August 2005

Inequalities in neighbourhood socioeconomic characteristics: potential evidence-base for neighbourhood health planning.

Agricola Odoi
University of Tennessee, Knoxville, aodoi@utk.edu

R Wray
Hamilton District Health Council

M Emo
Hamilton District Health Council

S Birch
McMaster University

B Hutchison
McMaster University

See next page for additional authors
Follow this and additional works at: https://trace.tennessee.edu/utk_compmedpubs

Part of the Epidemiology Commons, Other Public Health Commons, and the Veterinary Preventive Medicine, Epidemiology, and Public Health Commons

Authors

Agricola Odoi, R Wray, M Emo, S Birch, B Hutchison, J Eyles, and T Abernathy

Research

Inequalities in neighbourhood socioeconomic characteristics: potential evidence-base for neighbourhood health planning

Agricola Odoi*[1,2], Ron Wray[2,3], Marion Emo[2], Stephen Birch[4], Brian Hutchison[4], John Eyles[5] and Tom Abernathy[1]

Address: [1]Department of Clinical Epidemiology and Biostatistics, McMaster University, Hamilton, Ontario, Canada, [2]Formerly with Hamilton District Health Council, Hamilton, Ontario, Canada, [3]Canadian Institute for Health Information, Toronto, Ontario, Canada, [4]Centre for Health Economics and Policy Analysis, McMaster University, Hamilton, Ontario, Canada and [5]School of Geography and Geology, McMaster University & McMaster Institute of Environment & Health, Hamilton, Ontario, Canada

Email: Agricola Odoi* - odoiag@mcmaster.ca; Ron Wray - rwray@cihi.ca; Marion Emo - memo@cogeco.ca; Stephen Birch - birch@mcmaster.ca; Brian Hutchison - hutchb@mcmaster.ca; John Eyles - eyles@mcmaster.ca; Tom Abernathy - tomab@telus.net

* Corresponding author

Abstract

Background: Population health planning aims to improve the health of the entire population and to reduce health inequities among population groups. Socioeconomic factors are increasingly being recognized as major determinants of many aspects of health and causes of health inequities. Knowledge of socioeconomic characteristics of neighbourhoods is necessary to identify their unique health needs and enhance identification of socioeconomically disadvantaged populations. Careful integration of this knowledge into health planning activities is necessary to ensure that health planning and service provision are tailored to unique neighbourhood population health needs. In this study, we identify unique neighbourhood socioeconomic characteristics and classify the neighbourhoods based on these characteristics. Principal components analysis (PCA) of 18 socioeconomic variables was used to identify the principal components explaining most of the variation in socioeconomic characteristics across the neighbourhoods. Cluster analysis was used to classify neighbourhoods based on their socioeconomic characteristics.

Results: Results of the PCA and cluster analysis were similar but the latter were more objective and easier to interpret. Five neighbourhood types with distinguishing socioeconomic and demographic characteristics were identified. The methodology provides a more complete picture of the neighbourhood socioeconomic characteristics than when a single variable (e.g. income) is used to classify neighbourhoods.

Conclusion: Cluster analysis is useful for generating neighbourhood population socioeconomic and demographic characteristics that can be useful in guiding neighbourhood health planning and service provision. This study is the first of a series of studies designed to investigate health inequalities at the neighbourhood level with a view to providing evidence-base for health planners, service providers and policy makers to help address health inequity issues at the neighbourhood level. Subsequent studies will investigate inequalities in health outcomes both within and across the neighbourhood types identified in the current study.

Background

Traditional health planning has typically focused on the practice and delivery of health care services. Population health planning, on the other hand, aims to improve the health of the entire population and to reduce health inequities among population groups [1]. The health of a population is influenced by several factors including but not limited to socioeconomic status [2], social support networks [3] education [4], ethnicity [5], employment [6], working conditions [7-10], physical environment [11], personal health behaviours [12,13], health care services [14] and individual coping skills [15,16]. Therefore, health planning, policy and interventions need to take into consideration not only the health care services, but also these broad determinants of health.

