A feasibility study using geographic access to general practices and routinely collected data in public health and health services research

Shanley Chong\textsuperscript{a,b}, Roy Byun\textsuperscript{a,c} and Bin B Jalaludin\textsuperscript{a,d,e}

\textsuperscript{a} Epidemiology, Health, People and Places Unit, Population Health, South Western Sydney Local Health District, NSW, Australia
\textsuperscript{b} South Western Sydney Clinical School, University of New South Wales, Sydney, Australia
\textsuperscript{c} Faculty of Dentistry, The University of Sydney, NSW, Australia
\textsuperscript{d} Ingham Institute for Applied Medical Research, University of New South Wales, Sydney, Australia
\textsuperscript{e} Corresponding author: b.jalaludin@unsw.edu.au

Abstract

\textbf{Background:} Advances in geographic information systems (GIS) and increased availability of routinely collected data have the potential to contribute to public health and health services research. The aim of this feasibility study was to explore the use of GIS to measure access to general practices and its relationship to selected antenatal behaviours.

\textbf{Methods:} We obtained the Perinatal Data Collection, 2004−2008, for South Western Sydney Local Health District from the New South Wales Ministry of Health. Using a GIS, we generated circular and polygon-based 1 km and 2 km road network buffers from a maternal residential address and then calculated the number of general practices within those buffers (density measures). Distances to general practices from maternal residences were also calculated in the GIS as another measure of geographic access to general practices (proximity measures). We used generalised estimating equations logistic regression to examine associations between access to general practices and its relationship to selected antenatal behaviours.

\textbf{Results:} The numbers of general practices within circular buffers were significantly greater than within polygon-based road network buffers. The average distance to the five closest general practices from a mother’s residence was 2.3 km ± 3.0 km (median 1.4 km, range 0.05 km to 40.0 km). For two of the five measures of access, a higher number of general practices was associated with increased risk for late first antenatal visit. However, the associations were weak, with odds ratios very close to unity. Choice of metrics of geographic access to general practices did not alter the associations with smoking during pregnancy or late first antenatal visit.

\textbf{Conclusion:} GIS can be used to quantify measures of geographic access to health services for public health and health services research. Geographic access to general practices was generally not associated with smoking during pregnancy or late first antenatal visit. GIS and routinely collected health data can be used to answer novel questions in public health and health services research.

Key points

- We used routinely collected health data and geographic information systems (GIS) to demonstrate the feasibility of using metrics of geographic access to health services in public health and health services research.
- GIS was used to quantify density and proximity metrics as measures of geographic access to general practices.
- There were generally null, or weak, associations between geographic access to general practices and smoking during pregnancy or late first antenatal visit.
Introduction

Routinely collected health data have been widely used in public health and health services research.\(^1\)\(^3\) Routinely collected data have the advantages of being relatively cost-effective and easily accessible, allowing large sample sizes and – importantly – enabling whole populations to be studied. However, there are also some disadvantages to using routinely collected health data: they are often collected for administrative purposes rather than for research or health services planning, so may not be fit for purpose, and there may be no or limited information on potential confounders, effect modifiers and other variables of interest.

Use of geographic information systems (GIS) in public health and health services research, and particularly in environmental health, is a recent development.\(^4\)\(^5\) The potential uses of GIS in health include simple mapping of diseases, identifying associations between exposures and spatial distribution of disease, estimating exposures in individuals within defined geographic regions, identifying populations at risk for environmental hazards, and identifying disease clusters.\(^6\) However, the use of GIS in health is not new. An early example of geomapping use was John Snow’s investigation of the 1854 cholera outbreak in London.\(^7\) Snow mapped households where cholera deaths had occurred, as well as the location of water pumps that provided potable water to those households, and demonstrated that most deaths were clustered around one water pump – the Broad Street pump.

Efficient and equitable access to health services is important for individuals and communities. Geographic access is one of many dimensions of access to health services. Commonly used geographic access metrics are proximity (or distance) and density. Two widely used proximity metrics are Euclidean (straight line between two points) and road network (the shortest possible route between two points along the road network) distances. Density metrics are calculated using spatial buffers – for example, the number of general practices within a certain radius of a residence. A circular buffer is simple to compute, whereas a road network buffer is more computationally intensive.

