLE

Self-rated health values for Oklahoma counties with limited access to health care facilities

ABSTRACT

Approximately 30% of the population of Oklahoma lives in rural areas of the state. Rural communities face many challenges in accessing adequate health care services. This limited access results in overall medical neglect and death due to undiagnosed, and untreated ailments. Self-rated health is a general measure of overall health that can be answered with “excellent, very good, good, fair, or poor.” Self-rated health is a subjective measure that can yield objective health data, but can also show cultural attitudes toward health. This study examines self-rated health values within rural Oklahoma counties to better understand regional requirements for healthcare access. This study looks at the location of healthcare facilities within counties that had a large portion of the population reporting that their health values were in the fair and poor categories. This study shows that accessibility to healthcare is more than just a spatial concern and action needs to be taken to change cultural attitudes toward healthcare in this state.

INTRODUCTION

In 2022 it was recorded that 88% of counties in Oklahoma were experiencing a primary care shortage. The limited accessibility of healthcare services for residents results in overall medical neglect and death due to undiagnosed, and untreated ailments. Rural counties in Oklahoma are faced with more challenging aspects of accessibility to healthcare. The rural counties have a small number of operational healthcare facilities in general. These facilities face their own internal challenges of maintaining a workforce across these regions. Rural residents often have to travel long distances to access the already limited facilities that are available to them. In Oklahoma we must also be mindful of the residents that are members of the tribal

nation and depend on the use of the Indian Health Services (IHS). The IHS has an even more limited number of service stations in Oklahoma. There are 20 healthcare facilities across the 77 counties that serve members of the tribal nations. The challenge of accessing proper healthcare services is not entirely a locational problem. We must also consider the socioeconomic factors that limit accessibility. The citizens must have the financial means, the ability to communicate clearly, and trust in the establishment providing services.

Self-rated health (SRH) is a general measure of overall health that can be answered with "excellent, very good, good, fair, or poor." SRH measurements require the individuals to rank their own perception of their health status. This measure is completely subjective but is a valid and reliable measurement for individuals that are not cognitively impaired. Self-rated health assessments are a fascinating tool that have shown a strong correlation with mortality. These assessments provide a wealth of objective measures of the physical health of residents, but also creates a window into the social and psychological aspects of a person and can show overall attitudes of the region.

The aim of this study is to examine self-rated health values within Oklahoma counties to better understand regional requirements for healthcare access. Examining the SRH values by region can help answer the spatial question "Is there a correlation between the accessibility (location and number) of healthcare facilities within the county and the self-rated health values?" and "Do areas with more healthcare facilities have higher self-rated health values?". In analyzing this data we can also observe correlations in mortality rate and other general health outcomes for a region.

RESEARCH BACKGROUND:

A similar study was performed in China determining the relationship between self-rated health values of individuals who live in rural areas compared to values of individuals that live in urbanized areas. While urbanization provides more access to healthcare facilities and better

opportunity to receive these services, urbanization also brings about many new health risk factors including air pollution and sedentary lifestyles for the residents (Chen et al 2017). Areas with high levels of air pollution have been shown to negatively impact the health of the Chinese people. The results of this study shows a major health gap between rural and urban residents. Despite the environmental quality of urban areas, health values were substantially higher for urban residents compared to rural residents. The major health disparities between the two groups was not so much in relation to access to health care but in access to freshwater, which is a topic not necessarily in the forefront of our minds in the United States.

There have been 106 rural hospital closures in the United States since 2010, which creates geographical barriers when trying to access healthcare services (McCarty et al 2021. This study focused on the transportation times required for community members to reach secondary healthcare facilities. The study showed that rural facility closures resulted in almost 1% of the population having to travel more than 15 minutes to the nearest secondary facility. The travel times varied across different regions of the United States but the most affected communities were located in the south central region of the United States.

Another study analyzed the self-rated health values recorded in rural Appalachia. Rural Appalachia is characterized by overall poor health, poor health practices, and health disparities. This study surveyed rural Appalachian adults on their health behaviors including: soda consumption, fast food consumption, tobacco use, activity level, blood pressure medication use and overall self-rated health status. (Griffith et al 2011). The results of this study showed a distortion between self-rated health values and actual health status. Respondents reported being overall healthy while leading sedentary lifestyles, being hypertensive, and overweight. The frequency of visits to healthcare providers for this region is also disproportionately low. This study provides a unique insight that shows how socioeconomic variables and regional attitudes can affect self-rated health values. Public health programs will have to find ways to educate

these residents to make healthier choices, when they are already convinced that they are healthy.

MATERIALS AND METHODS

In planning this study, data needed to be compiled that represented the location of healthcare facilities across the state, county boundaries as well as some demographic information about the residents of these counties. I required the data compiled to be collected post COVID-19 pandemic because I feel that attitudes towards health and demand for healthcare changed drastically during this time. I have some experience using the open science framework, which can be used as a space that can store data, preserve data, share data, and make this data collaborative. So my data will be stored and accessible through the open science framework.

The collection portion of this study was not carried out by collecting new data, but by compiling previously collected data and analyzing this data. The data compiled included an Oklahoma county boundary TIGER/Line shapefile data. Population data was compiled from the United States census bureau. The GEOID in the TIGER/Line shapefile aided in linking population data to the shapefile. The shapefile uses a projected coordinate system NAD 1983 StatePlane Oklahoma North FIPS 3501 (US Feet).

