Assessment of Choice of Units of Analysis for Studying Associations between Geographic Accessibility to PHC Services and Socio-demographic Factors

Tayyab Ikram Shah[1], Lindsay Aspen[2], and Scott Bell[1, 3]

1 Geography, University of Saskatchewan, Saskatoon, SK, tayyab.shah@usask.ca
2 University of Saskatchewan, Saskatoon, SK, lindsay.aspen@usask.ca
3 Population Studies and Training Center, Brown University, Scott_bell@brown.edu

Abstract

The aim of this investigation is to assess the choice of geographical areal units for studying associations between geographic accessibility to primary health care (PHC) services and socio-demographic factors in a Canadian urban context. To achieve this, an accessibility score determined by physician to population ratios was calculated at both locally defined neighbourhood and census tract levels in two Canadian cities. The influences of units of analysis on accessibility score were analyzed empirically and a combination of global and local regression models (i.e., OLS and GWR) were applied to both types of units. Regression results demonstrate that the statistical modeling outcomes can be influenced by using different units of analysis which emphasize the use of units of analysis that are pertinent to policy and planning purposes.

Keywords: health geography, MAUP, GWR, units of analysis.

Background and Relevance

In large urban areas, geographic areal units that characterize suburban communities play an important role in the process of localization of health care resources with respect to population needs. In geographical studies, analytical and statistical results can be influenced by the geographical scale and zoning scheme (i.e., modifiable areal unit problem 'MAUP') used to delineate suburban communities. In recent years, considerable progress has been made in geographic accessibility research in addressing conceptual and methodological issues (For example S. Bell, Wilson, Bissonette, & Shah, 2012; Luo & Whippo, 2012; McGrail, 2012; Sara L., 2003; Fahui Wang, 2012); however, at a local scale such as suburban communities, more investigations are required to address the problems that arise with respect to the geographic areal units used to analyze the distribution of healthcare resources according to population needs. Generally, the modifiable areal unit problem (MAUP) can be categorized based on the contributing spatial aggregation factors which can modify analytical and statistical results: 1) scale effect, related to the number of areal units used(S. Bell, et al., 2012; Kwan & Weber, 2008; Schuurman, Bell, Dunn, & Oliver, 2007; Smiley et al., 2010), and 2) zonation effect, referring to the choice of boundaries or aggregation (Flowerdew, Manley, & Sabel, 2008; Stafford, Duke-Williams, & Shelton, 2008). This research investigates whether the associations between geographic accessibility to PHC services and socio-demographic characteristics vary in using different areal units for analysis in two Canadian cities.

Study Area and Data

This study is conducted in two Canadian cities: Toronto and Calgary (Figure 1). To investigate the MAUP effects, we selected two commonly used areal units of analysis in Canadian urban research: Neighbourhoods (NHs) and Census Tracts (CTs). 2006 census data were obtained at both dissemination area (DA) and Census tract (CT) levels; the DA data is used to prepare the neighbourhood figures. MAUP effects can be either scale- or zone-based in nature; the units of analysis for our research present a scale effect in Toronto (Population mean: NH=7839 and CT=4747) and a zonation effect in Calgary (Population mean: NH=4837 and CT=5313) (Table 1). A local form of the physician-to-population ratio (i.e., accessibility score) was used as the dependent variable. In order to calculate the accessibility score, a GIS based Three-Step Floating Catchment Area (3SFCA) method was applied (Aspen, Shah, Wilson, & Bell, 2012; S. Bell, et al., 2012; Bissonnette, Wilson, Bell, & Shah, 2012). The spatial patterns of the 3SFCA accessibility scores for both cities are mapped using a manual classification scheme (For Neighbourhood, see Figure 1a (Calgary) and 1b (Toronto); Census Tract, see Figure 1c (Calgary) and 1d (Toronto)). Units with higher accessibility scores indicate comparatively better access to health care resources for local residents.

