Using GIS To Analyze Physician Shortage Areas In Minnesota

Thomas J. Sandberg[1,2]
[1]*Department of Resource Analysis, Saint Mary's University of Minnesota, Minneapolis, MN 55410;* [2]*Blue Cross Blue Shield of Minnesota, Eagan, MN 55122*

Keywords: GIS, Healthcare, Health Plans, Demographic Analysis, Access, Medically Underserved Area, Physician Shortage, Rural Health Care, Health Professional Shortage Areas (HPSA), Physician Assistant, Advanced Nurse Practitioner, Primary Care.

Abstract

Physician shortages are a looming national problem. The current landscape of physicians in Minnesota is one that varies by region and within counties. Minnesota has several metropolitan areas that serve as bases for large provider concentrations and rural parts of the state where provider coverage is scant. The state has enough physicians to adequately serve the needs of its population; however, there is a problem of distribution. There is an assumption that paraprofessionals make up for some of the physician shortages, but the geographic extent is unknown.

Introduction

Physician shortages exist in a spatial context, and measuring them has been an inconsistent task. For the purposes of determining areas of physician shortage, the US Department of Health and Human Services has established Health Professional Shortage Areas (HPSAs). In an act of Congress, HPSAs were originally created in the late 1960's. Only minimal adjustments have occurred over the past 40 years. The earliest record of a set ratio is from legislation in 1972 that indicated a 4,000:1 ratio (US Congress, Office of Technology and Assessment, 1990). Current HPSAs were defined in 1980 and are defined as regions (states, counties, census tracts) where a 1:3500 ratio of primary care physicians to population are not met. Other non-geographic criteria exist for measuring HPSAs that include areas with large institutional populations (prisons, state hospitals), areas with unusually high population density and areas with large racial populations such as tribal areas.

Counties are among the most common geographic units for measuring, however, variation exists that makes standardization difficult. This research attempt will introduce a new standardized geographic unit to base HPSAs off of and compare these results to the county-based measurement system.

The relatively recent introduction of advanced paraprofessionals (physician assistants and advanced nurse practitioners, primarily) as primary care deliverers has an impact on HPSAs by increasing the headcount of providers. It is unknown if these paraprofessionals have an impact on the geographic distribution of shortage areas.

Background

This research focuses only on the state of Minnesota and its population. Shortage areas are a national issue, with great variation from state to state. Certain states, mostly those in the Western half of the United States, have great variation in the area, shape and population of their counties.

Many issues can contribute to an area being designated as a shortage area. This research focuses solely on the geographic aspect and does not seek to answer why any specific HPSAs exist.

One specific limitation of the federally defined, county-based system is that it does not account for physicians that traverse multiple states. Providers from Wisconsin, Iowa, North Dakota, and South Dakota that cross the Minnesota border to practice will be captured in the Minnesota provider data set from Blue Cross Blue Shield of Minnesota (BCBSM). Conversely, there is no way to capture the extent of Minnesota residents that cross into these states and receive primary care services. For the purposes of this model, it will be assumed that health care consumption occurs within the state border of Minnesota.

Methods

Data Acquisition

The federal HPSA model uses counties and census tracts as its primary geographic unit. These units of analysis are available via a data set included from the Environmental Systems Research Institute's (ESRI) ArcGIS software platform. US Population, 2000 is an attribute of the US Census Tract shapefile and is used for all references to population in this research.

The distribution of physicians and paraprofessionals are calculated using an October 2005 output from the Provider Administration and Reimbursement (PA&R) database courtesy of Blue Cross Blue Shield of Minnesota (BCBSM). HPSAs are calculated using data available internally to the US Department of Health and Human Services Bureau of Health Professions (2005).

A non-governmental, geographical unit of analysis is introduced in this research. In order to gauge physician shortages on a more homogenous level, a vector-based grid was created. This grid allows for an even spatial distribution of geographic units, unlike the irregular distribution of counties or tracts. Cell size of the grid is set at 30 by 30 miles and sets out to use standards set forth by both the US Department of Health and Human Services. US Population, 2000 from the census tract shapefile will be reassigned into grid cells.

Software Requirements and Data Manipulation

All GIS functions included in this research were performed using ArcGIS 9.1. The ET Vector Grid extension, available from www.esri.com, was used to create an arbitrary, homogenous geographic unit of analysis. To calculate shortages within individual grid cells, the general population by census tract needed to be reprojected into the grid cells. A sequence of data preparation was executed in order to estimate population.

The census tract shapefile was the source for all non-physician based statistics including area and population. This file in its normal state is projected in the GCS North American 1983 latitude/longitude projection. In order to represent Minnesota's parallel and perpendicular borders, particularly its long, straight Southern border and to avoid the East-West distortion inherent in latitude/longitude projections, the census tract shapefile was reprojected into a Mercator projection.