The Oklahoma State department of health has a directory of healthcare facilities that has been updated and in effect since December 2021. This data set contained latitude and longitude points of locations, facility name, city, and phone number, and was a CSV format. This CSV file was converted from X and Y coordinate points to a shapefile with the use of ArcGIS data management tool X and Y table to point. This point shapefile was also in the projected coordinate system NAD 1983 StatePlane Oklahoma North FIPS 3501 (US Feet).

The records compiled for use in this study from countyhealthrankings.org include self-rated health values, rank of health outcomes, rank of health factors, length of life rankings, quality of life rankings, covid 19 deaths, and average life expectancies. The self-rated health

scores are recorded as the percentage of the population that ranks their health fair to poor. The health rankings represent a ranking of 1-77 with number one being the highest ranked county and 77 being the lowest ranked county. I also compiled a list of the presence of county health departments from the Oklahoma state department of health and added this record as a simple “Y” or “N” to my CSV file that contains the health rankings. The CSV column headers followed consistent formatting in the style of ‘firstname_lastname”. The county boundary files listed county names as “Adair County”, so the CSV file that I created followed this same formatting.

For the assure stage in the data life cycle, I had to do some data cleanup on the healthcare facilities list. I used the OpenRefine software to filter out some results. There were several records in this dataset where there were slightly different names for the same location of a facility, and a few duplicates across the board. I also did some assurance checks to verify the measurements of the counties and distances between healthcare facilities located within. I checked distances between facilities by entering locations in google maps and was able to confirm area calculations (mostly unit conversions) of regions by quick google search.

Using the metadata wizard I was able to create metadata records for the csv files that I compiled. The healthcare facility directory had an original metadata record, however, after cleaning up the dataset to provide me with more accurate results I felt like I should create my own metadata record. There is a cross reference section of the metadata wizard where I was able to create a citation for the original dataset. I also created a metadata record using the metadata wizard for the county health ranking data compiled. This metadata record also was cross referenced and a citation was created for the original dataset. The metadata records that I created followed the current acceptable standards for metadata records and are available for others to access online.

My datasets and metadata records are preserved on the open science framework website. The metadata records provide cross referenced citations to the original datasets and maintain detailed explanations of the datasets. The data on OSF is able to be shared and

editable by other allowed users and creates a version control history which aids in the preservation of this information.

For the final stage of the data life cycle the data was analyzed, and integrated. I performed many tests in order to answer the question “Is there a spatial relationship between self-rated health values and geographical accessibility to healthcare facilities?” There were several tests performed to weed out some of the data. The first tests were conducted to determine which counties have the worst self-rated health values in the population but a limit of 10. I thought that ten would be a more manageable number of data to work with but would still be a good representation. The self-rated health values for the top ten lowest ranked counties included values ranging from 33%-26%. This measurement represents the percent of the population who ranked their health as poor-fair, which are the lowest measures on the scale. Then the count of healthcare facilities was calculated for these lowest ranked regions. I performed the same tests for counties with the lowest self-rated health percentages to help confirm my hypothesis that there is a relationship between geographical access and self-rated health values. The locations of healthcare facilities were joined and represented by points within the polygon county boundary file. The healthcare facilities were buffered to see if there was overlap between facilities, and distance was calculated for the next closest healthcare facilities for residents of a particular county. I calculated the area of these counties and the distribution of these facilities within the county. I also performed some tests to see if I could see a pattern between self-rated health values and population values in the counties and was able to transform the data and create some visual representations of these tests.

QUERIES:

Query Editor | Query History

```
1  SELECT county_nam AS County,
2  county_val AS SRH
3  FROM counties
4  ORDER  BY county_val DESC
5  Limit 10;
```

Data Output | Explain | Messages | Notifications

	county character varying (100)	srh character varying (254)
1	Adair County	33%
2	Choctaw County	29%
3	Sequoyah County	28%
4	McCurtain County	28%
5	Okfuskee County	28%
6	Hughes County	27%
7	Seminole County	26%
8	Coal County	26%
9	Delaware County	26%
10	Tillman County	26%

QUERY 1. Counties with the largest portions of the population rating health as poor or fair

Query Editor | Query History

```
1  SELECT county_nam AS County,
2  county_val AS SRH,
3  health_dep AS Health_department
4  FROM counties
5  ORDER  BY county_val ASC
6  Limit 10;
```

Data Output | Explain | Messages | Notifications

	county character varying (100)	srh character varying (254)	health_departr character varyi
1	Cleveland County	17%	Y
2	Canadian County	17%	Y
3	McClain County	18%	Y
4	Rogers County	18%	Y
5	Grant County	18%	Y
6	Woods County	18%	Y
7	Grady County	19%	Y
8	Logan County	19%	Y
9	Noble County	19%	Y
10	Major County	19%	Y

QUERY 2: Counties with the smallest portion of the population ranking health values as poor-fair.

```
1   SELECT ST_Area(geom)*0.00000009290304
2     FROM counties
3     WHERE county_nam = 'Adair County'
```

Data Output | Explain | Messages | Notifications

	?column? double precision
1	1494.437334208353

QUERY 3: Area of Adair County in sq km

```
1  SELECT ST_Area(geom)*0.00000009290304
2  from counties
3  WHERE county_nam= 'Cleveland County';
4
```

Data Output | Explain | Messages | Notifications

	?column? double precision
1	1445.6860668804172

QUERY 4: Area of Cleveland County in sq km

```
1  SELECT county_nam, COUNT(city)
2    FROM okhcf
3  JOIN counties
4    ON ST_Contains(counties.geom, okhcf.geom)
5    GROUP BY DISTINCT county_nam;
```