Table 1. Descriptive statistics – 2006 Population Census

Statistics	Toronto		Calgary	
	Neighbourhood	**Census Tract**	**Neighbourhood**	**Census Tract**
Count (n)*	318 (325)	527 (531)	204 (223)	186 (186)
Population**	2,492,815	2,501,540	986,770	988,165
Mean	7,839	4,747	4,837	5,313
Median	5,273	4,640	3,863	4,873
Std. Deviation	8,415	1,850	3,607	2,749
Range	69,865	22,570	17,580	20,635
Minimum	325	155	300	310
Maximum	70,190	22,725	17,880	20,945

* Neighbourhood counts: non-zero counts (total counts)
**NH population is derived from DA level datasets

In preparing the socio-demographic variables, 2006 census data were used and eight variables were shortlisted (Andersen & Davidson, 2001; N. Bell & Hayes, 2012; Chateau, Metge, Prior, & Soodeen, 2012; Field, 2000; Matheson, Dunn, Smith, Moineddin, & Glazier, 2012; Pampalon et al., 2012; PHAC, 2012; F. Wang & Luo, 2005). Table 2 indicates some of the key characteristics of these variables. Note that all of the explanatory variables were expressed as percentages and higher values indicate higher health care needs. Units with no population data were excluded in this analysis. The DA data was used to prepare the neighbourhood variables.

Methods and Results

This research was performed by first determining an OLS regression model using a forward step-wise approach in SPSS (IBM Corp, 2011). Next, to explore whether the regression coefficients vary across space or not, we applied a GWR method with an adaptive spatial kernel in ArcGIS (ESRI, 2012). Further, to detect the presence of spatial autocorrelation in the regression residuals, we calculated the Moran's Index (MI) with queen contiguity spatial weight matrix. Maps were prepared using a quantile classification scheme with four classes. The results obtained from the OLS regression analysis of accessibility score can be compared in Table 3. Comparisons between Neighbourhood and Census tract areal units for both Toronto and Calgary cities based on GWR models are shown in Table 4 and 5, and Figures 2 and 3 respectively.

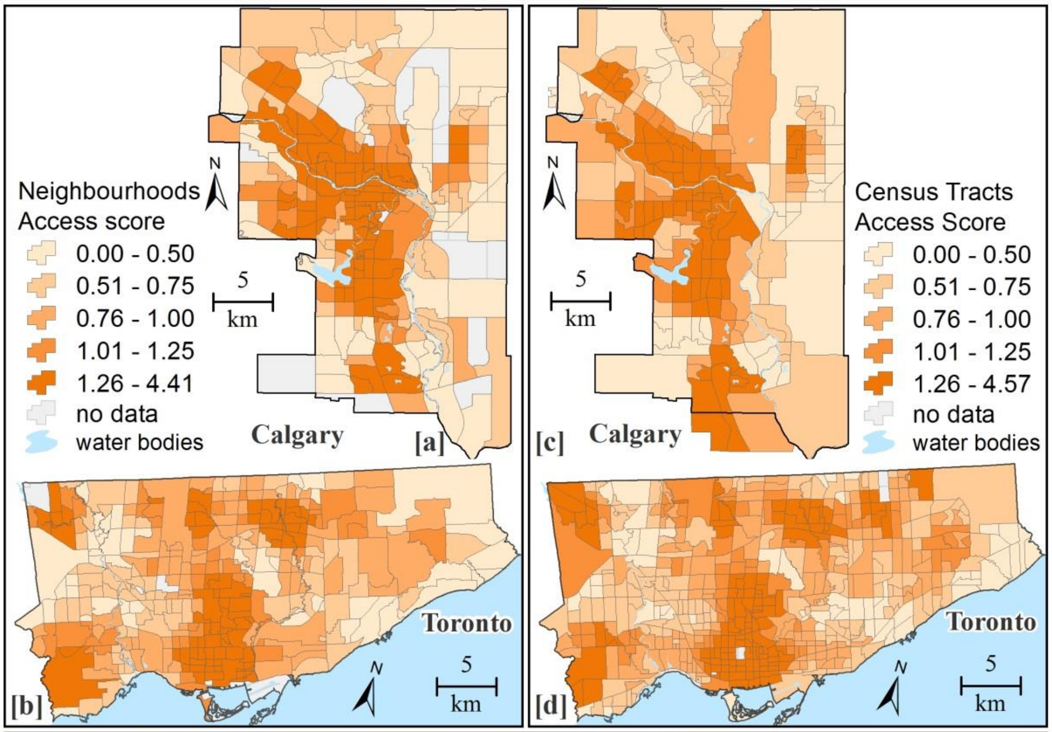

Figure 1 Neighbourhood accessibility score - Calgary (a), and Toronto (b), and Census Tract accessibility score - Calgary (c), and Toronto (d)