Once the tract shapefile was reprojected, the new grid was created using the ArcGIS extension, ETGeoWizards (figure 1). ETGeoWizards is a free extension available via the Environmental Systems Research Institute's (ESRI) website. This is a free alternative to the Spatial Analyst extension, available from ESRI for a fee. Figure 1 shows the inputs required by this extension to produce the grid. Grid extent was calculated based on the newly reprojected MN tract shapefile. Additionally, the Mercator projection uses meters as its unit of measurement; the grid cell size field is adjusted accordingly (48,240 meters is approximately 30 miles). Once the cell size, extent and file type (this example uses polygons) was established, a vector grid was created.

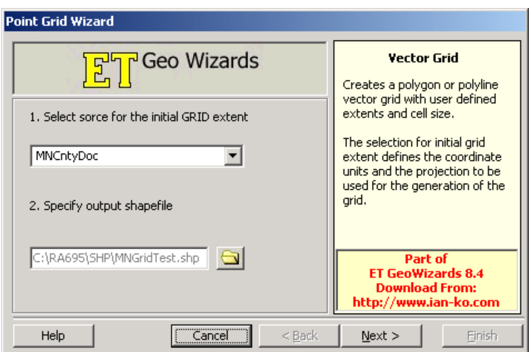


Figure 1. ETGeoWizards Vector Grid dialog box.

For illustrative purposes, figure 2 shows the grid placement over the Minnesota Counties shapefile. Only cells that intersected or were contained in the county shapefile were kept, cells that did not intersect any counties were deleted in edit mode. The final grid shapefile contains 234 polygons.

To most accurately attribute population into the grid cells, this project assumed that population was evenly distributed within census tracts. In reality, this is not always the case, however, high density population centers tend to have very small tracts based on area and the likelihood of these small tracts being divided among multiple grid cells is small. It was assumed that census tracts that were more likely to fall across multiple grid cells are the large, less populated tracts. The veracity of this theory can be tested by analyzing the number of unsplit tracts and by analyzing the percent of population that is wholly attributed into grid cells.

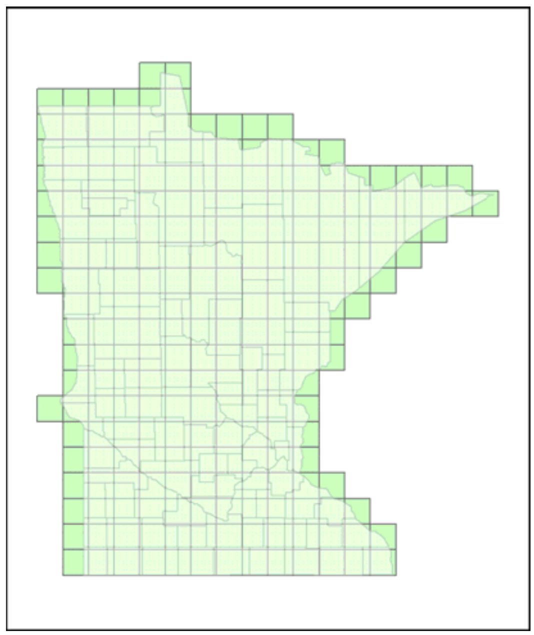


Figure 2. Projected 30x30 mile grid.

As previously stated, strong perpendicular and parallel border lines minimizes irregular grid cell line up, which was especially important when the cells were intersected as described later. One negative aspect that the Mercator projection has on any analysis is that it distorts area, especially over large extents like the state of Minnesota.

Area will be important when the population is attributed from census tracts to grid cells based on percentage of area. To minimize the distortion of area, the tract shapefile and grid shapefile were temporarily reprojected into the Albers Equal Area projection.

The Calculate Areas function of ArcGIS was used to calculate square meters of individual tracts in the census tract shapefile. Areas were also calculated for the vector grid shapefile to check the validity of the grid cell size. Because the grid shapefile was created in a Mercator projection and due to the nature of projection in general, there will be some natural distortion of area. However, due to the uniformity of the grid cells, variation will be minimal. A new column in the tract shapefile named F_AREA was created by the Calculate Areas function. Based on the projection, the calculated values are in square meters.

To calculate the percentage of a tract that falls into each grid cell, new polygons needed to be created that were essentially the combination of the two files. To accomplish this, the Intersect function was used to create a new shapefile that combined the grid and tract shapefiles.

Figure 3 shows the resulting polygon file created by the intersect function. The combined file contained 2,266 polygons. Data derivation shows that 60% of all tracts were unintersected. This also accounted for 60% of the population. Further derivation shows that the denser the population of a tract was, the less likely it was to be split just as previously hypothesized.