QUERY 5: Spatial Join of health care facilities within Oklahoma counties

```
1  CREATE TABLE hcfradius AS
2  SELECT ST_Buffer(geom,3281)::geometry(Polygon,2267) AS geom
3  FROM okhcf;
```

QUERY 6:Creating a one kilometer buffer around health care facilities

To create the following 3 visual representations (FIG 1.0-1.3) of the locations of health care facilities with a 1 km buffer the following queries were performed in this order.

```
SELECT *
FROM hcfradius
JOIN counties
ON ST_Contains(counties.geom, hcfradius.geomm)
ORDER BY county_nam ASC;
```

Step 1 a spatial join of the buffered health care facility locations to the counties. From this join a new table was created called countyhcfjoinbuffer

```
SELECT ST_BOUNDARY(geom)
FROM countyhcfjoinbuffer
WHERE county_nam ='Adair County';
```

Adair county was selected and a boundary of this county was created.

```
SELECT ST_Union(st_boundary, geomm)
FROM countyjoinbuffer;
```

Then a union was created for the county boundary with the buffered health care facilities contained within the boundary.

```
SELECT county_nam AS county, county_val AS SRH, COUNT(city), ST_Area(counties.geom) * 0.00000009290
FROM okhcf
JOIN counties
ON ST_Contains(counties.geom, okhcf.geom)
GROUP BY DISTINCT county_nam, county_val, counties.geom
ORDER BY SRH DESC;
```

QUERY 7: Area calculation of counties in sq kilometers

```
SELECT city
FROM okhcf
WHERE ST_intersects(geom, ST_geomfromtext('MULTIPOLYGON(((2999418.026340112 348226.5319143832,2999652.728183642
GROUP BY DISTINCT(city);
```

Data Output | Explain | Messages | Notifications | Geometry Viewer

	city character varying (2
1	Stilwell

QUERY 8: The only city in Adair County where there are health care facilities

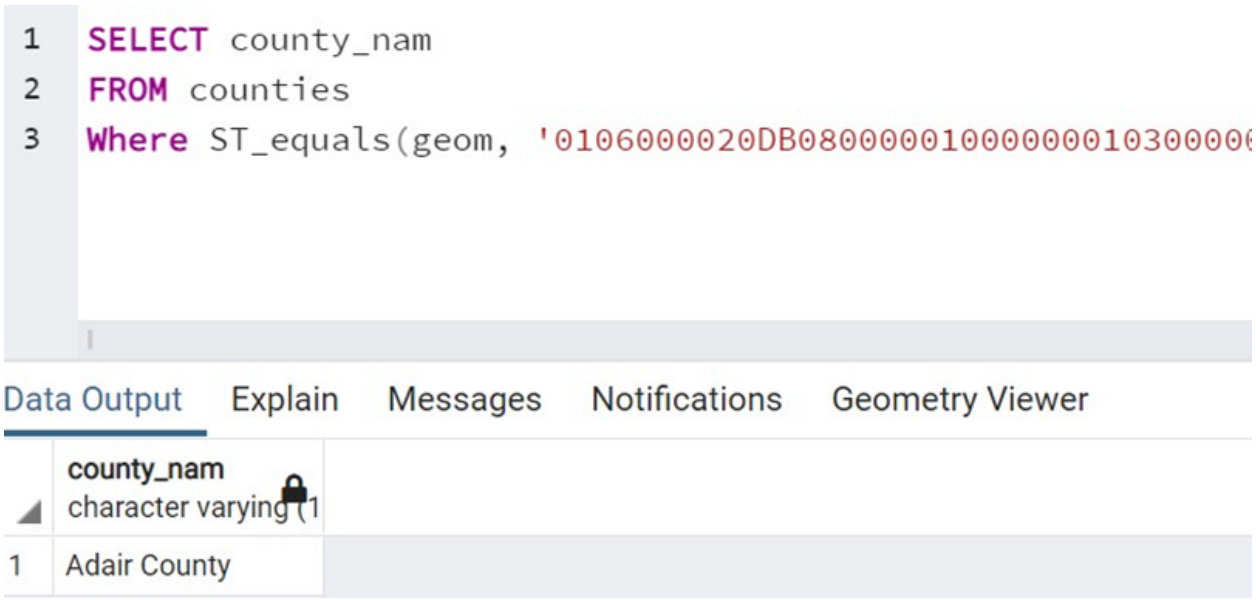
```
SELECT county_val, county_nam
FROM counties
WHERE ST_DWithin(geom,
                 ST_GeomFromText('POINT(2965030.5062470585 311633.78793489933)',2267),
                                 105600
                                 );
```

Data Output | Explain | Messages | Notifications | Geometry Viewer

	county_val character varying (2	county_nam character varying (1
1	33%	Adair County
2	28%	Sequoyah County
3	26%	Cherokee County

QUERY 9: County SRH values within a 20 miles radius of Stilwell Memorial Hospital

QUERY 10: ST_distance from Stilwell Memorial Hospital in Adair county to the next closest hospital, Tahlequah City Hospital converted to kilometers



QUERY 11: ST_Equals geometry test

```
1  SELECT ST_Overlaps(geomm, geom) As overlap,
2    ST_Crosses(geom, geomm) As crosses,
3      ST_Intersects(geom, geomm) As intersects,
4      ST_Contains(geom,geomm) As contains
5  FROM countyhcfjoinbuffer;
```

Data Output | Explain | Messages | Notifications | Geometry Viewer

	overlap boolean	crosses boolean	intersects boolean	contains boolean
1	false	false	true	true
2	false	false	true	true
3	false	false	true	true
4	false	false	true	true
5	false	false	true	true
6	false	false	true	true
7	false	false	true	true
8	false	false	true	true
9	false	false	true	true
10	false	false	true	true
11	false	false	true	true

QUERY 12: overlap, crosses, intersects, contains tests geometry tests.

```
SELECT (
    ST_Transform(
      ST_SetSRID(geom,2267),
      4326)
    )
  FROM top10worst;
```

Query 13: ST_transform to transform projected coordinates.