Table 2. Descriptive Statistics: Mean (Standard deviation 'SD')

Variables	**Toronto**		**Calgary**	
	Neighbourhood (n=318)	**Census Tract (n=527)**	**Neighbourhood (n=204)**	**Census Tract (n=186)**
Access Score	1.05 (0.64)	1.05 (0.66)	1.21 (0.98)	1.17 (0.92)
Percent of Population with High Needs	42.4 (4.03)	42.35 (4.03)	39.72 (5.67)	38.81 (4.41)
Percent of Home Owners	60.17 (21.09)	57.4 (23.18)	73.19 (21.76)	73.77 (19.96)
Percent of Lone Parents	18.94 (7.64)	19.91 (7.56)	14.84 (7.85)	15.75 (6.03)
Percent of Aboriginal Population	0.41 (0.46)	0.55 (0.6)	2.53 (2.2)	2.59 (1.82)
Percent of 65+ Living Alone	3.68 (2.47)	3.61 (2.38)	2.84 (3.42)	2.72 (2.87)
Percent of Recent Immigrants	9.5 (7.09)	9.97 (7.65)	5.07 (3.98)	5.32 (3.59)
Percent of 15+ less than High-school education	19.48 (10.06)	20.46 (9.71)	17.27 (8.11)	18.85 (8.09)
Unemployment rate	7.11 (2.52)	7.57 (2.76)	3.97 (1.75)	4.23 (1.29)

To assess MAUP effects and variability across space, we compared the measures of fit (Adj. R-squared), the number of significant variables found, coefficient estimates, and local coefficients for the predictors. Disparity in our results was observed with respect to the areal unit utilized in both cases (table 2 and 3). In both cities, the NH OLS models performed better over the CT models (Adj. R-squared for Toronto, NH=0.281 > 0.239=CT; for Calgary, NH=0.450 > 0.443=CT). In Toronto, five variables in the NH model (in comparison to four in the CT model) were found associated with accessibility score; whereas in Calgary, three variables (two common and one different) were found for both the NH and CT models. In all four models, predictors were found to be negatively associated with accessibility score, with the exception of the Living Alone variable in Calgary CT model. All coefficient estimates for the Toronto NH model were comparatively stronger than in the Toronto CT model (except No high-school education). The two common predictors in the Calgary models indicated stronger coefficient estimates for Home owners in the NH model and No high-school education in the CT model. Furthermore, our Moran's Index results point towards the presence of spatial autocorrelation in the residuals for the NH and CT models in both cities (Calgary, NH=0.601 < 0.729=CT; Toronto, NH=0.415 < 0.456=CT).

A comparison of adj. R-squared values from our GWR analysis indicate better performance for the CT model in Toronto (NH=0.425 < 0.571=CT) and better performance for the NH model in Calgary (NH=0.669 > 0.523=CT). The GWR method associates better model fit with increased variance and non-stationary as reported by the coefficients analyzed. Lone Parents in Toronto, as well as the No High-school and Home Owners variables in both cities follow the Adj. R-squared pattern mentioned above. An example of this can be seen in the No High-school variable for the Calgary NH model (table 5). Variation in coefficient was from -0.005 to 0.133 for NH, and -0.048 to 0.061 for CT. Spatial non-stationarity is observed in the NH model by a switch from negative to positive values from the Median (-0.013) to the 75th percentile (0.013). This is in contrast to the observed mean (-0.048), as well as the corresponding result for the OLS model (-0.036). These results can be visualized in Figure 3, where we see the Southwestern area of Calgary with a high concentration of values in the 0.013 – 0.133 range. Our Moran's Index results point towards improved spatial autocorrelation in the NH models for both cities (Calgary, NH=0.542 < 0.597=CT; Toronto, NH=0.252 < 0.311=CT).