Since socioeconomic and demographic characteristics are important determinants of population health, adopting a population health approach to health planning at the neighbourhood level requires improved knowledge of the distribution of population socioeconomic and demographic characteristics at this level. Globally, there is an increasing interest in understanding the relationship between neighbourhood of residence and health of the population [17-19]. To this end, some researchers have suggested that improving the health of those living in the worst areas calls for systematically exploring area differences to inform social and health policy [20].

Currently, the lowest geographical level at which most health-planning data in Canada are analyzed is the municipal (city) level. Obviously, the use of such a large unit of analysis limits the ability to identify specific population characteristics as well as health variations and needs at the lower levels. The implication is that disparities in health outcomes and access to health care services across population sub-groups at these lower levels are unclear. Moreover, most large cities have diverse populations [21-24]; therefore the neighbourhoods within them have diverse socioeconomic and demographic characteristics that may influence neighbourhood population health outcomes and therefore health needs [25-27]. A number of studies have shown the extent and causes of neighbourhood socioeconomic inequalities [21-23,28-31]. In Canada, there is evidence that neighbourhood socioeconomic inequality has been rising since 1970 [32,33]. Moreover, numerous studies have reported associations between neighbourhood socioeconomic characteristics and various health outcomes [34-42]. Therefore, taking into consideration the diverse socioeconomic and demographic characteristics of the different neighbourhoods during health planning would ensure that planning and health services are tailored to the unique needs of the local residents of each neighbourhood.

Studies of geographical distribution of determinants of health have mainly used one of three approaches. The first involves either production of a single map showing the spatial distribution of a single variable (determinant of health) or production of a series of maps each showing the distribution of a single determinant of health [43-45]. The limitations of this method are that only one determinant can be assessed at a time(if a single variable approach is used) and assumes each determinant is independent of other determinants (if a series of maps is used). Moreover, when a series of maps is used, interpretation may be difficult. In the second approach, a composite index is created from combining two or more variables [46-48]. Although this approach mitigates the limitations of approach 1 and is effective in highlighting areas considered to be "high-risk", its drawback is that specific population characteristics (e.g. education, ethnicity, income, etc) are rolled into an index so that one cannot identify distinct characteristics (with respect to these variables) attributable to specific geographical areas. A third analytical approach uses factor analysis (or principal components analysis) to investigate several determinants [49,50]. This approach is a data reduction technique used to reduce the dimensionality of the data from several variables to a few factors (or principal components) that explain most of the variability in the original data. The current study uses principal components analysis and cluster analysis and Geographical Information Systems (GIS) to mitigate the limitations of the above approaches.

The objectives of this study were to use multivariate statistical techniques to identify the socioeconomic and demographic characteristics of neighbourhoods in the city of Hamilton, Ontario, Canada; and to classify the neighbourhoods based on similarities of these characteristics. Potential applications of the methodology in needs-based neighbourhood population health planning, service delivery, and policy development are proposed.

Methodology

Study area and geographical scale of analysis

The study was carried out in the city of Hamilton, Ontario, Canada. The city has a population of over 490,000 people and spans over 1,117 square kilometres [51]. A number of factors were considered in selecting the appropriate level of geography for the study. These included homogeneity of socioeconomic variables within the geographical unit; large enough population size to minimize the "small number problem"; data availability; acceptability by health planners and health service providers; and stability of the boundaries over time for maximum temporal data comparability for future analyses.