Antenatal care early in pregnancy is important for ultrasound assessment of gestational age, and providing information about nutritional supplements (e.g. folic acid supplementation) and screening programs (e.g. for Down syndrome and haemoglobinopathies).\(^8\)\(^9\) Smoking during pregnancy is associated with a higher risk of growth restriction, low birthweight and orofacial clefts.\(^10\)

In this feasibility study, we used routinely collected health data (perinatal data) and GIS to demonstrate how geographic metrics of access to general practices can be used in public health and health services research. We investigated associations between proximity and density metrics of geographic access to general practices, and late first antenatal visit and smoking during pregnancy. We hypothesised that increased geographic access to general practitioners before and during pregnancy would lead to early consultations and lower rates of smoking during pregnancy.

Methods

Maternal and infant data

We obtained the Perinatal Data Collection (PDC), 2004–2008, for South Western Sydney Local Health District (SWSLHD) from the New South Wales (NSW) Ministry of Health. SWSLHD is about 6231 km\(^2\) in area, with seven local government areas and, in 2011, had a population of 840 602.

The PDC is a population-based surveillance system of all live births and stillbirths of at least 20 weeks gestation or birthweight of at least 400 g. We obtained information on maternal demographic characteristics (age, country of birth), pregnancy, labour, delivery, the newborn (birthweight, gestational age), number of cigarettes smoked during second half of pregnancy, duration of gestation at first antenatal visit, plurality, parity and mother’s residential address. First antenatal visit was categorised into <14 weeks and ≥14 weeks (late first antenatal visit) gestation.\(^10\) Smoking during second half of pregnancy was categorised as ‘yes’ (any smoking) and ‘no’. A validation study of PDC against information from medical records showed low rates of missing data, and high levels of agreement for pregnancy duration at first antenatal visit and smoking during pregnancy.\(^11\) A previous study of women attending antenatal clinics in south western Sydney demonstrated good agreement between self-reported smoking status and smoking status classified by urinary cotinine levels.\(^12\)

Area socio-economic deprivation

We used the 2006 Index of Relative Socio-economic Disadvantage (IRSED) at the census collection district (CCD) level in the analyses as an ecological measure of area deprivation.\(^11\) The IRSED was categorised into tertiles – tertile 1 was designated as most disadvantaged, tertile 2 as middle disadvantaged and tertile 3 as least disadvantaged. All women were assigned an IRSED based on CCD residential address.

Access to general practitioners

We used density (number of general practices within buffers) and proximity (distance by road network to general practices) measures as proxies for geographic access to general practitioners. General practice addresses, provided by the NSW Ministry of Health, were geocoded, and then ArcGIS v10 (Esri Australia) was used to generate metrics for access to general practices. These metrics were number of general practices within 1 km and 2 km circular buffers (CBs) around mothers’ residences, number of general practices within 1 km and 2 km polygon-based road network buffers (PRNBs) around mothers’ residences,
and average distance from mothers’ residences to the five closest general practices. The PRNBs were created using the StreetPro Navigation (PitneyBowes Inc., US) road network file and an ArcGIS network analyst to calculate the endpoints of all possible routes up to the specified distance (1 km or 2 km) along the road network for each mother’s residence. The endpoints were then connected to form irregular polygons. Using the same StreetPro Navigation file, the distance to the five closest general practices from each mother’s residential address was computed using an ArcGIS network analyst.

In the example shown in Figure 1, there are five general practices within the 1 km CB and three general practices within the 1 km PRNB. The five closest general practices from the mother’s residence are shown, with 1 indicating the closest general practice and 5 indicating the furthest. ArcGIS was also used to generate kernel density surface maps.

Figure 1. Measures of geographic access to general practices

![Figure 1](image-url)  
Mother's residence  
GP  
Route to general practice  
1 km polygon-based road network  
1 km circular buffer  
Route

Statistical analysis

We used the Wilcoxon signed-rank test to examine differences in the number of general practices between CBs and PRNBs, and the Kruskal–Wallis test to compare the number of general practices within buffers by smoking status and late first antenatal visit. We used generalised estimating equations (GEE) logistic regression to assess associations between access to general practices, and smoking during pregnancy and late first antenatal visit. The final model adjusted for maternal age, whether the mother was born in Australia (yes/no), plurality (yes/no), first pregnancy (yes/no), population density within the buffers and area deprivation. Results from GEE models are presented as odds ratios (ORs), together with their 95% confidence intervals (95% CI). All analyses were conducted using Proc Genmod in SAS v9.4 (SAS Institute Inc., Cary, NC).