The calculate areas function was used to calculate areas of the tract-grid intersect polygons of this new file and attributes are collected in a new field named F_NAME_1. After the areas were calculated, another new field

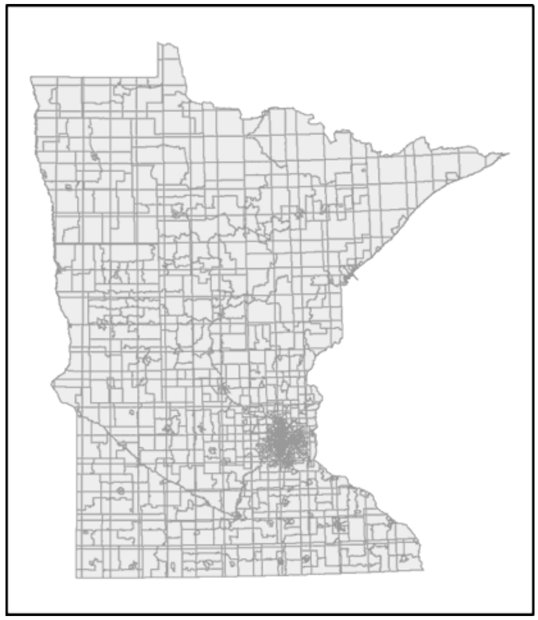


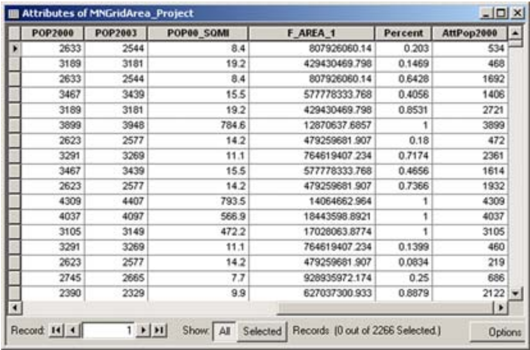
Figure 3. Tract/Grid intersection.

named PERCENT was manually created using ArcGIS. Values for this field are equal to the percentage of the tract that exists in the corresponding grid cell. The equation used to populate this field was:

[PERCENT] = [F_AREA_1] / [F_AREA].

The resulting percentage, as seen in figure 4, will be used to calculate the 2000 Population column to achieve population by intersected polygon. The next step was to then summarize the file by adding population back into the grid cells. The Dissolve command in ArcGIS was used to create the new vector grid shapefile, the old GRID_ID field being the Dissolve By field.

For data derivation purposes, the population is displayed as seen in figure 5. This map essentially mimics a population density map. This map will be helpful in identifying population


Attributes of MNGridArea_Project

POP2000	POP2003	POP00_SQMI	F_AREA_1	Percent	AttPop2000
2633	2544	8.4	807926060.14	0.203	534
3189	3181	19.2	429430469.798	0.1469	468
2633	2544	8.4	807926060.14	0.6428	1692
3467	3439	15.5	577778333.768	0.4056	1406
3189	3181	19.2	429430469.798	0.8531	2721
3899	3948	784.6	12870637.6857	1	3899
2623	2577	14.2	479259681.907	0.18	472
3291	3269	11.1	764619407.234	0.7174	2361
3467	3439	15.5	577778333.768	0.4656	1614
2623	2577	14.2	479259681.907	0.7366	1932
4309	4407	793.5	14064662.964	1	4309
4037	4097	566.9	18443598.8921	1	4037
3105	3149	472.2	17028063.8774	1	3105
3291	3269	11.1	764619407.234	0.1399	460
2623	2577	14.2	479259681.907	0.0834	219
2745	2665	7.7	928935972.174	0.25	686
2390	2329	9.9	627037300.933	0.8879	2122

Record: 1 Show: All Selected Records (0 out of 2266 Selected.) Options

Figure 4. Example of percent of land area recalculation into population by grid.

centers and gives the viewer an idea of the distribution of people throughout the state of Minnesota. For consistency and aesthetic purposes, the new vector grid shapefile was reprojected back into the Mercator projection.

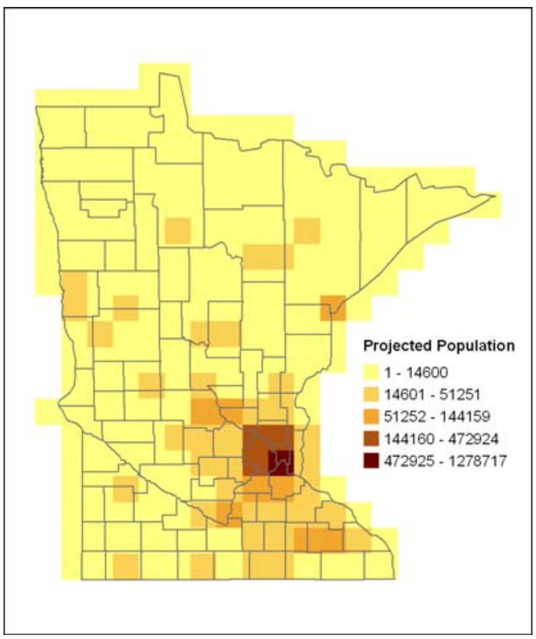


Figure 5. 2000 Population by grid cell.

The final data set that needed to be processed was the practitioner distribution file from BCBSM. Fields available include specialty type, practitioner name and practice address. The file was first limited to primary care defined as the following specialty types: general practice, family practice, internal medicine, pediatrics and geriatrics.