RESULTS

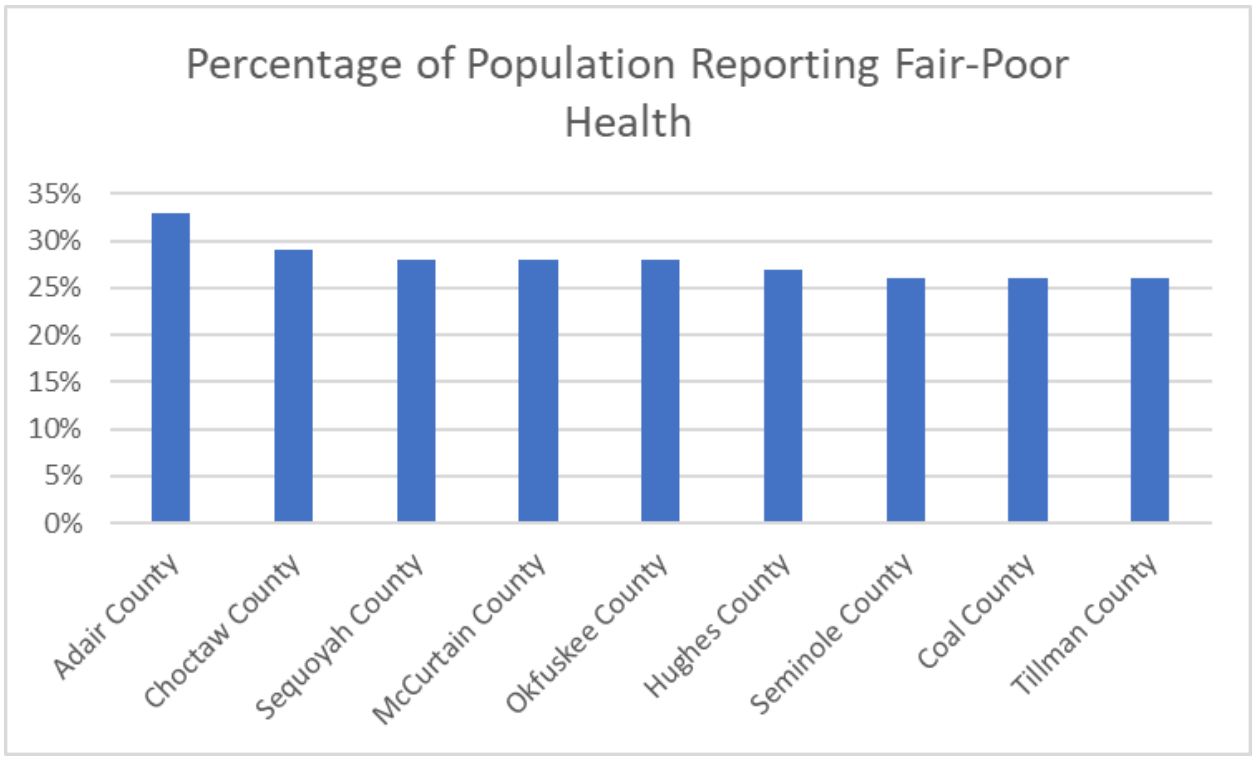


CHART 1 Percentage of the population within counties who report their health and fair-poor