Table 3. Results of OLS regression models

Variables	Toronto		Calgary	
	NH	CT	NH	CT
Constant	3.09***	2.43***	4.18***	3.51***
Home Owners	-0.018***	-0.011***	-0.028***	-0.022***
Lone Parents	-0.022***	-0.015**		
Recent Immigrants	-0.015**	-0.009*		
No High-school	-0.015***	-0.018***	-0.036***	-0.045***
Living Alone	-0.030*			0.054**
Aboriginal Status			-0.104***	
Adj. R-squared	0.281	0.239	0.450	0.443
AICc	525.9	912.9	455.6	393.3
Moran's Index (residuals)	0.601***	0.729***	0.415***	0.456***

***p<0.001; **p<0.01; *p<0.05

Table 4. Results of GWR models -

GWR Models	Toronto		Calgary	
	NH	CT	NH	CT
Neighbors	208	187	46	75
Adj. R-squared	0.425	0.571	0.669	0.523
AICc	464.1	629.9	386.5	376.7
Moran's Index (residuals)	0.542***	0.597***	0.252***	0.311***

***p<0.001

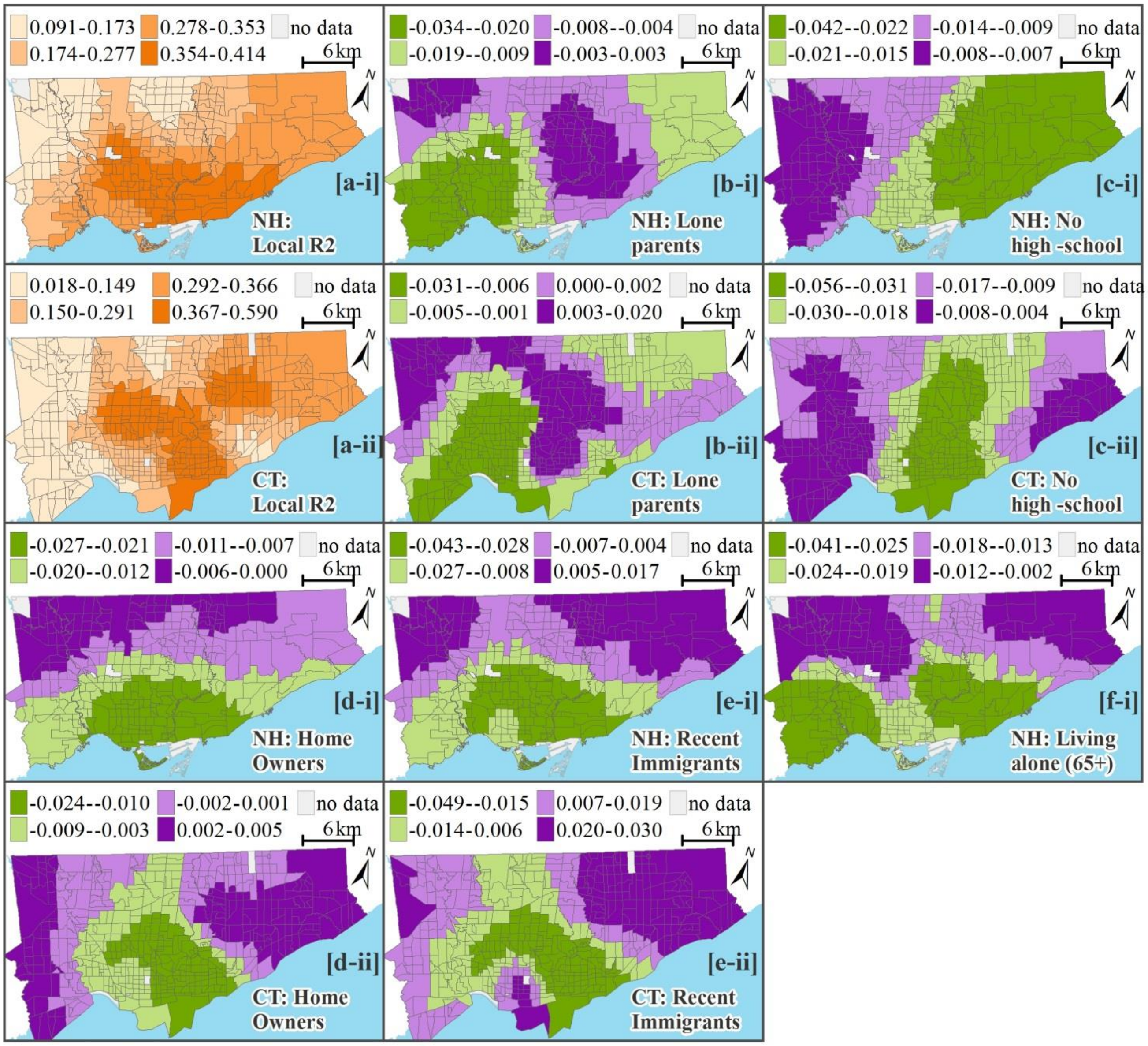

Figure 2. Results of Toronto GWR model by Neighbourhood (i) and Census Tract (ii)

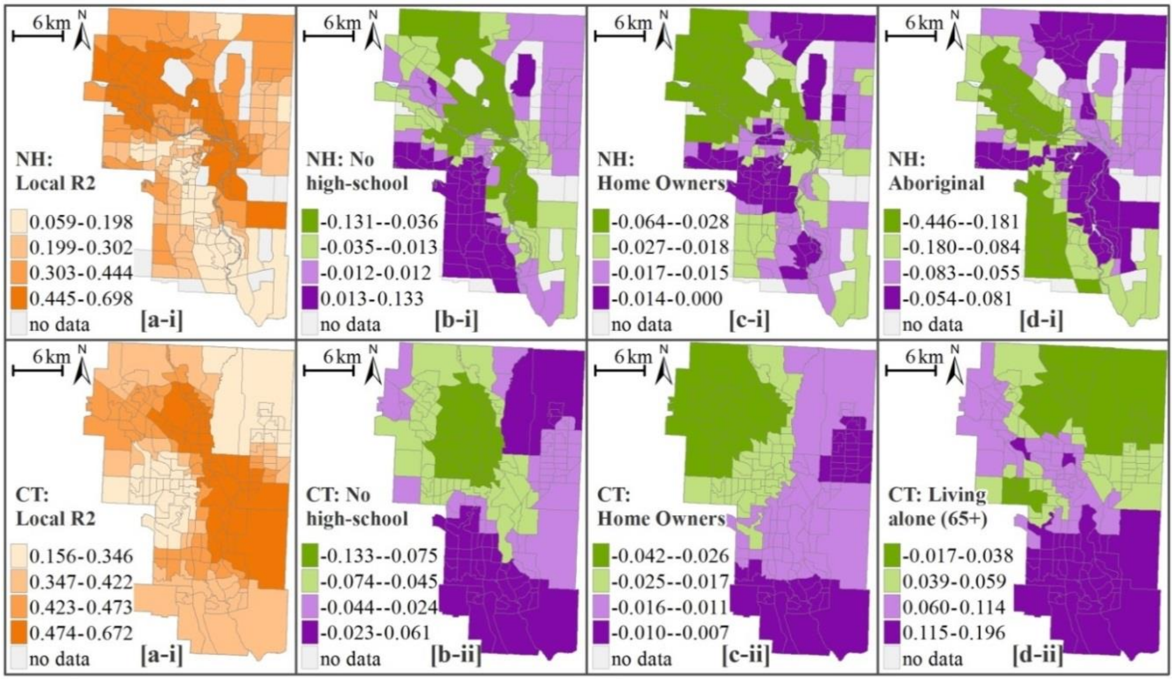

Figure 3. Results of Calgary GWR model by Neighbourhood (i) and Census Tract (ii)