This study was approved by the NSW Population and Health Services Research Ethics Committee.

Results

Between 2004 and 2008, there were 63 074 births in SWSLHD. Most women were aged between 20 and 34 years (79.1%), 44% were born overseas, and most delivered a singleton infant (97.2%) (Table 1). Twelve per cent (n = 7649) smoked during pregnancy, and 33% (n = 21 067) had a late first antenatal visit (≥14 weeks).

Figures 2a and 2c are maps of the study’s more densely populated area, with the location of mothers’ residences and general practices, respectively. Figures 2b and 2d show the same information following kernel density estimation. Red areas indicate greater numbers of maternal residences (and hence births) and general practices; blue areas indicate fewer births and general practices.

Table 1. Smoking during pregnancy and late first antenatal visit by maternal characteristics, 2004–2008

<table>
<thead>
<tr>
<th>Maternal characteristic</th>
<th>Smoking during pregnancy</th>
<th>Late first antenatal visit</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>n</td>
<td>%</td>
</tr>
<tr>
<td>Age group (years)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;20</td>
<td>2 373</td>
<td>28.0</td>
</tr>
<tr>
<td>20–35</td>
<td>49 869</td>
<td>12.0</td>
</tr>
<tr>
<td>&gt;35</td>
<td>10 832</td>
<td>9.9</td>
</tr>
<tr>
<td>Country of birth</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Australia</td>
<td>35 135</td>
<td>17.3</td>
</tr>
<tr>
<td>Overseas</td>
<td>27 939</td>
<td>5.9</td>
</tr>
<tr>
<td>IRSED&lt;sup&gt;a&lt;/sup&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Most disadvantaged</td>
<td>20 986</td>
<td>13.3</td>
</tr>
<tr>
<td>Middle disadvantaged</td>
<td>20 704</td>
<td>12.6</td>
</tr>
<tr>
<td>Least disadvantaged</td>
<td>20 868</td>
<td>10.7</td>
</tr>
<tr>
<td>Plurality</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>1 772</td>
<td>12.6</td>
</tr>
<tr>
<td>No</td>
<td>61 302</td>
<td>12.2</td>
</tr>
<tr>
<td>Previous pregnancy&lt;sup&gt;b&lt;/sup&gt;</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>39 275</td>
<td>13.6</td>
</tr>
<tr>
<td>No</td>
<td>23 799</td>
<td>9.9</td>
</tr>
</tbody>
</table>

95% CI = 95% confidence interval; IRSED = Index of Relative Socio-economic Disadvantage; na = not applicable; OR = odds ratio  
<sup>a</sup> Data missing for 521 women  
<sup>b</sup> Data missing for 422 women
The mean number of general practices was 5.3 ± 8.4 (median 2, range 0–48) within the 1 km CBs and 14.9 ± 14.5 (median 10, range 0–73) within the 2 km CBs. For the 1 km PRNBs, the mean number of general practices was 2.9 ± 6.0 (median 1, range 0–43); for the 2 km PRNBs, the mean was 9.7 ± 11.7 (median 6, range 0–55). The number of general practices within CBs was significantly higher than within PRNBs ($p < 0.0001$ for all buffer sizes). The average distance of the five closest general practices to a mother’s residence was 2.3 km ± 3.0 km (median 1.4 km, range 0.05–40.0 km).

Women who smoked during pregnancy had significantly fewer general practices in their neighbourhood and longer average distances to the closest five general practices than those who did not smoke during pregnancy (Table 2). Conversely, mothers who delayed their first antenatal visit had significantly more general practices close to where they lived and shorter average distances to the closest five general practices (Table 2).
After adjusting for potential confounders, although the ORs were greater than 1, the associations between access to general practices and smoking during pregnancy were not significant, and all were close to unity (Table 3). The association between access to general practices and late first antenatal visit showed an unfavourable effect, with a higher number of general practices associated with an increased risk for late first antenatal visit. This was significant for the 1 km PRNB and the distance measure (Table 3).