The second set was limited to physician extenders and included physician assistants and advanced nursing practitioners. Latitude and longitude were assigned after being processed by Centrus Desktop geocoding software (Boulder, CO). 92% of all practitioner records were matched to the addresses and the remaining 8% were assigned to the centroid of the zip code available from the data set. It is important to note that this 8% is distributed throughout the state and not concentrated in any particular area.

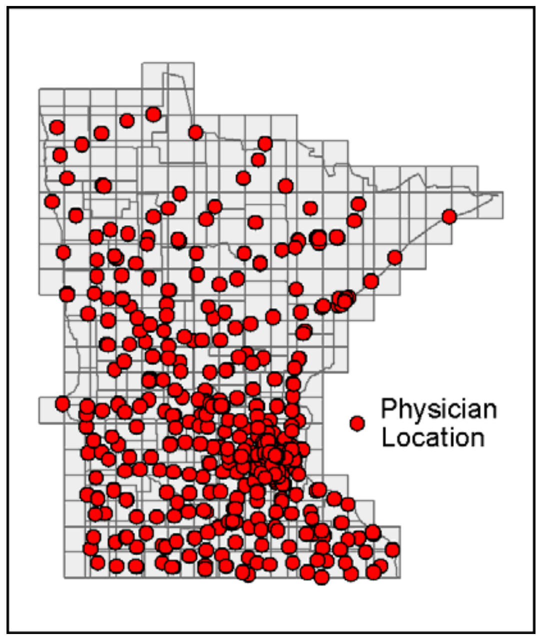


Figure 6. Geocoded location of physicians.

Figure 6 shows the distribution of physicians in the state of Minnesota.By spatially joining the practitioner shapefile to the vector grid cell shapefile, a sum of the number of practitioners was obtained by grid cell. Because this file captured all of the known practice sites for a practitioner, individuals may be counted more than once. This data sought only to provide a representation of providers by location.

It was assumed that most of the over representation of practitioners occurred in highly populated areas, areas that were unlikely to experience shortages.

Analysis

Using county and tract data to calculate HPSAs can be misleading. Figure 7 shows 2005 HPSAs according to the US Department of Health and Human Services' Bureau of Health Professions. Under this scenario, 32 counties were listed as either partial or full HPSAs. This suggests that over 41% of the Minnesota population was at least partially medically understaffed for primary care. The area that is considered fully under staffed comprises 14.2% of the state's land area and 3.3% of the Minnesota population.

An analysis of the provider data file from BCBSM showed a much different landscape. Figure 8 shows a distribution that is concentrated to just a few counties. Under this model, four counties accounted for 4% of the area and 1.1% of the population, a 65% percent decrease from the federally defined, full HPSAs.

Historical figures for 1988 listed the state of Minnesota as having 25 full and partial Health Manpower Shortage Areas (HMSAs – later renamed HPSAs). Unfortunately, individual counties were not listed. However, the total population affected was listed. A total of 243,199 persons lived in areas designated as either partial or full HMSAs (US Congress, Office of Technology and Assessment, 1990).

When compared against the 1990 US Census, this population accounted for 5.6% of the total. For comparison, the United States as a whole in 1988 had 31,357,445 persons living in HMSAs for

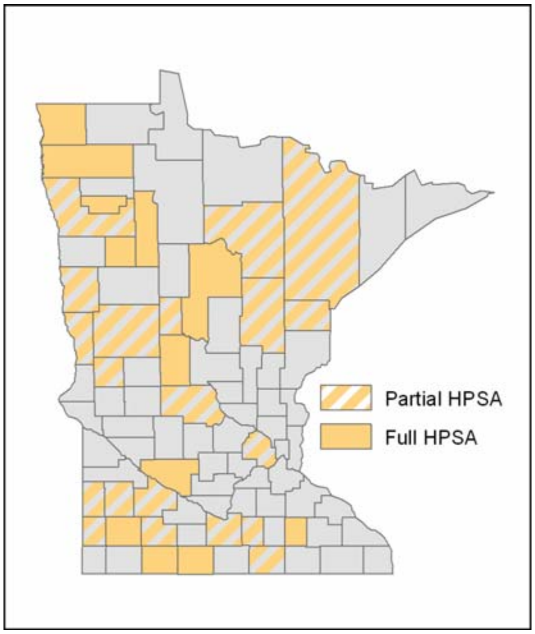


Figure 7. 2005 Health Professional Shortage Areas.

a total of 12.6% of the population. Recent research suggests that the current US percent is closer to 20%, an increase over the last decade (Laditka, et al., 2005). Minnesota is similar to the national trend. Although the exact methodology could not be obtained for calculating the partially under-staffed population in 1988, the 3.3% of the fully under-staffed population should be an increase over current numbers. For reference, historical practitioner data sets beyond the year 2000 were not available from BCBSM.

HPSA ratios, when applied to counties and other governmental boundaries, can produce skewed results. A county with a large area, such as St. Louis in Northeast Minnesota, qualifies as a partial HPSA. Further analysis showed that the population in this county is concentrated in two areas, Duluth in the extreme South and the Iron Range in the center. The federal, county-based model assumes that population is evenly distributed throughout its 6,737 square miles. However, anyone familiar with the Boundary Waters Canoe Area knows that vast portions of this county are uninhabited. The application of grid cells allowed for a standardization of area. Grid cells also attempted to mimic user consumption patterns. The Minnesota Department of Health, in its regulation of managed care organizations, suggests that primary care be accessible within 30 miles. By creating cells 30 miles square, the model attempts to capture this effect.