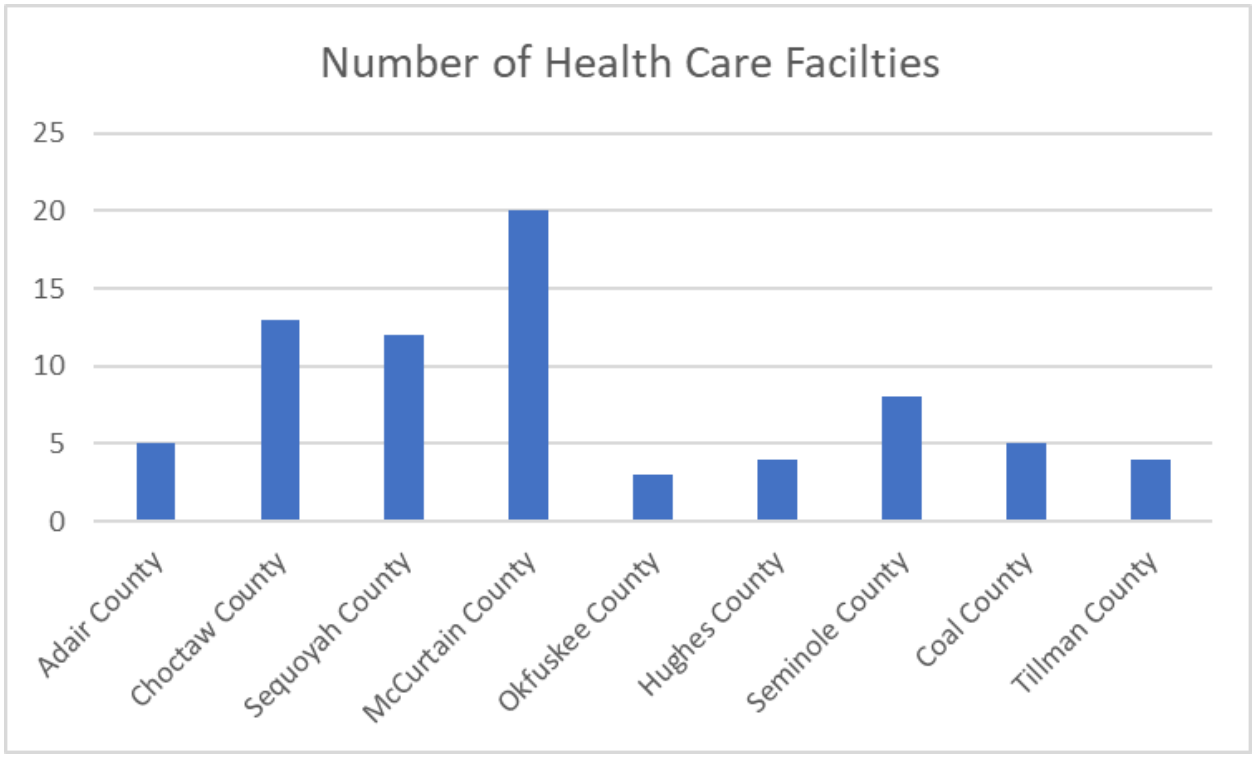


CHART 1.1 The Number of healthcare facilities within counties with highest percentage of the population reporting their health status as fair-poor

FIG 1.4 Distribution of 10 counties with the highest percentages of population reporting fair to poor health

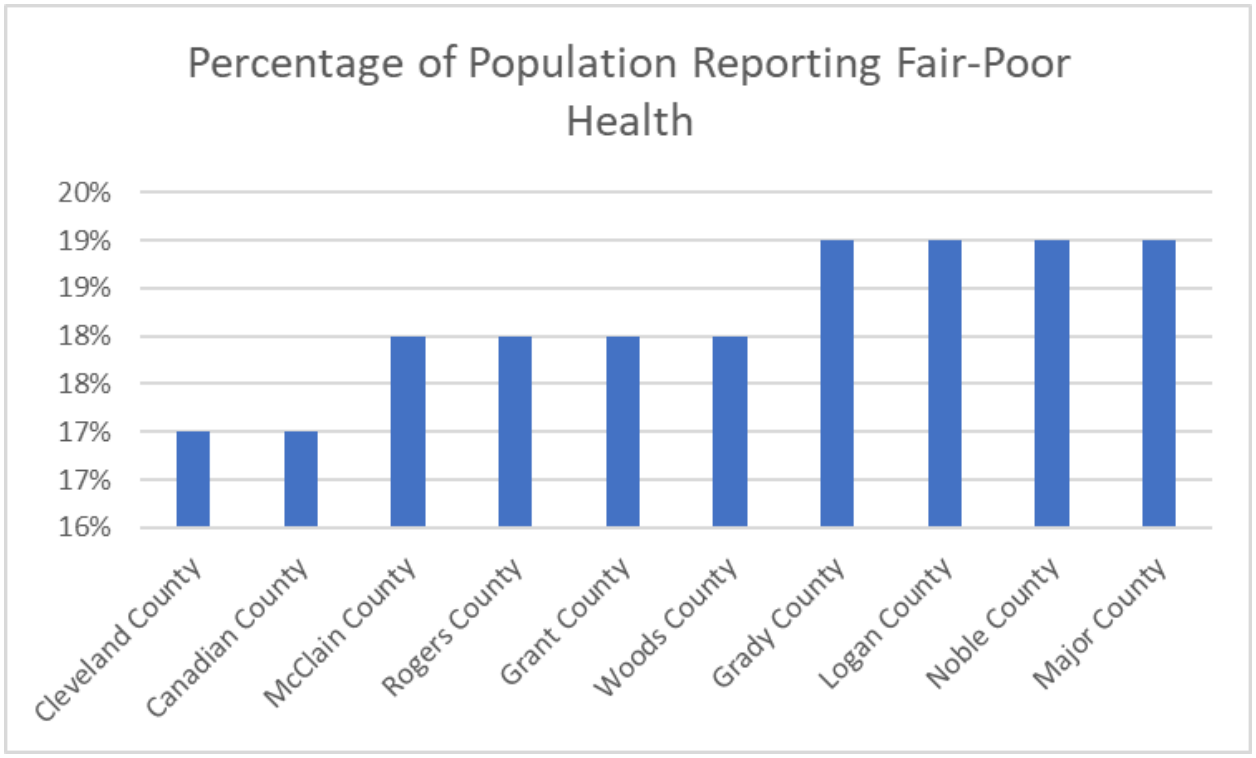


CHART 2: Percentage of the population within counties who report their health and fair-poor

