ORIGINAL ARTICLE

ARTICLE ORIGINAL

Research methodology for the investigation of rural surgical services

Erik Ellehoj, PhD

Research and Evidence Branch, Alberta Health and Wellness

Joshua Tepper, MD

Institute for Clinical Evaluative Sciences, Toronto, Ont.

Brendan Barrett, MB

Professor of Medicine, Memorial University of Newfoundland, St. John's, Nfld.

Stuart Iglesias, MD

Rural Family Physician, Gibsons, BC

Correspondence to: Dr. Erik Ellehoj, 11456 43 Ave., Edmonton AB T6J 0Y4

See also Editorial on page 183, and accompanying articles, pages 195 and 207.

This paper describes a functional approach to the definition of rural populations for purposes of rural health care research. Rather than define "rural" directly, we created a definition of urban populations and our research target became the non-urban component. Using Geographic Information Systems technology, isochrones (drivetime zones) were created that attached suburban populations to urban centres and mapped non-urban populations into rural hospital catchment areas.

For population-based analyses, we have proposed a methodology for constructing catchment areas attached to Rural, Regional and Metropolitan services. We have developed a model for calculation of travel time for patients required to travel for care. We successfully applied these methodologies to the disparate regions of rural Alberta and Northern Ontario in 2 papers that investigated the delivery of rural surgical services.

This methodology represents a durable and portable designation of "rural" with potential for research applications in other areas of health research. By defining "urban" rather than "rural," we avoided many of the methodological conundrums in this research field.

Cet article décrit une approche fonctionnelle de la définition des populations rurales aux fins de la recherche sur les soins de santé en milieu rural. Au lieu de définir directement le mot «rural», nous avons créé une définition des populations urbaines et la composante non urbaine est devenue la cible de notre recherche. Nous avons créé, au moyen de la technologie des systèmes d'information géographique, des isochrones (zones de durée des déplacements) qui ont jumelé des populations suburbaines à des agglomérations urbaines et cartographié des populations non urbaines en bassins hospitaliers ruraux.

Pour les analyses démographiques, nous avons proposé une méthodologie de construction de bassins rattachés à des services ruraux, régionaux et métropolitains. Nous avons mis au point un modèle de calcul de la durée des déplacements pour les patients qui doivent se déplacer pour obtenir des soins. Nous avons appliqué avec succès ces méthodologies aux régions disparates des milieux ruraux de l'Alberta et du Nord de l'Ontario dans deux communications où nous avons étudié la prestation des services de chirurgie en milieu rural.

Cette méthodologie propose une désignation durable et transférable du mot «rural» qu'il pourrait être possible d'appliquer dans d'autres domaines de la recherche sur la santé. En définissant le mot «urbain» plutôt que le mot «rural», nous avons évité un grand nombre d'énigmes méthodologiques dans ce domaine de recherche.

Introduction

Canada's expansive geography is a defining feature of the country and an important consideration in the delivery of health services. A number of important health care decisions are related to geography and distance. Geographic areas can be defined for regional distribution of resources. Distances are used to help determine the accessibility of services to the population. Geographic patterns of health utilization can identify overlapping service provision and opportunities for rationalization of services. Distances are also used in determining funding models as well as defining recruitment and retention initiatives for health care providers.

With funding from Health Canada,

the Society of Rural Physicians of Canada and the Canadian Association of General Surgeons invited interested parties to collaborate on research that would examine (i) whether utilization and outcomes of surgical services were the same for rural and urban Canadians, and (ii) whether the presence of local rural surgical programs or the distance to travel for surgical care has an impact on utilization and outcomes.[1,2] To conduct this research we had to address 3 methodological issues, which are the focus of this paper. Our overall research targeted surgical services in Alberta and Northern Ontario, with a special emphasis on rural populations. Our research questions centred around the issues of access to surgical services by these rural populations and the health human resource profiles of surgical care in communities outside urban Canada.

Alberta and Northern Ontario were chosen, in part because of their contrast in geography, population distribution and rural surgical delivery systems. There is little agricultural land in Northern Ontario, due to the Canadian Shield. The non-urban population is clustered in small resource-based communities, with populations of usually less than 5000, with little or no population in the surrounding hinterland. Rural Alberta has an agriculturally based population. The non-urban communities are larger, with substantial catchment from the surrounding farms and ranches.