Based on the above criteria, census tracts were chosen as the most appropriate level of geography for the analysis. A

census tract (CT) is a small, relatively stable geographic unit usually having a population of 2,500 to 8,000 persons with an average of approximately 4,000 [52]. There are 132 CTs in the City of Hamilton. Census tracts are used in this study to represent neighbourhoods because CTs have 'neighbourhood-like' characteristics due to their homogeneity with respect to socioeconomic and demographic characteristics [53]. Therefore, throughout this paper, CTs and neighbourhoods are used interchangeably. There were several advantages of adopting this as the level of geography for analysis and future health planning: (i) Most census data are reported at this level of geography; (ii) Administrative health data can easily be aggregated to this level, if the postal codes of the health care recipients are known (this is because, in Canada, the postal code areas are smaller than CTs are so postal code data can easily be aggregated to the CT level); (iii) Census tracts are homogeneous with respect to socioeconomic and demographic characteristics; (iv) Their population sizes are large enough to allow calculation of relatively stable rates of most health events; (v) The boundaries of the CTs follow permanent and easily recognizable physical features and changes to their boundaries are discouraged to maintain maximum data comparability over time [52].

Data source and variable selection
Socioeconomic and demographic data for 833 dissemination areas (DA) in the city of Hamilton, Ontario, Canada, were extracted from the 2001 Canadian census data [52]. A DA is a small area composed of one or more neighbouring blocks and is the smallest standard geographic area for which all census data are disseminated in Canada [52]. The data were then aggregated to the CT level at which all analyses were performed. The variables used in the analyses were chosen based on their usefulness as determinants of health [54], reliability, and availability at the DA level [55]. Care was taken to include as many socioeconomic and demographic variables as possible in order to enhance the highest statistical differentiation between the CTs [55]. A total of 18 variables measuring the following characteristics were included in the analyses: demographic structure, social status, economic status, ethnicity, aboriginal status, and housing (see Table for a complete list of variables).

Statistical analyses
Variable standardization and correlation analysis
All statistical analyses were performed in STATA [56]. To overcome the impact of differing variances and different scales of measurements (e.g. dollars vs percentages), all variables used in the analyses were standardized to mean 0 and unit variances [57]. Had standardization not been performed, variables with high variances would unduly dominate the results of the analyses. A correlation matrix was constructed to explore relationships among the variables.

Principal components analysis
Principal component analysis (PCA) was used to reduce the dimensionality of data and to investigate the nature of the relationships among the CTs with the main objective of isolating the general features that best describe the variations in the data. Using this method, 18 inter-correlated socioeconomic and demographic factors were reduced to 5 principal components each of which represented different aspects of the original data. Kaiser criterion (eigenvalue one test) was used to guide the decision on the number of principal components to retain; all components with eigenvalues equal to or less than 1 were not retained since they explained variations equal to or less than any one of the original variables [58,59]. To maximize the variance of factor loadings and therefore aid the separation of CTs (or neighbourhoods) into homogeneous groups, varimax rotation was used [60].

Cluster analysis
Cluster analysis is a multivariate statistical technique used to organize observations into groups (or clusters) such that observations within a cluster have a high degree of similarity (or natural associations) among themselves while the clusters are relatively distinct from each other [57]. There are many different definitions of a cluster [57]. For the purpose of this study, we define a cluster as a set of entities that are alike.

There are two major classifications of cluster analysis techniques: hierarchical and non-hierarchical (or partition) techniques. This study adopted the partition cluster analysis, using k-means clustering methodology, to group CTs (or neighbourhoods) based on socioeconomic and demographic characteristics into clusters or neighbourhoods types. Using this methodology, the user specifies the number of clusters, say x, to create. These x clusters are formed through an iterative process. The algorithm begins with x seed values which act as the initial x group means. Observations are then assigned to the nearest group seed. After all observations have been assigned, cluster means are computed for each group. The initial cluster seeds are then replaced by their respective cluster means. The observations are then re-assigned to the nearest cluster mean. These steps continue until no observations change groups (clusters).