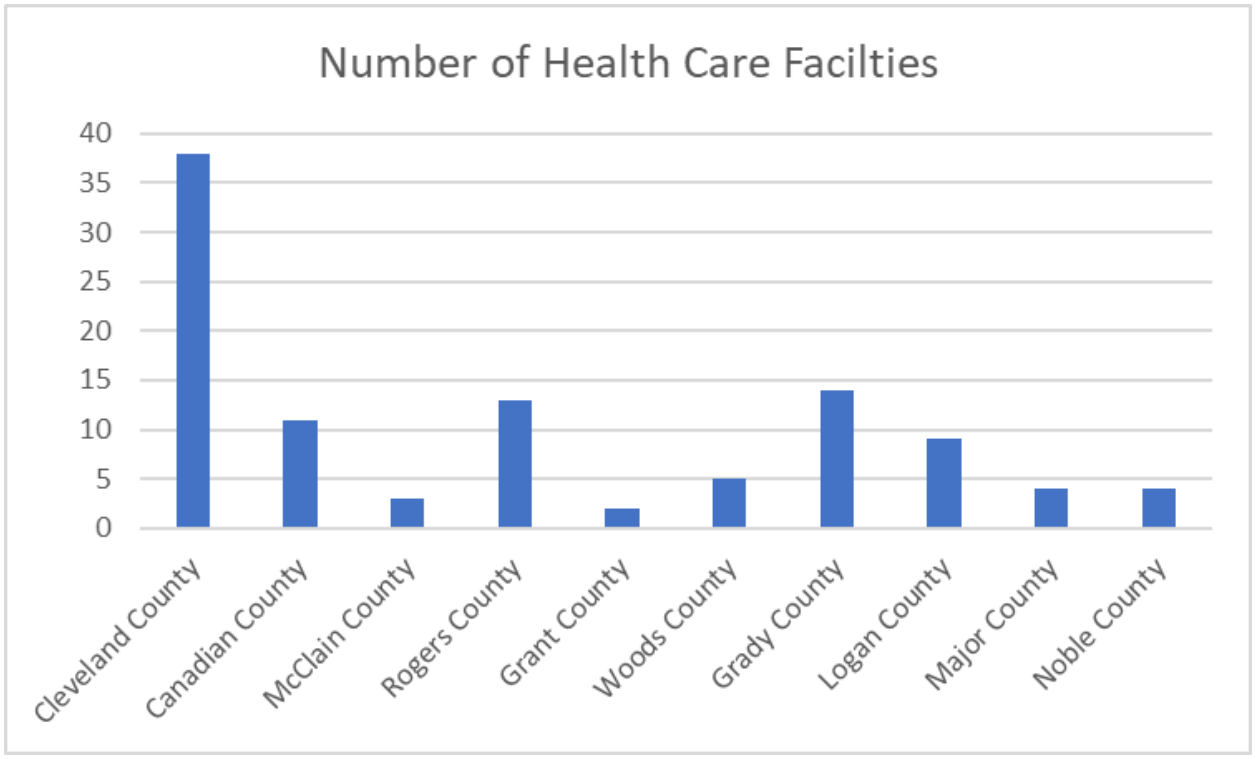


CHART 2.1 Number of healthcare facilities within counties with lowest percentage of the population reporting health values as fair-poor

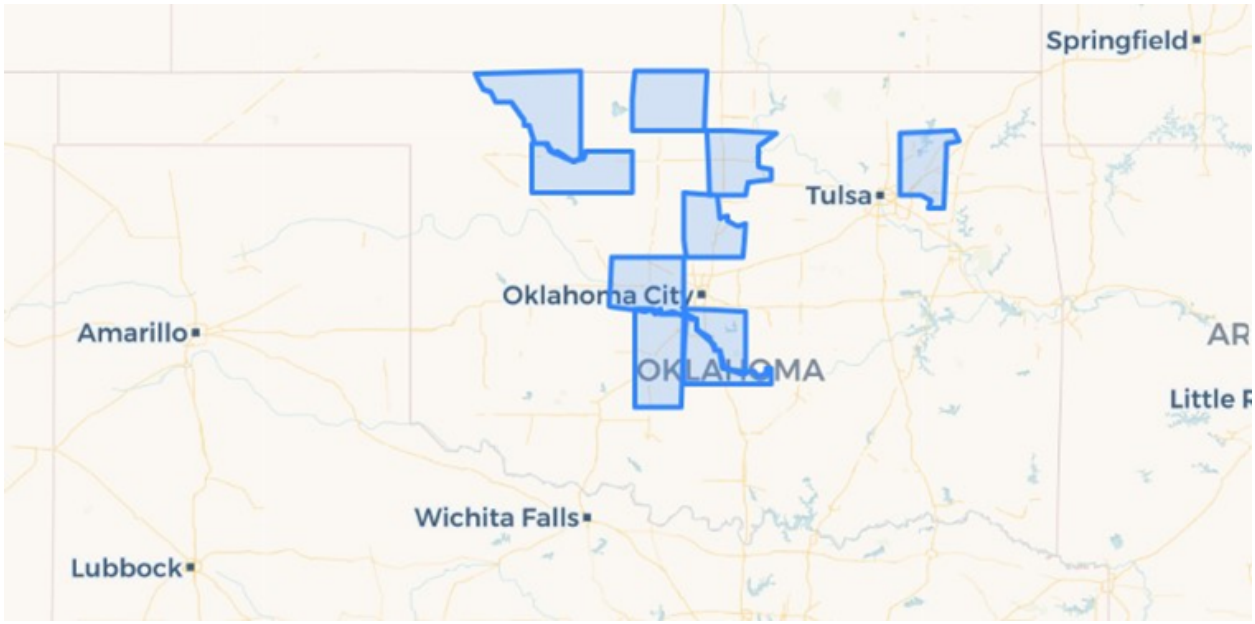


FIG 2.0 Distribution of 10 counties with the lowest percentages of population reporting fair to poor health .

The following images represent the locations of all the healthcare facilities within 3 counties with the largest portion of the population rated their health as poor-fair.