Three methodological issues

This paper explains the 3 methodologies used to i) define "rural," ii) define "catchment areas" (CAs), and iii) measure distances travelled by patients for health care.

i) First, we required a definition of "rural." There are many such definitions.[3] Ultimately, the most appropriate definition is usually chosen with regard to the available data and the context of the research questions.

ii) If access to surgical services locally is to be tested as an independent variable for utilization and outcome results, then geographic CAs must be created that attach populations to local hospitals. The challenges associated with identifying unique "markets" for individual hospitals are outlined by Thall and colleagues.[4]

iii) These projects required a methodology to measure distance. This measure had 2 specific applications in our research. First, we considered distance from an urban centre to be relevant to our designation of non-urban populations. Second, we required a measure of distance when we tested whether the obligation to travel for testing, consultation or surgery might influence utilization or outcome.

Defining "rural" by defining "urban"

Our methodology is based on a definition of "urban." The remainder of the population (non-urban) in both provinces became, by design, our rural population. Why non-urban Canada? In the evolution of the popular and prevalent primary/secondary/tertiary care paradigm, developed by/for urban Canada, training and privileges in procedural care is restricted to hospital-based specialist providers. However, rural Canada has few specialists. Our inquiries about access to and utilization of surgical services in communities with few or no specialists and our interest in generalists working outside the primary/secondary/tertiary care model required us to look outside of urban Canada. This huge (largely heterogenous in most other dimensions), non-urban population collectively shared the characteristic of having to access procedural care without a significant local specialist presence. The definition avoided the problems usually associated with "rural" and provided remarkable consistency within and between provinces. Acknowledging the heterogeneity of this default population, we built into our modelling exercise other independent descriptive variables, such as travel time to definitive surgical care and the size and scope of local surgical services.

We included in the urban population all of the Census Metropolitan Areas and a subset of the Census Agglomeration Areas (Table 1). We needed a decision rule to identify which of the Census Agglomeration Areas would be considered urban. In previous studies[5,6] of rural surgical services the authors have specified that health care programs are rural if they are provided exclusively, or almost exclusively, by family physicians with no, or few, local specialists. We adopted this decision rule and restricted our inclusion of Census Agglomeration Areas in our urban population only if they had a significant (>2) specialist presence (outside of general surgery) on the local medical staff. Ultimately, this cut-off was a population of 35 000 for a Census Agglomeration Area to be "urban."

Once the urban centres were identified, it was necessary to remove the surrounding population served by the facilities from the equation. We had, for other purposes, established a difference between Metropolitan centres (that included a medical school) and Regional centres (>35 000 with a significant specialist presence other than general surgery). In Alberta, the Metropolitan centres were Edmonton and Calgary. There are no Metropolitan centres in Northern Ontario that meet this definition. Instead, the city of London in Southwestern Ontario was selected.

We chose a 60-minute drivetime around the Metropolitan centres and a 30-min drivetime around the Regional centres to capture the urban population (60:30 rule). Our rationale for a longer drivetime for the Metropolitan centres was their significant size (>400 000) and services. These were expected to have a greater gravitational pull than the Regional centres. The choices of 60 and 30 minutes were arbitrary. The drivetime zones were validated by comparison with the 5 rural categories assigned* by Statistics Canada to the Enumeration Areas (EAs). The non-rural EAs occupied zones similar in size to the drivetime zones. EAs were used instead of Dissemination Areas because EAs were used for postal code assignment by Alberta Health and Wellness until late 2005. The use of a consistent set of boundaries was considered more important than the availability of more current boundaries, especially when these were amalgamated into groups.

The use of Statistics Canada's urban–rural categories was carefully explored, but they were unsuited to our purposes because of their inconsistent size. The 5 urban–rural categories were derived from workplace and residence location as indicated in the census. Some cities have large commuting areas, and others do not; the goal of this analysis was to investigate access to services that is based on travel time.

The isochrones for the urban population were created using the road networks surrounding each urban Census Agglomeration Area and Census Metropolitan Area, along with appropriate speed limits. Delays were added at intersections according to posted right-of-way rules. The hospital facility was chosen as the starting point for the drivetime calculations.