The optimum number of groups or clusters or "neighbourhood types" to be identified was decided upon using Calinski-Harabasz pseudo-F test [61] as well as the distribution of the CTs within the cluster. Five clusters (groups) were found to provide the most optimal separation of the CTs within the clusters. This would allow a rea-

Table 1: Variables included in the Hamilton neighbourhood analysis. Most definitions were adopted from Statistics Canada health region peer groups Study [55]. Those not adopted from the Statistics Canada study are: median income, married, live alone, population under 20 and non-official language population.

Variable name	Definition
Persons with less than grade 9 education	Percentage of the pop 20 years and over with less than grade 9 education
New immigrants	The percentage of immigrants who came to Canada from 1996 to 2001
Visible minority	Percentage of the population belonging to a visible minority group. As defined by the employment equity act (1986), visible minorities are persons (other than Aboriginal people) who are non-Caucasian in race or non-white in colour.
Aboriginal persons	Percentage of population reporting at least one Aboriginal origin (North American Indian, Métis or Inuit)
Median income	Median personal income for persons aged 15 and over, from all sources.
Government transfer income	Percentage of all income that came from government transfers (e.g., Canadian pension plan (CPP), guaranteed income supplement (GIS), old age security, etc.) for the population 15 years of age and older.
Incidence of low income (LICO)	Percentage of persons in economic families and unattached individuals with 2000 incomes below the Statistics Canada low-income cut-off (LICO). The cut-offs represent levels of income where people spend disproportionate amounts of money for food, shelter, and clothing. LICOs are based on family size and degree of urbanization; cut-offs are updated to account for changes in the consumer price index.
Non-official language pop	Percentage of the population not speaking any of the two official languages.
Unemployment rate	Total number of unemployed individuals 15 and older divided by the total number of individuals 15 and older participating in the labour force.
Average dwelling value	Average expected value of an owner-occupied, non-farm, non-reserve dwelling (including the value of the land the dwelling is on) at the time of the census
Owner-occupied dwellings	Percentage of dwellings in which the owner also lives. Band housing and collective dwellings (i.e. rooming houses, nursing homes, military camps etc.) Are excluded from both numerator and denominator.
Population under 20 years old	Percentage of the population under the age of 20 years
Population 65 years or older	Percentage of the population aged 65 years or older
Single-parent Families	Percentage of single-parent families among all census families living in private households. A census family refers to a married or common-law couple or lone parent with at least one never-married son or daughter living in the same household.
Married	Percent of legally married persons 15 & over
Live alone	Percent of persons living alone in private households
Internal migrant mobility	Percent of the population that lived in a different Canadian municipality at the time of the previous census. Excludes Canadians in households outside Canada (military & government personnel)
Population density	Total population of a census tract divided by its area in Km^2

sonable number of relatively homogeneous CTs (or neighbourhoods) per group (or neighbourhood type). Formation of more groups (neighbourhood types) resulted in some with too few CTs whereas formation of fewer groups resulted in loss of homogeneity within the groups.

The similarity (or dissimilarity) measure (also known as distance measure) used for the classification of the CTs was Minkowski distance metrics with argument 2 (L2) [57]. This measure, commonly known as euclidean distance, was calculated as follows:

$$L2 = \left\{ \sum_{k=1}^{p} \left(x_{ik} - x_{jk} \right)^2 \right\}^{1/2}$$

where: p (in this case 18) is the number of variables included in the cluster analysis; x_{ik} and x_{jk} are the values of variable k for CTs i and j respectively. The summations are over the p (or 18) variables involved in the cluster analysis. The 5 initial cluster centres were obtained randomly from among the CTs or neighbourhoods in the study area. For reproducibility a random number seed was applied before the 5 CTs were randomly chosen.