Discussion

We used routinely collected health data and GIS in this study to demonstrate the feasibility of using metrics of geographic access to health services in public health and health services research. We used two GIS functionalities to quantify geographic access to general practices by density and proximity. Previous studies used simple CBs as density metrics to capture variables of interest. Although easy to implement in a GIS, this approach does not take into consideration the road network, and may include general practices that are not readily accessible by road. Therefore, we constructed PRNBs to generate more realistic buffers. However, not surprisingly, the numbers of general practices within CBs were significantly greater than within PRNBs. This is because CBs created using Euclidean distances have larger surface areas than PRNBs, and hence more general practices. We also used GIS functionality to adjust for population density in each buffer and generate a proximity metric (average road distance from a mother’s residential address to the closest five general practices) as a further proxy measure of access to primary health services.

Women who smoked during pregnancy had significantly fewer general practices close to their residence than women who did not smoke during pregnancy. This finding makes intuitive sense. However, a regression model showed no associations between the number of general practices and smoking during pregnancy. Women who delayed their first antenatal visit until ≥14 weeks gestation had more general practices in their vicinity than women who did not delay their first antenatal visit. In the adjusted regression model, these two associations to number of general practices in the vicinity remained significant. This was an unexpected finding because we had hypothesised that increased general practice numbers would mean greater access to general practitioners and hence visits earlier in pregnancy.

Our findings may have been unexpected but are not surprising. Our general practice metrics (density and proximity measures) are geographic or physical measures, and are only one component of accessibility to health services. Other important aspects include affordability, cultural acceptability and whether care is available when needed. The choice of which general practice to visit depends on many factors other than

Table 2. Access to general practices by buffer type and outcome variables

<table>
<thead>
<tr>
<th>Access to general practices</th>
<th>Smoking during pregnancy</th>
<th>Late first antenatal visit</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Yes</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Mean ± SD</td>
<td>Median (range)</td>
</tr>
<tr>
<td>Number of general practices within 1 km CB</td>
<td>3.9 ± 6.3</td>
<td>2.0 (0–47)</td>
</tr>
<tr>
<td>Number of general practices within 1 km PRNB</td>
<td>2.1 ± 4.3</td>
<td>1.0 (0–43)</td>
</tr>
<tr>
<td>Number of general practices within 2 km CB</td>
<td>12.3 ± 12.5</td>
<td>8.0 (0–73)</td>
</tr>
<tr>
<td>Number of general practices within 2 km PRNB</td>
<td>7.6 ± 9.6</td>
<td>5.0 (0–51)</td>
</tr>
<tr>
<td>Distance to closest five general practices (km)</td>
<td>2.5 ± 3.1</td>
<td>1.5 (0–24.4)</td>
</tr>
</tbody>
</table>

CB = circular buffer; PRNB = polygon-based road network buffer; SD = standard deviation
Note: p-values were <0.0001 using the Kruskal–Wallis test for smoking during pregnancy and late first antenatal visit for all metrics of geographic access to general practices.

After adjusting for potential confounders, although the ORs were greater than 1, the associations between access to general practices and smoking during pregnancy were not significant, and all were close to unity (Table 3). The association between access to general practices and late first antenatal visit showed an unfavourable effect, with a higher number of general practices associated with an increased risk for late first antenatal visit. This was significant for the 1 km PRNB and the distance measure (Table 3).
geographic access. For example, in the US, African Americans’ proximity to high-quality renal dialysis facilities did not lead to their receiving care at those facilities, and the study authors suggest that institutional and social barriers may play a role in health-seeking behaviour.\textsuperscript{16}

We have demonstrated some relatively straightforward methods to calculate density and proximity measures in GIS as metrics for geographic access to general practitioners. With increasing sophistication of GIS software and computational hardware, car travel time on the road network (the quickest route between two points) is another geographic access metric for health services that is now being widely used.\textsuperscript{17} Travel time can also be used as a density measure – for example, for the number of general practices accessible within 30 minutes. GIS-calculated travel time is highly correlated with reported travel time (\( r = 0.86 \)) and may be a more relevant metric than simple measures such as Euclidean or road network distance.\textsuperscript{18}

We did not use travel time in our calculations. Fortney et al.\textsuperscript{19} have demonstrated that road distances to the closest health provider explain about 97\% of variation in travel times to the closest health provider. In addition, travel time computation in GIS does not take into account unpredictable local conditions (such as traffic congestion and inclement weather) or temporal conditions (such as weekdays and weekends).\textsuperscript{18,20} It is open to question whether more resource-intensive and sophisticated travel time measures that incorporate traffic congestion and inclement weather are worthwhile.