In summary, the urban population was defined as those people with postal codes within a 1-hour drivetime of a Metropolitan centre (medical school) and/or within a 30-min drive to the remaining Census Metropolitan Areas and urban Census Agglomeration Areas (pop. >35 000). This definition takes into consideration appropriate speed limits and delays at intersections. Our rural population is the non-urban population.

*Statistics Canada defines urban to rural areas in 5 categories: urban core, secondary urban core, urban fringe, rural fringe, and rural. Available: www.statcan.ca/english/census20001/dict/geo050.htm

Hospital catchment areas

Studies of health service delivery are typically designed to look at boundaries that reflect the "market" for each hospital. Hospital CAs ensure that there is only one facility within each reporting boundary. Hospital CAs allow the data to be analyzed in the context of the facility providing the service. Any changes that may be contemplated must be examined in the context of the data for the facility and for other relevant facilities. Any analysis of service disparity,[7] service access[8–10] or facility market

Table 1. Designation of Census Metropolitan Areas and Census Agglomeration Areas for Alberta and Northern Ontario

	Alberta		Northern Ontario	
Designation	City / town	Population	City / town	Population
CMA	Calgary	951 395	Sudbury	155 601
	Edmonton	937 845	Thunder Bay	121 986
CAA / Urban	Red Deer	67 707	North Bay	63 861
	Lethbridge	67 374	Sault Ste. Marie	78 908
	Medicine Hat	61 735	Timmins	43 680
	Fort McMurray	42 602		
	Grande Prairie	36 983		
CAA / Non-urban	Brooks	11 604	Elliot Lake	11 956
	Camrose	14 854	Haileybury–New Liskerd	12 867
	Cold Lake	27 935	Kenora	15 838
	Wetaskiwan	11 154		

CMA = Census Metropolitan Areas; CAA = Census Agglomeration Areas share[11,12] all require the availability of hospital CAs. Creating hospital CAs is, however, fraught with difficulties. The many challenges associated with this task are outlined by Thrall and colleagues[4] in an article aptly entitled "Delineating hospital trade areas: It's practically brain surgery."

Goody[13] summarizes much of the work performed in defining hospital CAs. The focus is on rural issues in the US. Goody's article lists a number of authors who have stated that the Metropolitan Statistical Area (MSA–US Census) is an appropriate estimate of market share for an urban area. A number of authors are also listed who claim that the County is an appropriate market area for rural hospitals (although problems with the methodology are also acknowledged).

Goody also explores the role of distance in the choice of facility selected by a client and lists several authors who have uncovered similar evidence. The author also uses utilization data to determine ZIP code assignments to hospital CAs. Cut-off percentages of 60% and 75% have been used to assign a ZIP code to a hospital CA.[13]

It is challenging to apply US findings to Canada because there are so many differences between the 2 countries. ZIP codes and postal codes are quite different, despite their common goal of identifying mailing addresses. In Canada, all facilities are publicly funded and the overall goal is to maximize access to services while minimizing costs. Much medical and population research is performed at the County level in the US, but in Canada counties serve little role in health care administration (with the obvious exception of ambulance services).

In the accompanying paper (p. 195) by Tepper and coworkers,[1] hospital CAs were an essential component of the analysis, since a critical question was "Does the level of service provided by a facility influence the observed rates of surgical procedures?" The facilities were categorized according to the level of service that they provide, but the population served by each facility was unknown, therefore hospital CAs needed to be created to do the research.

Alberta catchment areas

Alberta Health and Wellness created a set of hospital CAs in 1993, which were named General Hospital Districts (GHDs). These were created by examining the road connection, type of facility, capacity, population information, and services provided by each facility at the postal code level. The resulting boundaries contained the population that used the facility within the boundary as the primary hospital.

In 2002, an analysis using Canadian Institute for Health Information (CIHI) admission data for a package of primary care diagnoses was performed to see if these boundaries are still appropriate given changes in many of these facilities. The findings were that most of the facilities serve primarily their own GHDs.

Ontario catchment areas

No GHD-equivalent boundaries were available for Ontario, and therefore a set of hospital CAs was created for Northern Ontario. The rules used to create these boundaries were similar to those used in the original creation of the Alberta GHDs.