Since neighbourhoods belonging to the same group have certain socioeconomic and demographics characteristics

Table 2: Correlation matrix of variables used in multivariate analyses of socioeconomic and demographic variables in Hamilton neighbourhoods, 2004. Numbers not in brackets are pair-wise correlation coefficients whereas those in brackets are p-values. A: Persons with less than grade 9 education; B: New immigrants; C: Visible minority; D: Aboriginal persons; E: Median income; F: Government transfer income; G: Low-income persons; H: Non-official language population; I: Unemployment rate; J: Dwelling value; K: Owner-occupied dwellings; L: Population aged under 20 years; M: Population aged 65 years or older; N: Single-parent families; O: Married population; P: Population living alone; Q: Internal migrants R: Population Density

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
A	1.00																	
B	0.12 (0.166)	1.00																
C	0.28 (0.001)	0.69 (<.001)	1.00															
D	0.24 (0.006)	-0.04 (0.623)	0.12 (0.176)	1.00														
E	-0.65 (<.001)	-0.35 (<.001)	-0.48 (<.001)	-0.59 (<.001)	1.00													
F	0.70 (<.001)	0.31 (<.001)	0.37 (<.001)	0.11 (0.229)	-0.68 (<.001)	1.00												
G	0.52 (<.001)	0.50 (<.001)	0.61 (<.001)	0.22 (0.013)	-0.74 (<.001)	0.82 (<.001)	1.00											
H	0.66 (<.001)	0.59 (<.001)	0.62 (<.001)	0.04 (0.660)	-0.51 (<.001)	0.56 (<.001)	0.60 (<.001)	1.00										
I	0.49 (<.001)	0.34 (<.001)	0.44 (<.001)	0.11 (0.201)	-0.54 (<.001)	0.68 (<.001)	0.66 (<.001)	0.56 (<.001)	1.00									
J	-0.49 (<.001)	-0.45 (<.001)	-0.48 (<.001)	-0.19 (0.027)	0.68 (<.001)	-0.77 (<.001)	-0.82 (<.001)	-0.50 (<.001)	-0.54 (<.001)	1.00								
K	-0.25 (0.004)	-0.67 (<.001)	-0.61 (<.001)	-0.17 (0.054)	0.59 (<.001)	-0.62 (<.001)	-0.79 (<.001)	-0.50 (<.001)	-0.55 (<.001)	0.68 (<.001)	1.00							
L	-0.07 (0.426)	-0.12 (0.190)	0.10 (0.246)	0.09 (0.300)	0.14 (0.103)	-0.33 (<.001)	-0.16 (0.077)	-0.07 (0.413)	-0.12 (0.180)	0.28 (0.001)	0.38 (<.001)	1.00						
M	0.13 (0.132)	-0.06 (0.485)	-0.26 (0.003)	-0.22 (0.010)	0.38 (0.859)	0.38 (<.001)	0.01 (0.948)	-0.01 (0.906)	0.06 (0.484)	-0.16 (0.072)	-0.11 (0.195)	-0.76 (<.001)	1.00					
N	0.46 (<.001)	0.31 (<.001)	0.43 (<.001)	0.18 (0.039)	-0.57 (<.001)	0.67 (<.001)	0.80 (<.001)	0.41 (<.001)	0.57 (<.001)	-0.75 (<.001)	-0.58 (<.001)	0.08 (0.394)	-0.12 (0.184)	1.00				
O	-0.32 (<.001)	-0.36 (<.001)	-0.44 (<.001)	-0.23 (0.007)	0.64 (<.001)	-0.67 (<.001)	-0.86 (<.001)	-0.39 (<.001)	-0.56 (<.001)	0.79 (<.001)	0.81 (<.001)	0.40 (<.001)	-0.10 (0.240)	-0.73 (<.001)	1.00			
P	0.10 (0.252)	0.34 (<.001)	0.29 (0.001)	0.08 (0.344)	-0.42 (<.001)	0.58 (<.001)	0.62 (<.001)	0.27 (0.002)	0.41 (<.001)	-0.59 (<.001)	-0.77 (<.001)	-0.72 (<.001)	0.44 (<.001)	0.33 (<.001)	-0.80 (<.001)	1.00		
Q	-0.33 (<.001)	0.16 (0.071)	0.14 (0.103)	0.35 (<.001)	-0.03 (0.698)	-0.19 (0.032)	0.02 (0.791)	-0.04 (0.671)	0.00 (0.962)	0.10 (0.251)	-0.25 (0.004)	-0.03 (0.756)	-0.24 (0.005)	-0.11 (0.231)	-0.16 (0.076)	0.24 (0.006)	1.00	
R	0.18 (0.038)	0.42 (<.001)	0.43 (<.001)	0.11 (0.219)	-0.41 (<.001)	0.44 (<.001)	0.60 (<.001)	0.32 (<.001)	0.36 (<.001)	-0.62 (<.001)	-0.67 (<.001)	-0.28 (0.002)	0.05 (0.553)	0.48 (<.001)	-0.64 (<.001)	0.59 (<.001)	-0.02 (0.821)	1.00