FIG. 1.0 Adair County Health Care facilities

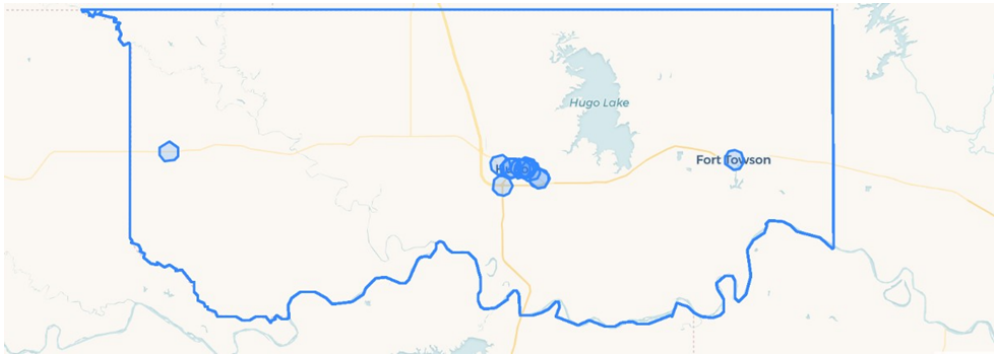


FIG 1.1 Choctaw County Health Care Facilities

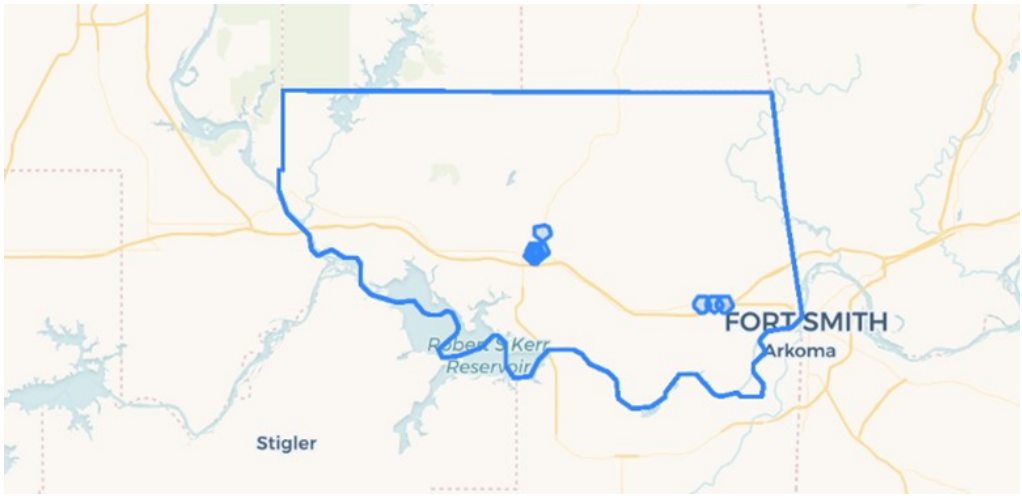


FIG 1.3 Sequoyah County Health Care Facilities.

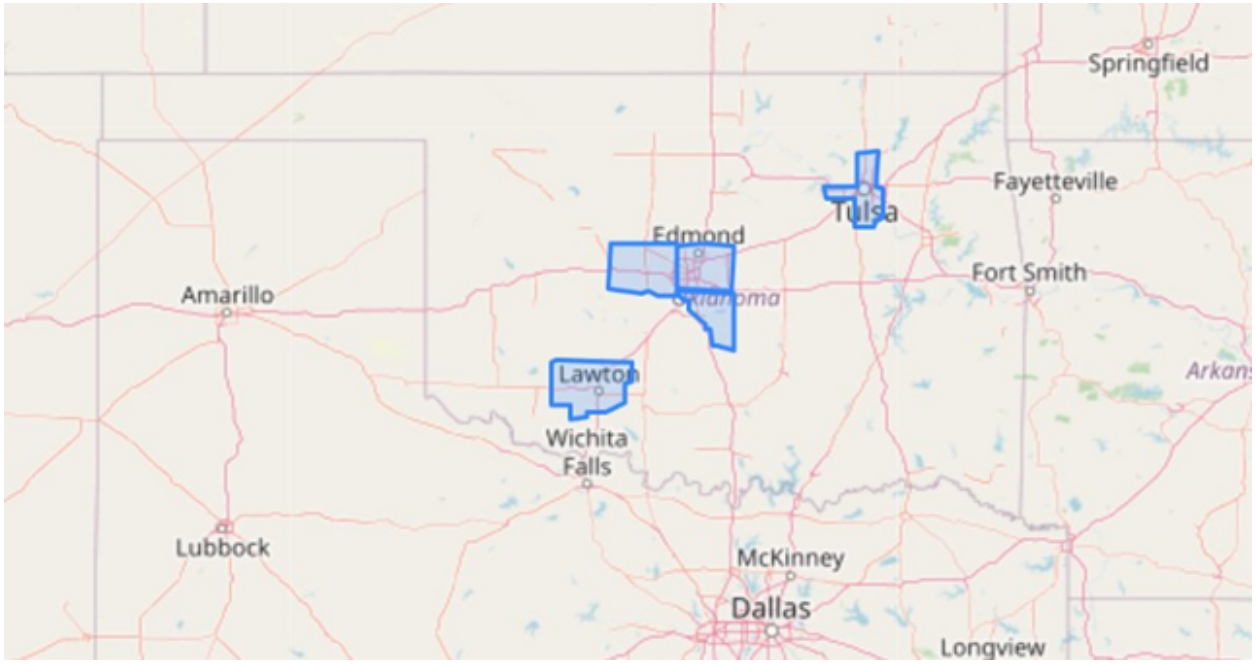


FIG 3 Union of counties with the highest population in the state.