There are some barriers (although not insurmountable ones) to using GIS. Its application in health research requires significant resources, technical skills and organisational commitment. Although some GIS software is freely available, it may have limited functionality, so may not be fit for purpose. GIS software with full functionality can be expensive to purchase and maintain. Further, application of network analysis on large datasets can be computationally intensive, requiring high-end, expensive computer hardware. High-level skills are required in preparing, exploring, modelling and presenting the data.\textsuperscript{21} In addition, access to address information or geocoded data is often denied (rightly or wrongly) because of the need to preserve confidentiality.

One limitation of our study is that we calculated geographic access to general practices, rather than access to general practitioners themselves, as measures of access to primary healthcare. Large group practices (five or more general practitioners per practice)\textsuperscript{22} are more common than solo or smaller group practices, and it would have been important to adjust for this. Unfortunately, we did not have the necessary information. Another potential limitation is that we measured geographic access to general practices from the mother’s residence. Women may have visited general practices and antenatal factors. Maternity units may or may not include visits to general practitioners for

<table>
<thead>
<tr>
<th>Access to general practices</th>
<th>Smoking during pregnancy</th>
<th>Late first antenatal visit</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Unadjusted OR (95% CI)</td>
<td>Adjusted OR (95% CI)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Number of general practices within 1 km CB</td>
<td>0.988 (0.981, 0.995)**</td>
<td>0.996 (0.988, 1.0041)</td>
</tr>
<tr>
<td>Number of general practices within 2 km CB</td>
<td>0.992 (0.987, 0.997)**</td>
<td>0.998 (0.991, 1.005)</td>
</tr>
<tr>
<td>Number of general practices within 1 km PRNB</td>
<td>0.990 (0.983, 0.997)**</td>
<td>0.998 (0.990, 1.005)</td>
</tr>
<tr>
<td>Number of general practices within 2 km PRNB</td>
<td>0.994 (0.989, 0.999)*</td>
<td>1.001 (0.995, 1.008)</td>
</tr>
<tr>
<td>Mean distance to closest five general practices (km)</td>
<td>1.001 (0.987, 1.015)</td>
<td>1.008 (0.994, 1.022)</td>
</tr>
</tbody>
</table>

* \( p < 0.05 \), ** \( p < 0.01 \), *** \( p < 0.001 \)
95\% CI = 95\% confidence interval; CB = circular buffer; OR = odds ratio; PRNB = polygon-based road network buffer

Note: Odds ratios were adjusted for maternal country of birth, maternal age, area deprivation, plurality, parity and population density.
antenatal care in their definition of ‘gestational age at first antenatal visit’, so this variable may not be reported consistently across maternity hospitals. Further, hospital factors, such as lack of resources, may be the cause of late first antenatal visits, rather than women choosing to delay antenatal visits. Finally, although not a study limitation, it should be noted that, because of the large number of subjects in our study, small differences, whether meaningful or not, are likely to be statistically significant.

Conclusion
This study aimed to demonstrate the feasibility of using metrics of geographic access to health services in public health and health services research. We used a number of GIS functionalities to quantify density and proximity metrics of access to general practices to determine whether geographic access to general practitioners was associated with smoking during pregnancy and late first antenatal visit. There were generally no, or only weak, associations between geographic access and smoking during pregnancy or late first antenatal visit. The choice of metrics of geographic access to general practices did not alter the associations with smoking during pregnancy or late first antenatal visit. Advances in statistical methods and GIS, as well as increased availability of routinely collected data, have the potential to contribute to our understanding of the relationship between geographic access to health services and health outcomes.

Acknowledgements
The authors would like thank the Centre for Epidemiology and Evidence, NSW Ministry of Health, for providing the birth data for the study and for approval to publish the manuscript.

Competing interests
None declared

Author contributions
SC drafted the manuscript and carried out the statistical analysis. RB gave feedback on the drafted manuscript and assisted in drafting the revised manuscript. BJ gave feedback on the drafted manuscript and assisted in drafting the revised manuscript.

References