The approach was based on amalgamated EAs. Using census boundaries as a foundation reduces the possibility of split populations. EAs surrounding a facility were amalgamated using a combination of travel distances (on roads), postal code assignments, and Voronoi polygon. Boundaries between any 2 polygons are established at the half-way point between 2 facilities using linear distances.[14]

Isochrones were created for each facility for a variety of travel times (15, 30, 45, 60, 90 min, and 2, 3, 4 h). The isochrones for each category were compared facility-by-facility in context of the surrounding EAs. The assignment of postal codes to EAs is often one-to-many, in other words, a single postal code may be assigned to more than one EA. These assignments were examined carefully to reduce the possibility of a postal code being assigned to EAs in different CAs.

Voronoi polygons were used to create regions surrounding each facility based on straight-line (crow-fly) distances. These were used as a general guide in southern facilities and in order to determine flight travel times† in northern facilities. The EAs within each Voronoi polygon were merged to form a first draft of the GHDs. The isochrones and utilization data were then used to refine these boundaries and thus change the membership of the associated EAs. The resulting GHDs were created by amalgamating adjacent EAs based on crow-fly distances and then adjusted using isochrones and utilization data.

Postal code admission data for the same package of primary care diagnoses were obtained from CIHI and used to adjust the boundaries and create the final GHDs.

†Remote communities are located far away from larger health facilities. Health services are often obtained using airplanes as transport rather than personal vehicles. Care was exercised to avoid using isochrones in locations where they were not appropriate.

For a graphic description of the final CAs see the companion paper (p. 195) by Tepper and coworkers.[1]

Calculation of distances

Most spatial analysis, including distance calculations, are performed using a Geographic Information System (GIS).[15] A GIS is a combination of software, hardware and data designed to display, store, create and analyze geographic data. Alternatives to GIS are time-consuming. However, most standard GIS packages are poorly suited to provide a variety of distance estimates without specific add-on software, and thus the default options are often used as the only distance estimate. The default options that are described below include: crow-fly, road network, time travel analysis, and isochrones or drivetimes.

1. Crow-fly distances

The default option in all GIS is to calculate distances without any barriers in any direction. A travel distance of 50 km is simply a circle centred on the starting point with a radius of 50 km.

Crow-fly distance can be calculated in several manners. Several GIS offer the option to calculate distances based on projected coordinates or on great distance routes. The projected (or cartesian) option assumes that the portion of the world being examined has been projected onto a flat x–y plane. The characteristics of this plane are controlled by the selection of projection.

An alternative to this option is the use of spherical methods to calculate distances. In this case, the calculations are based on the shortest travel route and are thus better estimates of distance over long distances. The differences in calculated distances for the 2 methods are small for short distances, which are typically the most relevant distances for health facility access research. A more detailed description of the differences and alternative methods has been published.[16]

The principal advantage of crow-fly distances lies in its computational simplicity. The disadvantage is that, in reality, people follow road networks to access care (or other services) and thus these distances may not reflect the true accessibility for a given location with unique characteristics. For example, Fort McMurray, Alta., has a single highway to connect it to the rest of the province. Any access calculations that reflect a distance of more than 50 km will not be a proper representation of the access of this community. Crow-fly distances may also be particularly inaccurate where natural geography provides barriers to routine road development, such as large bodies of water.

Even in conditions with good road access the crow-fly approach can have significant limitations. Taber, Alta., has good road access in all directions, and the crow-fly approach is illustrated in Fig. 1. However, in reality, the quality of the roads vary due to the use of construction materials and to the road width. These factors determine the maximum speed at which the roads may be safely travelled. Figure 1 also illustrates a modification of the circle based on such considerations and the use of drive-time zones (explained below).

2. Road network distances

The road network approach to calculating distance utilizes the real travel distances on the existing road network. A barrier is assumed to be present if there are gaps or overlaps at any intersection. Any gaps in the road network will be reflected by the calculations performed using the network. The road network analysis takes into account transportation barriers, and alternative distances calculations assume that there is always a road present to link any community to another.

The ability to calculate road network distances is typically not available in most GIS, but add-on modules can be purchased to allow the user such functionality. Add-on packages are available for ArcView, ArcGIS, MapInfo and other platforms. Also needed are road networks geographic files. Special care is needed with these files to ensure that all roads are perfectly connected at the intersections. The calculation of road network distances is rarely used because of problems with the data (availability, cost and quality) and limited access to the software.