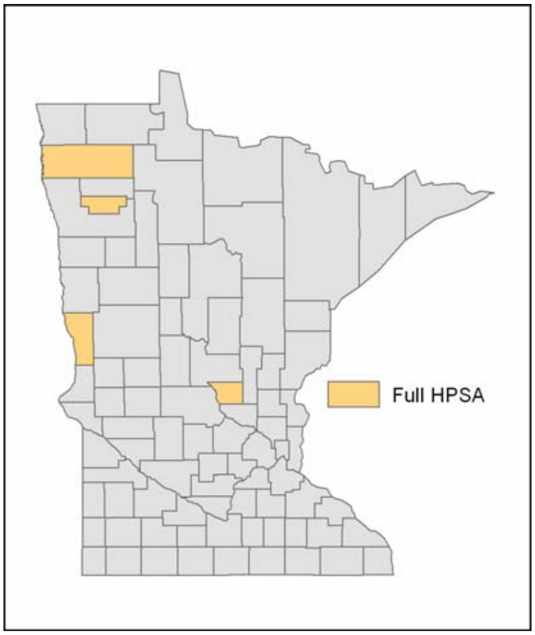


Figure 8. Health Professional Shortage Areas (HPSAs) using September 2005 Blue Cross Blue Shield of Minnesota Physician Data.

Calculating shortage areas in grid cells involved creating a few new fields and populating them with simple formulas. PHYSSHORTRT is the field that captured the ratio for each grid cell. The calculation was simply the attributed population field divided by the physician count and is expressed as:

[PHYSSHORTRT] = [ATTPOP200] / [PHYSCNT].

A second field, PHYSSHORTFL was created that populates with a "Y" any record that does not meet the 1:3500 ratio.

Figure 9 shows the distribution of grid cells that do not meet HPSA standards using the BCBSM primary care physician data set. Summary of the cells indicated that 192,824 people or 3.9% of the Minnesota population resided in these cells. Comparison against the federally defined HPSAs shows a 16.0% difference between the two populations.

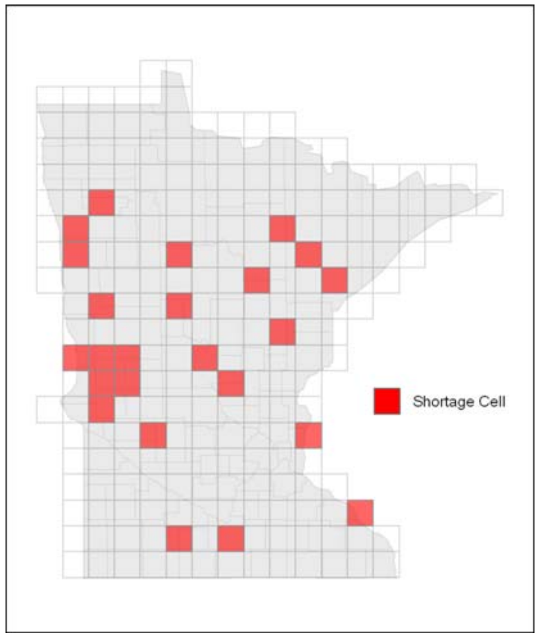


Figure 9. Physician shortages by grid cell using the BCBSM primary care physician data set.

One particular criteria for an area to be defined as an HPSA is that "primary care in contiguous areas are over utilized, excessively distant or otherwise inaccessible to the population of the area under consideration." (US Congress, Office of Technology and Assessment, 1990). Adjustment for grid cells that were adjacent to metropolitan areas was a process that attempted to capture the influence of having a large provider concentration within reasonable distance. To determine the locations of metropolitan centers, the distribution of cities with a population greater than 20,000 was displayed.

Figure 10 displays the manual selection of grid cells based on containment or immediate proximity to a population center. This selection is subjective, as no algorithm was available. Simply choosing cells that contained or bordered a population center greater than 20,000 would have lead to an overstatement of grid cells. Choosing cells that only contained a population center would have likely understated areas. A new field named METRO_ADJ was created that identified these subjectively selected cells and the results were displayed in figure 10.