Table 5. Results of GWR models – descriptive statistics

Items	City	Unit	Mean	Minimum	25th	Median	75th	Maximum
LocalR2	Toronto	NH	0.265	0.091	0.173	0.276	0.353	0.414
		CT	0.273	0.018	0.149	0.291	0.366	0.590
	Calgary	NH	0.328	0.057	0.201	0.304	0.449	0.694
		CT	0.409	0.156	0.343	0.421	0.473	0.672
Intercept	Toronto	NH	2.534	0.856	1.798	2.429	3.457	3.814
		CT	1.729	0.170	0.900	1.401	2.615	3.705
	Calgary	NH	3.342	0.110	2.061	2.882	4.437	6.968
		CT	3.246	0.301	2.254	3.322	4.268	5.921
Lone Parents	Toronto	NH	-0.012	-0.034	-0.020	-0.009	-0.004	0.003
		CT	-0.003	-0.031	-0.006	-0.001	0.002	0.020
	Calgary	NH						
		CT						
No High-school	Toronto	NH	-0.017	-0.042	-0.022	-0.015	-0.009	-0.007
		CT	-0.020	-0.056	-0.031	-0.018	-0.009	0.004
	Calgary	NH	-0.005	-0.131	-0.035	-0.013	0.013	0.133
		CT	-0.048	-0.133	-0.075	-0.045	-0.024	0.061
Home Owners	Toronto	NH	-0.013	-0.027	-0.021	-0.012	-0.007	0.000
		CT	-0.005	-0.024	-0.010	-0.003	0.001	0.005
	Calgary	NH	-0.023	-0.064	-0.028	-0.018	-0.015	0.000
		CT	-0.019	-0.042	-0.026	-0.017	-0.011	-0.007
Recent Immigrants (Toronto)/ Aboriginal Status (Calgary)	Toronto	NH	-0.011	-0.043	-0.028	-0.008	0.004	0.017
		CT	0.001	-0.049	-0.015	0.006	0.019	0.030
	Calgary	NH	-0.120	-0.446	-0.181	-0.084	-0.055	0.081
		CT						
Living Alone	Toronto	NH	-0.019	-0.041	-0.025	-0.019	-0.013	-0.002
		CT						
	Calgary	NH						
		CT	0.080	-0.017	0.038	0.059	0.114	0.196

Conclusions

A disparity in MAUP effects was observed in using Neighbourhood versus Census Tract boundaries with respect to population health care needs (accessibility) in Toronto and Calgary. In OLS regression, the use of Neighbourhood models indicated better representation of the data set over Census Tract models for both cities. A local form of regression (GWR) indicated that a Census Tract model performed better over a Neighbourhood model in Toronto; whereas the reverse was true in Calgary. It is unlikely that these results can be easily generalized for different cities, socio-economic variables, or dependent variables. Rather, it is important to understand the implications of our analysis towards areal unit choice, and to be aware of the difficulty in discerning which is the most appropriate without performing a proper analysis using both. In summary, using inappropriate areal units can result in poor interpretations of healthcare needs. This research highlights the importance of choosing an appropriate neighbourhood definition for suburban geographic areal units. As well, this research contributes to the existing body of literature on geographical accessibility to PHC services with a focus on large urban areas.

References

Andersen, R. M., & Davidson, P. L. (2001). Improving access to care in America. In R. M. Andersen, T. H. Rice & G. F. Kominski (Eds.), *Changing the US Health Care System: key issues in health services, policy, and management* (Vol. 1, pp. 3-30). San Francisco, CA: Jorrey-Bass.

Aspen, L., Shah, T. I., Wilson, K., & Bell, S. (2012). *Does Location Matter: Effects of Distance & Practice Size on Consumer Preferences for Seeking Primary Healthcare*. Paper presented at the HealthGIS 2012 - First SIGSPATIAL International Workshop on the use of GIS in Public Health, Redondo Beach, California. http://sysrun.haifa.il.ibm.com/hrl/healthgis2012/papers/healthgis-12.pdf

Bell, N., & Hayes, M. V. (2012). The Vancouver Area Neighbourhood Deprivation Index (VANDIX): A Census-based Tool for Assessing Small-area Variations in Health Status. *Canadian journal of public health, 103*(8), S28-S32.