3. Travel time analysis

Estimated travel times are based on road network distances. Travel time can be obtained after performing road network analysis and by using the speed limits on every road segment. The default analysis is the shortest route between any 2 points. However, most of the software modules also allow for the calculation of the quickest route as long as the speed limit has been entered for every road segment in the network. The speed limits used in the accompanying papers by Tepper and coworkers and by Iglesias and colleagues[1,2] are shown here in Table 2. These assignments result in a conservative estimate, which is essential since these estimates must not reflect travel time under ideal circumstances, but rather average- to worst-case conditions.

Our methodology used travel time analysis both to define the urban population surrounding urban centres as well as to measure the distance for those rural patients obligated to travel for testing, consultation and surgery.

4. Isochrones (drivetime zones)

Analysis of access to services will require an analysis of distances from all possible sources to all possible service centres. An alternative is to create isochrones (drivetime zones) around each of the service centres. GIS analysis can then be used to determine the status of each origin point against the service isochrones.

In a simple scenario, a speed of 80 km/h could be assumed to create a circle around Taber, Alta. Figure 2 shows the 15-min drivetime zone using speed limits and accounting for intersections. The starting point is at the intersection of Highway 3 and Hwy. 36.

The compression in the northwest zone of the isochrone exists because the virtual driver must cross the entire town at slow speed with a large number of intersections. The shape stretches on the main highways due to their higher speed limits.

As the drivetime is increased to 60 min, the region takes a diamond-shape in more densely populated areas (Fig. 3). This map shows compression on the west side, which is a result of crossing through Lethbridge and the fact that the main highway (#3) veers north and then west again. The shape of the isochrones provides visual evidence of the road connectivity in every direction. In Northern Ontario, the patterns are similar.

The use of isochrones provides an efficient method to determine the travel time estimates from service centres to large numbers of potential client communities. Isochrones can be determined for every hospital or service of interest. The patterns in each community are determined by road availability near the community. The isochrones account for the Canada–US border by incorporating estimations of the time required to clear customs.

The use of road networks requires more effort than simple straight-line calculations, but the results provide a better reflection of the road network pat-

Table 2. Speed limits used in travel time analysis

Type of road	Speed limit, km/h
Large highways	110
Other primary highways	100
Secondary paved highways	60
Arterial roads	60
Streets	50