in common, the resultant grouping provides useful insights into understanding the socioeconomic and demographic characteristics of each group (or neighbourhood type).

Cartographic manipulations
All cartographic manipulations were performed in ArcView GIS [62]. The principal components extracted during the PCA and the identified cluster resulting from the cluster analysis were exported to ArcView GIS. The geographical distribution of principal components 1–4 across the CTs was cartographically displayed in four different maps, one map per principal component. The spatial distribution of the five identified clusters were also displayed in one map. All CTs belonging to one cluster were represented with the same colour resulting in a map with five different colours each representing CTs belonging to the same cluster.

Results
Correlations
All the observed correlations were in the expected directions (Table 2). For instance, neighbourhoods with high proportions of low-income earners were more likely to have high percentages of non-official language population ($r = 0.62$), single-parent families ($r = 0.8$), population living alone ($r = 0.6$), and high population density ($r = 0.6$). Moreover, neighbourhoods with high percentage of visible minority population also tended to have high proportions of new immigrants ($r = 0.69$), low-income persons ($r = 0.61$), and low percentage of owner-occupied dwellings ($r = -0.61$). Additionally, neighbourhoods with high percentage of population with less than Grade 9 education, tended to have low median income ($r = -0.65$), and a high percentage of population receiving government transfer income ($r = 0.7$).

Principal components analysis
Only the first five principal components, that had eigenvalues greater than 1, were extracted (Table 3). The first PC

Table 3: Component loadings of socio-economic and demographic factors in the Hamilton, Ontario, Canada (2004)

Variable	Principal component					Uniqueness
	1	2	3	4	5	
Persons with <grade 9 education	0.575	0.176	0.659	-0.128	0.292	0.103
New immigrants	0.583	0.164	-0.337	0.539	0.243	0.171
Visible minority	0.643	0.419	-0.236	0.361	0.159	0.200
Aboriginal percentage	0.246	0.333	-0.165	-0.788	0.238	0.123
Median income	-0.790	-0.200	-0.125	0.378	-0.203	0.136
Government transfer income	0.847	-0.153	0.372	-0.053	0.017	0.118
Low income	0.941	0.110	0.018	-0.026	-0.144	0.081
Non-official language pop	0.675	0.255	0.197	0.342	0.424	0.143
Unemployment rate	0.721	0.103	0.165	0.036	0.064	0.437
Average dwelling value	-0.874	0.044	-0.103	0.042	0.203	0.181
Owner-occupied dwellings	-0.867	0.099	0.354	-0.096	0.000	0.104
Population under 20 years old	-0.324	0.867	0.089	0.045	-0.228	0.082
Population 65 years or older	0.139	-0.861	0.282	0.014	0.217	0.113
Single-parent families	0.759	0.252	0.160	-0.073	-0.437	0.138
Married percent	-0.875	0.148	0.190	0.172	0.251	0.084
Percent living alone	0.711	-0.552	-0.337	-0.075	-0.010	0.071
Percent of internal migrants	0.036	0.112	-0.790	-0.280	0.292	0.198
Population density	0.681	-0.113	-0.225	0.101	-0.325	0.357
Eigenvalue	8.30	2.42	1.92	1.47	1.06	-
Percentage of variation explained	46.10	13.42	10.69	8.14	5.88	-
Cumulative % of variation explained	46.10	59.52	70.35	78.35	84.23	-