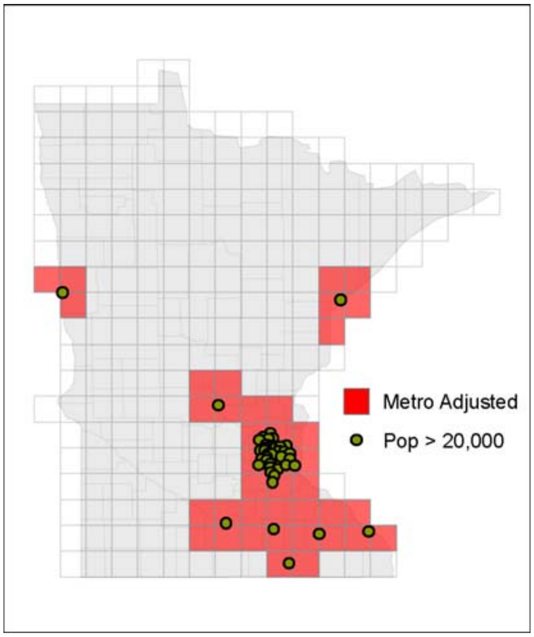


Figure 10. Cells considered “metropolitan” or adjacent to a metropolitan cell.

Figure 11 shows the physician distribution after the metropolitan area adjustment. A new field was created in ArcGIS named PHYSSHORTADJ that essentially voided an HPSA if it also was a cell that was metropolitan adjusted (METRO_ADJ=Y). Under this scenario, 129,823 people or 2.6% of the state’s population was classified as living in an HPSA. This is a 48.5% decrease in population over the model without metropolitan adjustment and a 24.8% decrease in population over the federally defined county model.

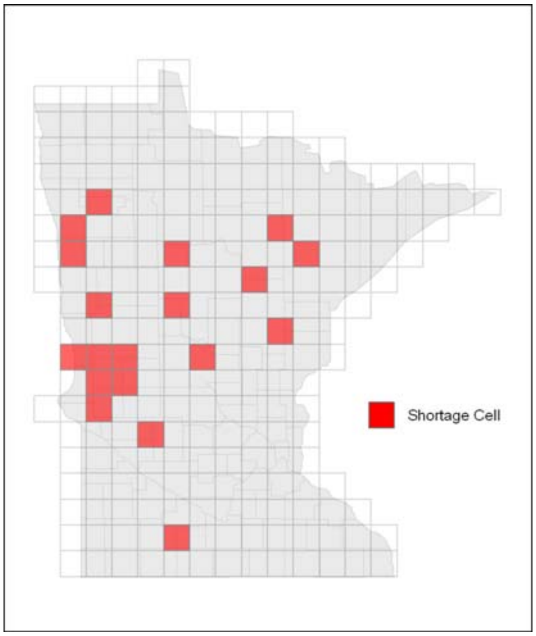


Figure 11. Physician shortages by grid cell, adjusted for proximity to metropolitan areas.

The use of paraprofessionals as primary care providers, most commonly as physician assistants and advanced nursing practitioners, has become a more acceptable practice. Deregulation and advanced training are allowing these individuals to participate more fully in the care of patients (Wessel, 2005).

The recent ability to prescribe drugs and provide education to patients has revolutionized this field. While their impact in patient care can be measured, it is relatively unknown what impact these practitioners have on the geographic distribution of shortage areas. In their research, Cooper, et al. (2002) suggest that physician extenders are providing care in rural areas where physician shortages are likely.

The addition of the paraprofessional population did in fact change the landscape of shortage areas, although not dramatically, as shown in figure 12. In this scenario, only 4 cells toggled from shortage to surplus. These 4 cells were not concentrated in any particular area, but were distributed throughout the state. The population now considered to be in shortage areas was 164,645, or approximately 3.3% of the total population. This accounts for a 17.1% improvement over the unadjusted, physician only model as was previously shown in figure 9.

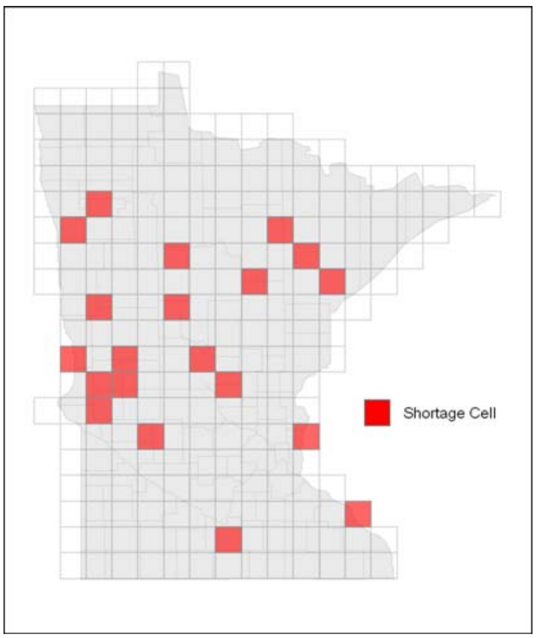


Figure 12. Physician/physician extender shortages by grid cell.

When the same metropolitan adjustment was applied to this combined practitioner group, results were comparable to the physician only model. Visual analysis of figure 13, however, shows that these shortages were more regionalized than in previous examples. Southern and much of Eastern Minnesota was void of any shortage areas. Shortage cells under this scenario now account for 114,868 people, or 2.3% of the total population. The metro adjustment was a 43.3% decrease from the unadjusted combined group. When compared to the physician only, unadjusted model, this scenario represents a 67.9% decrease in the population affected.