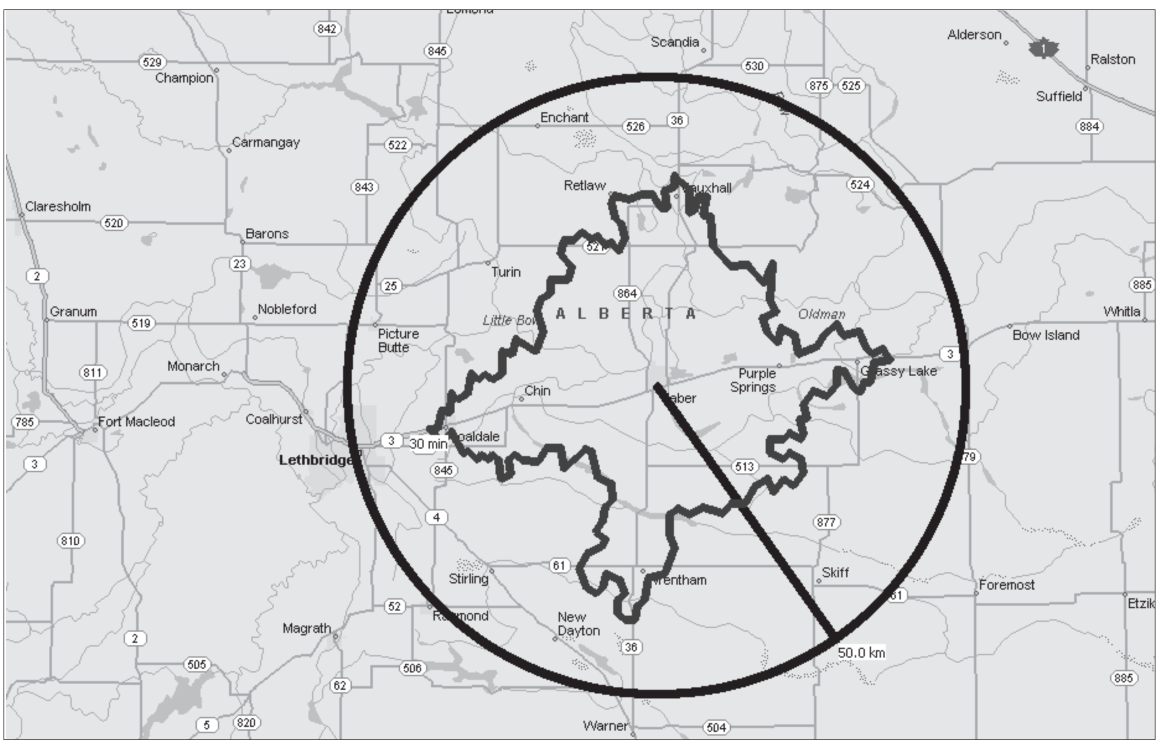

Fig. 1. Comparison of crow-fly (50 km radius) and drivetime (30 min) travel distance for Taber, Alta.

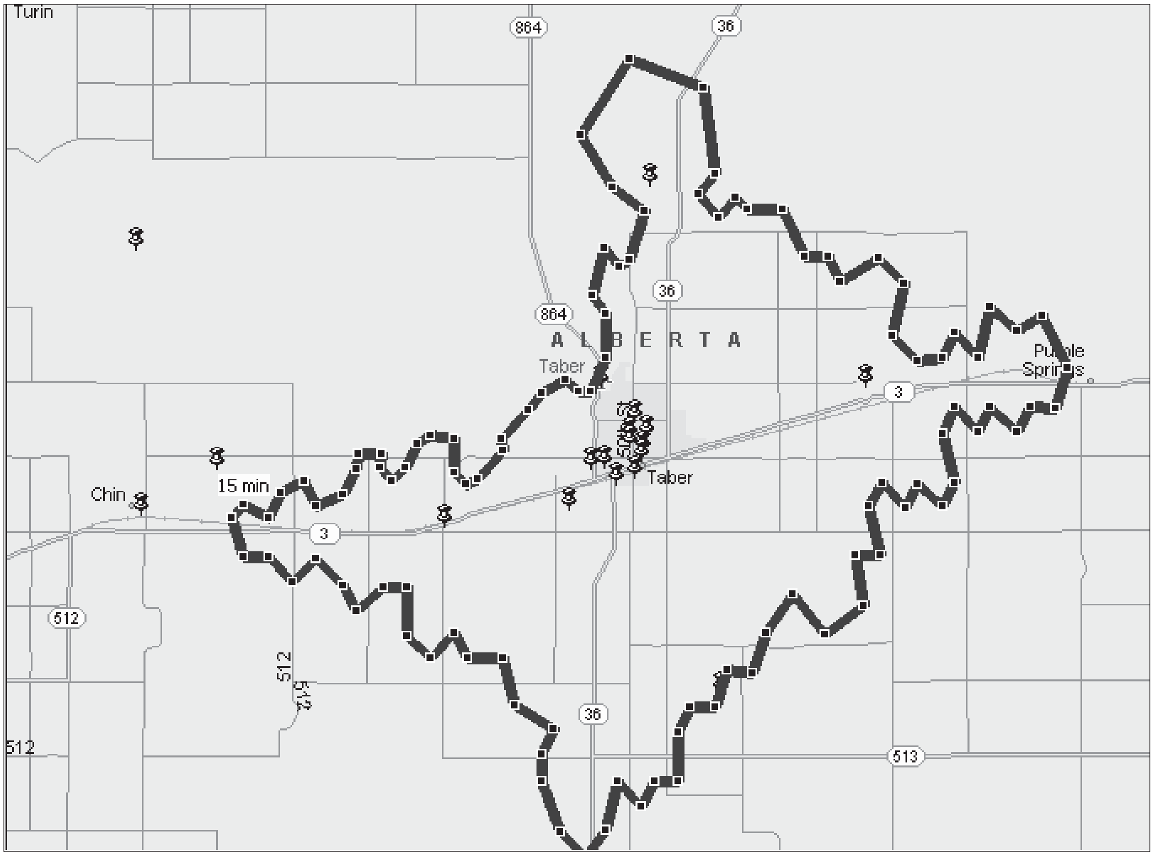

Fig. 2. 15-minute isochrone (drivetime zone) from Taber, Alta.