is associated with the largest eigenvalue. This PC is a linear combination of the variables that account for the highest variability (46.1%) in the data. The second PC explains the highest variability not accounted for by the first PC while the third PC explains the largest variability not accounted for by the first two PC and so on. The first five PCs together accounted for a total of 84.23% of the total variation in the data (Table 3).

The first Principal Component is the most important since most (12) of the variables loaded heavily on it (Table 3). The component loadings measure the relationships of the socioeconomic and demographic variables with each of the PCs. The values of the loadings range from -1 to 1. The uniqueness values of almost all the variables were relatively low with the highest being 0.437 and the lowest 0.071. Uniqueness is the percentage of variance for a variable that is not explained by the PCs. For instance, almost all the variations for low-income, population under 20, percent married and percent living alone are explained by the five PCs (Table 3).

Principal component 1 is mainly an economic status component but also had a social component. Neighbourhoods (CTs) with high values of this PC had high proportions of aboriginal people, low median income, high percentage of low-income earners, relatively high percentage of individuals receiving government transfer income, high unemployment rates, low dwelling values, a high percentage of single-parent families, few owner-occupied dwellings, few married people, high percentage of persons living alone and high population density (Figure 1 and Table 3). These neighbourhoods could be described as high risk neighbourhoods because they had high values of most of the undesirable socioeconomic determinants of health.

Principal component 2 was generally a demographic component. Neighbourhoods with high values of PC 2 had high percentages of children less than 20 years of age, but low percentages of seniors (Figure 2 and Table 3). Neighbourhoods with high values of PC 3 had high percentages of population with less than grade 9 education but low percentages of internal migrants (Table 3 and Figure 3). Finally, neighbourhoods that had high values of the fourth PC, an immigration and aboriginal status component, had high percentage of new immigrants but low population of people of aboriginal origin (Figure 4 and Table 3). The last PC (PC 5) did not load highly on any of the variables but was extracted because its eigenvalue was slightly higher than 1 (Table 3).