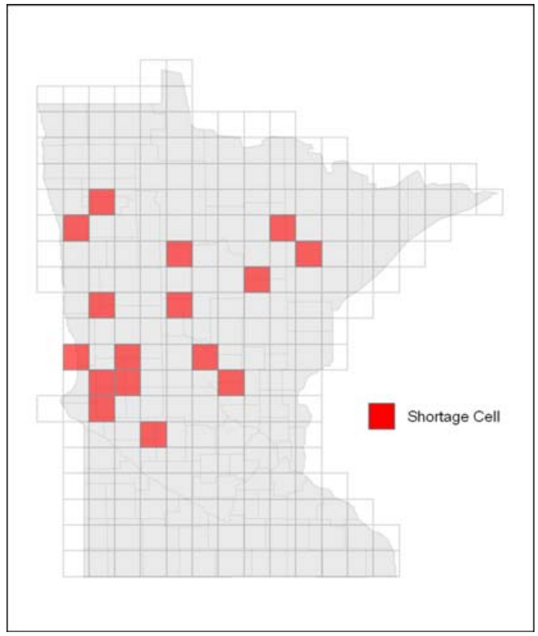


Figure 13. Physician/physician extender shortages by grid cell, adjusted for proximity to metropolitan areas.

Analytical comparison of the federally defined full HPSAs and the metropolitan adjusted combined group grid cells shows a 41.1% difference in the population affected. This suggests that either the HPSAs over inflated shortage areas with 162,048 people affected or that the shortage cell model under estimated these areas with 114,868 Minnesotan's affected. While the federal model is certainly easier to administrate by using widely available, county based data; the cell based analysis can show more accurate shortage distributions.

The practitioner population, similar to the general population, is one that is aging. Medical school output has remained relatively static and has not kept pace with the growth of the general population (Cooper, et al., 2002). Figure 14 shows the distribution of areas at risk for impending retirements.

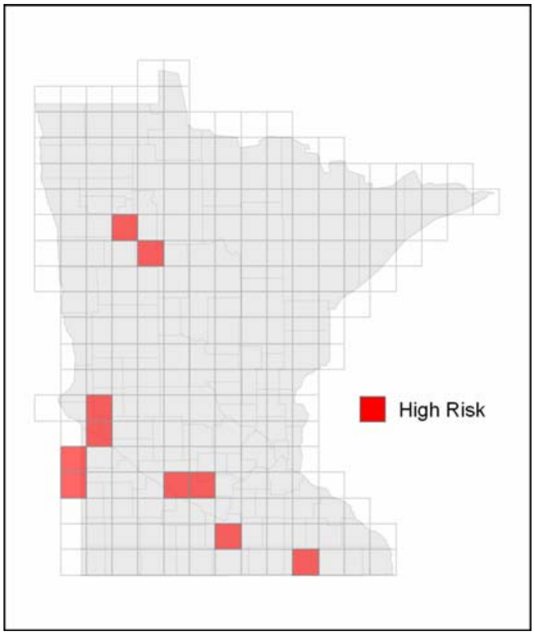


Figure 14. Areas of high risk/impending physician retirements.

This model summarized physicians by age to create an average by grid cell. Rates were not adjusted by using population projections for 2010 or beyond and were not adjusted for proximity to metropolitan areas. This model exists only to predict where shortages may occur based on impending physician retirements. This model also did not account for the impact that paraprofessionals will have on these grid cells. There is no simple way to predict where this population may move to and also the ages of physician extenders and nurse practitioners were unavailable.

Cells that have average ages greater than 54 years or older were flagged as high risk. A new field named AGE_FLAG was created and was populated by the results of the following equation:

[AGE_FLAG] = [PhysAvAge] > 54.

The geographic distribution of these shortages is now focused on the Southern and Western parts of Minnesota, contrary to the distribution of HPSAs. The Minnesota population that will be affected is 60,244, or 1.2% of the total population.

Results

By combining the results of the metropolitan adjusted HPSAs and the impending physician retirements by grid cells, a new pattern arises. Figure 15 shows that this distribution no longer is focused in Northwest Minnesota, but now includes a presence in Southern Minnesota.

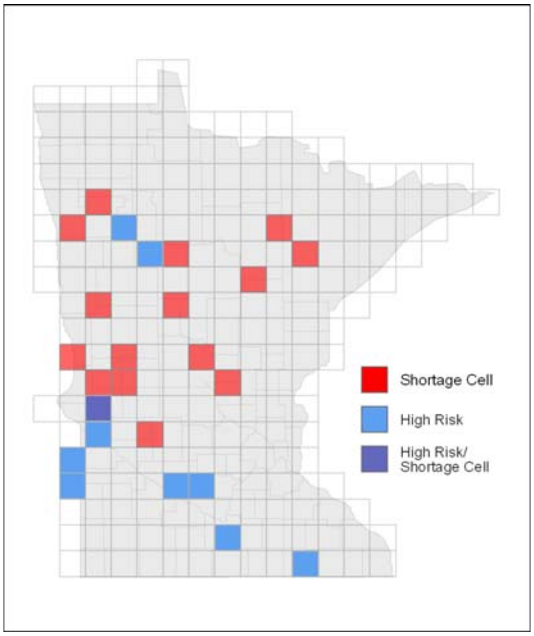


Figure 15. Current shortages and impending retirements.