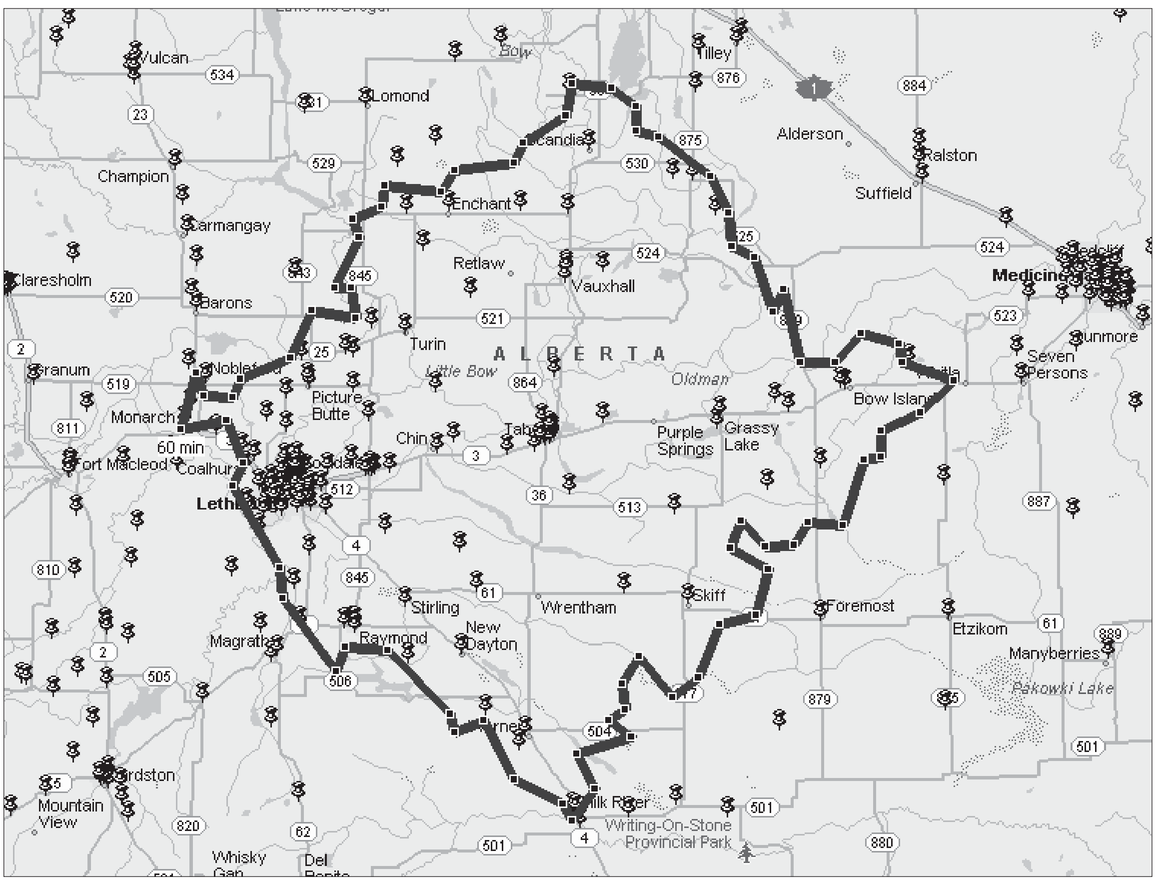

Fig. 3. 1-hour isochrone from Taber, Alta.

terns that patients will need to follow in order to gain access to health services.

Limitations

Our definition of "rural" required us to choose some arbitrary distance/drivetimes. Beyond these we would expect to see individuals seeking/requiring procedural care to access it through a rural hospital, rather than travelling directly to an urban clinic or emergency department. For those of us who have worked in rural Canada, the choice of a 60:30-min drivetime for Metropolitan/Regional hospitals made sense on an intuitive level. Our validation exercise for our CAs using CIHI separation data seemed to validate this drivetime: the patients outside the 60:30 drivetime did attend the local rural hospitals for their primary care. However, we did no sensitivity testing. It is possible that by making changes in the 60:30 rule we might have altered our findings.

Our CAs represent boundaries that are impossible to measure accurately. In reality, rural citizens "belong" to a hospital CA because of their perception of where their point of entry into the health care system might be, and not because of a drivetime zone. Depending on a variety of issues, most particularly the complexity and severity of the clinical problem, this point of entry will be different between persons from the same location. It will also be different for the same person with a spectrum of clinical illnesses. CAs, at best, represent approximations of the person-by-person resolution of these issues.