Bell, S., Wilson, K., Bissonette, L., & Shah, T. (2012). Access to primary health care: does neighbourhood of residence matter? *Annals of the Association of American Geographers*. doi: 10.1080/00045608.2012.685050

Bissonnette, L., Wilson, K., Bell, S., & Shah, T. I. (2012). Neighbourhoods and potential access to health care: The role of spatial and aspatial factors. *Health & Place, 18*(4), 841-853. doi: 10.1016/j.healthplace.2012.03.007

Chateau, D., Metge, C., Prior, H., & Soodeen, R.-A. (2012). Learning From the Census: The Socio-economic Factor Index (SEFI) and Health Outcomes in Manitoba. *Canadian journal of public health, 103*(8), S23-S27.

ESRI. (2012). ArcMap Desktop Software (Version 10.1). Redlands, CA: ESRI Inc.

Field, K. (2000). Measuring the need for primary health care: an index of relative disadvantage. *Applied Geography, 20*(4), 305-332. doi: Doi: 10.1016/s0143-6228(00)00015-1

Flowerdew, R., Manley, D. J., & Sabel, C. E. (2008). Neighbourhood effects on health: Does it matter where you draw the boundaries? *Social science & medicine, 66*(6), 1241-1255. doi: DOI: 10.1016/j.socscimed.2007.11.042

IBM Corp. (2011). IBM SPSS Statistics for Windows (Version 20.0). Armonk, NY: IBM Corp.

Kwan, M. P., & Weber, J. (2008). Scale and accessibility: Implications for the analysis of land use-travel interaction. *Applied Geography, 28*(2), 110-123.

Luo, W., & Whippo, T. (2012). Variable catchment sizes for the two-step floating catchment area (2SFCA) method. *Health & place, 18*(4), 789-795. doi: http://dx.doi.org/10.1016/j.healthplace.2012.04.002

Matheson, F. I., Dunn, J. R., Smith, K. L. W., Moineddin, R., & Glazier, R. H. (2012). Development of the Canadian Marginalization Index: A New Tool for the Study of Inequality. *Canadian journal of public health, 103*(8), S12-S16.

McGrail, M. (2012). Spatial accessibility of primary health care utilising the two step floating catchment area method: an assessment of recent improvements. *International Journal of Health Geographics, 11*(1), 50.

Pampalon, R., Hamel, D., Gamache, P., Philibert, M. D., Raymond, G., & Simpson, A. (2012). An Area-based Material and Social Deprivation Index for Public Health in Québec and Canada. *Canadian journal of public health, 103*(8), S17-S22.

PHAC. (2012). What determines Health? In P. H. A. o. Canada (Ed.). Ottawa, ON: Public Health Agency of Canada.

Sara L., M. (2003). GIS and health care. *Annual Review of Public Health, 24*, 25-42.

Schuurman, N., Bell, N., Dunn, J., & Oliver, L. (2007). Deprivation Indices, Population Health and Geography: An Evaluation of the Spatial Effectiveness of Indices at Multiple Scales. *Journal of Urban Health, 84*(4), 591-603.

Smiley, M. J., Diez Roux, A. V., Brines, S. J., Brown, D. G., Evenson, K. R., & Rodriguez, D. A. (2010). A spatial analysis of health-related resources in three diverse metropolitan areas. *Health & place, 16*(5), 885-892. doi: 10.1016/j.healthplace.2010.04.014

Stafford, M., Duke-Williams, O., & Shelton, N. (2008). Small area inequalities in health: Are we underestimating them? *Social science & medicine, 67*(6), 891-899. doi: DOI: 10.1016/j.socscimed.2008.05.028

Wang, F. (2012). Measurement, Optimization, and Impact of Health Care Accessibility: A Methodological Review. *Annals of the Association of American Geographers*. doi: 10.1080/00045608.2012.657146

Wang, F., & Luo, W. (2005). Assessing spatial and nonspatial factors for healthcare access: towards an integrated approach to defining health professional shortage areas. *Special section: Geographies of Intellectual Disability, 11*(2), 131-146. doi: DOI: 10.1016/j.healthplace.2004.02.003