Under this scenario, 157,493 people or 3.2% of the population lived in a current or impending shortage area. This accounted for a mere 2.9% decrease from the current, federally defined HPSA model. While state figures remain similar, the grid cell model was better equipped to analyze micro scale shortage situations.

Appendix A provides a summarization of the statistics presented in this analysis. Statistics were calculated using ArcGIS 9.1 and Microsoft Access 97.

Deliverables

Since it is county health boards that file to declare HPSAs, it is unlikely that a model that is not county based would be implemented on the state or national level. Non governmental agencies such as health plans may find use in such a model for a number of different situations. This model may be useful when looking at shortages of provider specialties. Rather than use the general population, membership could be substituted. This model would be an improvement over simple buffer models that have a difficult time portraying density of either members or providers.

Health plans may also find this model useful when analyzing contracting and marketing opportunities. Areas that have high provider counts and low membership may be targeted at open enrollment time. This data combined with employer locations and census demographic variables could be a useful tool for direct marketing purposes.

Discussion and Conclusion

The introduction of a new geographic unit produced very different results from the federally defined, county/tract based model. While the overall numbers only changed slightly when comparing HPSAs against grid cells, the geographic distribution varied greatly. The use of a standardized geographic unit not only allowed easy visual analysis; it also estimated density more accurately than the county/tract based model of the HPSA.

A certain limitation of the grid cell model is that it would need to be accepted at a large-scale level. Alignment of cells, size of cells, map projection and data source of the population are all aspects that would certainly introduce variation and thus different results. This model will work best in a controlled situation where results can easily be repeated.

Acknowledgements

I would like to thank Nancy Garrett, Boyd Lebow and Lynn Brown of Blue Cross Blue Shield of Minnesota for their help and advice towards the direction of this project. I would like to also thank Blue Cross and Blue Shield of Minnesota for allowing me the use of their provider data set and for supporting me in my career and educational development. Finally, I would like to thank my colleagues at St. Mary's University of Minnesota and the staff, especially John Ebert, Patrick Thorsell and Dr. Dave McConville for their project support, editing and encouragement.

References

Cooper, R., Getzen, T., McKee, H., Laud, P. 2002. Economic and Demographic Trends Signal An Impending Physician Shortage. Health Affairs, 21 (1): 140-154.

Laditka, J., Laditka, S., Probst, J. 2005. More May Be Better: Evidence of a

Negative Relationship between Physician Supply and Hospitalization for Ambulatory Care Sensitive Conditions. *Health Services Research*, 40(4): 1148-1166.

U.S. Congress, Office of Technology Assessment. 1990. Health Care in Rural America,OTA-H-434. U.S. Government Printing Office, Washington, DC.

U.S. Department of Health and Human Services, Bureau of Health Professions. 2005. Health Professional Shortage Area Primary Medical Care Designation Criteria, 42CFR. Retrieved October, 2005 from http://bhpr.hrsa.gov/shortage/hpsacritpcm.htm.

Wessel, L. 2005. Nurse Practitioners in Community Health Settings Today. *Journal of Health Care for the Poor and Underserved*, 16: 1-6.

Appendix A. Summarization of Statistics.

Geography	**Counties**	**Population**	**% of Total**
State of MN	87	4,919,471	100.0%
HPSA – Full County, 2005	12	162,048	3.9%
HPSA – Partial County, 2005	20	1,886,537	38.3%
HPSA – Combined, 2005	32	2,048,585	41.6%
BCBSM HPSA – Full, 2005	4	55,818	1.1%
HMSA – Combined, 1988	25	243,199	5.6%
Grid Cells	Cells		
Physician Shortage	24	192,824	3.9%
Physician Shortage - metropolitan adjusted	19	129,823	2.6%
Combined Shortage	20	164,645	3.3%
Combined Shortage – metropolitan adjusted	15	114,868	2.3%
Impending shortage (retirement)	10	60,244	1.2%
Final shortage (combined adjusted, retirement)	24	157,493	3.2%

Key Findings:

- 2005 Partial and Combined HPSAs are heavily skewed with the addition of Hennepin County. According to the US Department of Health and Human Services website, three extremely dense, urban census tracts meet special, non-geographic criteria that designate this county as an HPSA. The county population is just over 1 million persons and accounts for over 50% of the 20 partial county count. Removal of this county changes the HPSA partial, and the HPSA combined percent of total columns to 15.7% and 19.0%, respectively.
- There is a slight decrease in the percent of population living in HPSAs when paraprofessionals are added.
- There is a slight decrease in the percent of population living in HPSAs when comparing the full county, federal model to the combined, metropolitan adjusted vector grid cell model.
- While there appears to be an improvement from the historical 5.6% HMSA count, adjustment for impending retirements indicates that the percent of population in HPSAs will increase in the future.