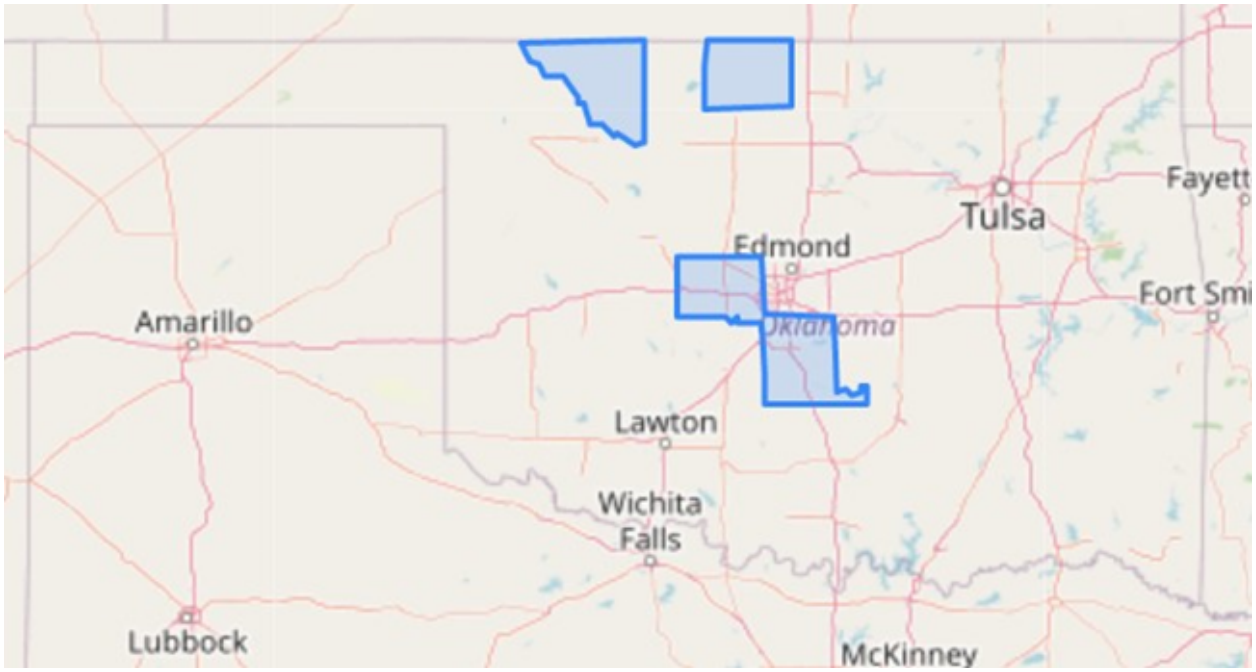


FIG 3.1 Counties with lowest percentage of the population rating health as fair-poor.

FIG 3.2 Counties with lowest population

FIG 3.3 Counties with largest percentage of the population ranking health values as fair-poor

FIgure 3-3.3 represents a test to see if there is any county overlap between the self-rated health values and the populations of the counties. There was some overlap between the counties with the lowest percentage of self-rated health values.

DISCUSSION

While implementing the data, I began to realize this spatial question has a more complex answer. Accessibility has many different meanings beyond just local availability. The tests implemented did show that there are fewer options for healthcare services in these regions that have poor self health ranking. However, there were many areas with a smaller number of facilities in the region that had higher self-rated health values. The results do not show me a definitive relationship between location of healthcare facilities within the county and the percentage of residents who report their health as fair-poor. The only sound observation that can be made from this study is that the counties with large percentages of the population reporting fair-poor health values are rural counties. Therefore, I think that the lower rating of self health stems from cultural attitudes and socioeconomic variables within these rural communities.

This study could be improved if it focused more on regions of the state and not at such a small county scale. If we continue this study on the small scale, I think it could be improved by surveying the location distribution of the residents, and any public transportation services and routes offered within the county, the percent of population who are utilizing Indian Health Services. I also think it would be interesting to see the distribution of residents within the state that are dependent on the very limited network of Indian Health services facilities.

CONCLUSION

The tests performed were overall unsuccessful in determining a relationship between the location of healthcare facilities within a county and the values of self-rated health. The counties with the highest percentage of the population rating their health poorly were all rural counties. Accessibility to healthcare is a major issue facing rural communities. However, this study showed that accessibility has more meanings than just location. Health insurance coverage, health care literacy, and social stigma and privacy are other factors that can play a part in accessibility. This study confirms the importance for creating public health programs and activities that educate members of the rural communities. These programs can educate community members of all ages, which can instill new healthy habits. Above anything else, I think this study reflects a need for universal health care in our country. A universal health care system would eliminate financial accessibility limitations, which could improve the overall health of a region.

Bibliography

Griffith, B. N., Lovett, G. D., Pyle, D. N., & Miller, W. C. (2011). Self-rated health in rural Appalachia: health perceptions are incongruent with health status and health behaviors. BMC Public Health, 11(1), 229–236. https://doi-org.ezproxy.lib.ou.edu/10.1186/1471-2458-11-229

Hongsheng Chen, Ye Liu, Zhenjun Zhu, & Zhigang Li. (2017). Does where you live matter to your health? Investigating factors that influence the self-rated health of urban and rural Chinese residents: evidence drawn from Chinese General Social Survey data. Health & Quality of Life Outcomes, 15, 1–11.

McCarthy, S., Moore, D., Smedley, W. A., Crowley, B. M., Stephens, S. W., Griffin, R. L., Tanner, L. C., & Jansen, J. O. (2021). Impact of Rural Hospital Closures on Health-Care Access. Journal of Surgical Research, 258, 170–178. https://doi-org.ezproxy.lib.ou.edu/10.1016/j.jss.2020.08.055