Isochrome analysis assumes that people will behave consistently throughout the year. The travel times outlined in this analysis were based on a yearly average. These travel times will vary within the year, based on season and weather. Another limitation is the accuracy of the speed limits, stop signs, and traffic lights used in the road network file. Our specifics were validated by comparison with Microsoft Map Point. Finally, isochrome analysis assumes that travel will occur by road, whereas it sometimes occurs by air. Our analysis used ">3 hours" as a highest category. For most remote communities, the total time of connecting to an airport, flying time and connecting to a referral hospital is >3 hours.

Conclusion

This methodology represents a durable and portable designation of "rural" with potential for research applications in other areas of health research. By defining urban rather than rural, we avoided many of the methodological conundrums in this research field. For population-based analyses, we have proposed a methodology for constructing CAs attached to Rural, Regional and Metropolitan services. Finally, we have developed a model for calculation of travel time for patients required to travel for care. We successfully applied these methodologies to the disparate regions of rural Alberta and Northern Ontario in a study of rural surgical services.[2]

Competing interests: None declared.

This article has been peer reviewed.

References

1. Tepper J, Pollett W, Yan J, et al. Utilization rates for surgical procedures in rural and urban Canada. *Can J Rur Med* 2006;11(3):195-203.
2. Iglesias S, Tepper J, Ellehoj E, et al. Rural surgical services in two Canadian provinces. *Can J Rur Med* 2006;11(3):207-17.
3. Pitblado JR, Pong R. Geographic distribution of physicians in Canada. Sudbury: Centre for Rural and Northern Health Research, Laurentian University; 1999.
4. Thrall G, Borden E, Thrall S. Delineating hospital trade areas: It's practically brain surgery. *Geospatial Solutions* July 2002. Available: www.geospatial-online.com/ (accessed 2006 June 12).
5. Iglesias S, Saunders LD, Tracy S, et al. Appendectomies in rural hospitals: safe whether performed by specialist or GP surgeons. *Can Fam Physician* 2003;49:328-33.
6. Iglesias S, Jones L. Rural surgical programs in Western Canada. *Can J Rur Med* 2002;7:103-7.
7. Soret S, McCleary K, Rivers P, et al. Understanding health disparities through geographic information systems. In: Khan O, Skinner R, editors. *Geographic information systems (GIS) and health applications*. Hersey (PA): Idea Publishing Group; 2003. p.12-42.
8. Luo W, Wang F. Spatial accessibility to primary care and physician shortage area designation: a case study in Illinois with GIS approaches. *Environ Plann B Plann Des* 2003;30:865-84.
9. Lin G. Using GIS to unveil distance effects on hospitalizations in Victoria. In: Khan O, editor. *GIS applications in health*. Hershey (PA): Idea Group Publishing; 2003. p. 243-59.
10. Mooney C, Zwanziger J, Phibbs C, et al. Is travel distance a barrier to veterans' use of VA hospitals for medical surgery care? *Soc Sci Med* 2000;50:1743-55.
11. Wolarski E. Using GIS with publicly available data for hospital market share analysis. In: Khan O, editor. *GIS applications in health*. Hershey (PA): Idea Group Publishing; 2003. p. 279-88.
12. Gesler W, Meade M. Locational and population factors in health care-seeking behaviour in Savannah, Georgia. *Health Serv Res* 1988; 23:443-62.
13. Goody B. Defining rural hospital markets. *Health Serv Res* 1993;28: 183-200.
14. Okabe A, Boots B, Sugihara K, et al. *Spatial tessellations: concepts and applications of Voronoi diagrams*. Hoboken (NJ): John Wiley and Sons; 1992.
15. Burrough PA, McDonnell RA. Principles of geographical information systems. Oxford University Press; 1988. p. 132-61.
16. Ellehoj EA, Schopflocher D, et al; for the Health Surveillance Branch, Alberta Health and Wellness. *Calculating distances in Alberta*. Alberta Health and Wellness Geographic Methodology Series, no 4. 2004. Available: www.health.gov.ab.ca/resources/publications/GeoDistances.pdf (accessed 2006 June 12).