Cluster analysis
Figure 5 shows the geographical distribution of the identified neighbourhood types (clusters). The detailed descriptive statistics of the socioeconomic and demo-
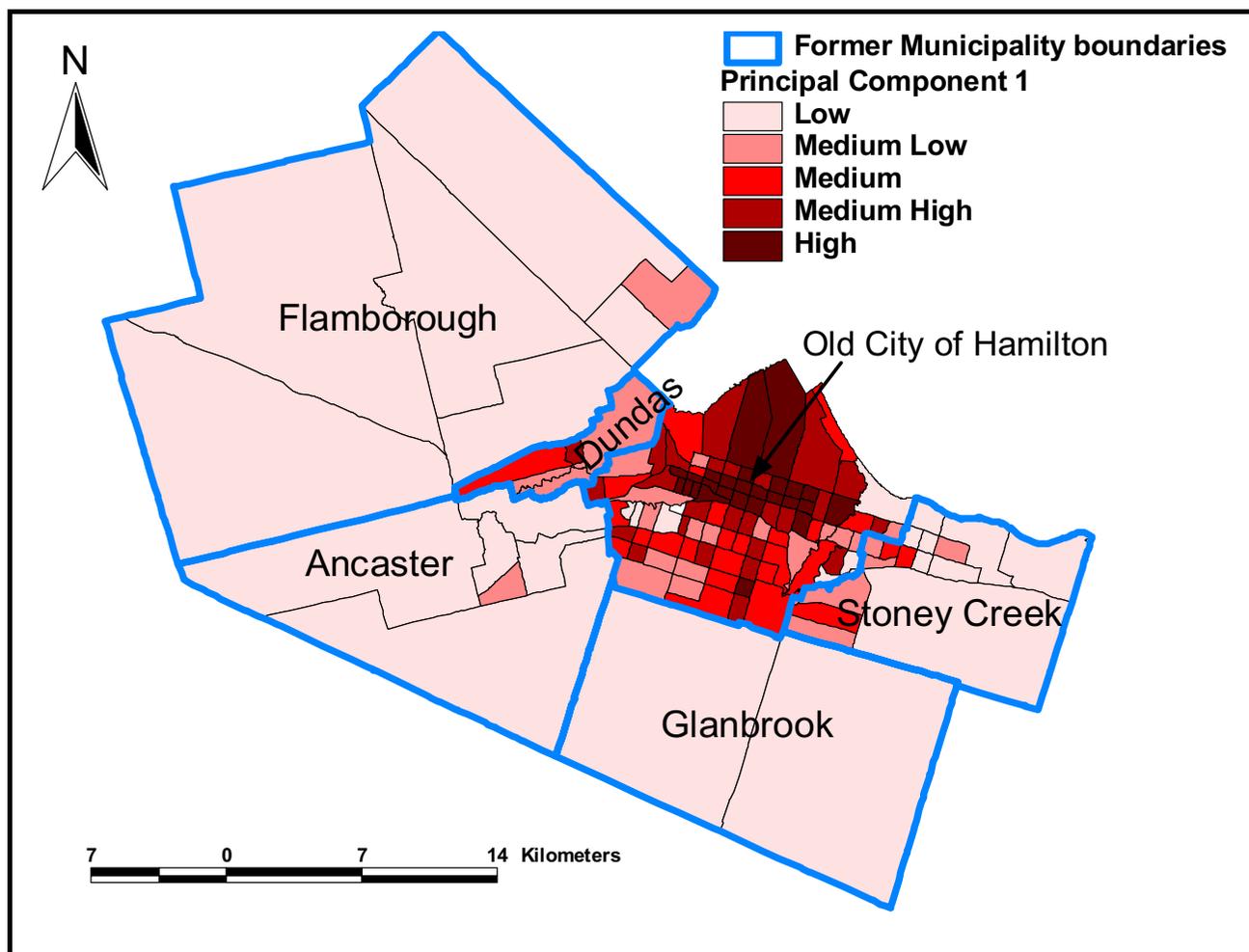
Figure 1
Principal component 1. Spatial distribution of first principal component extracted in the principal components analysis of socio-economic factors in Hamilton, Ontario, Canada (2004)

graphic factors for each of the neighbourhood types compared to the entire city of Hamilton is presented in Table 4. Statistical significance tests were performed to compare the characteristics of each of the neighbourhood types with the Hamilton average. Measures that were significantly different from the Hamilton average are described as high or low otherwise they are described as medium.

Neighbourhood Type A is primarily located within an inner ring surrounding the downtown core. These neighbourhoods can be described as "mature" areas (i.e., many seniors) with some indications of transition (i.e., neighbourhood turnover with arrival of new immigrants). Neighbourhood type A consists of 46 census tracts. A high population density, percentage of seniors and low-

income earners, and a medium percentage of new immigrants characterize this neighbourhood type. It also has a high percentage of single-parent families, low dwelling values and a medium percentage of persons not able to speak English or French. Approximately 32.7% of Hamilton residents live in this neighbourhood type.

Neighbourhood Type B includes high economic status neighbourhoods in low-density rural or suburban environments. Since it covers the largest geographical area, its low population density has a great impact on the overall population density of Hamilton. These areas constitute the geographic periphery of the city, forming an outer ring from east to west. Neighbourhood type B is comprised of 24 census tracts and has approximately 23% of Hamilton
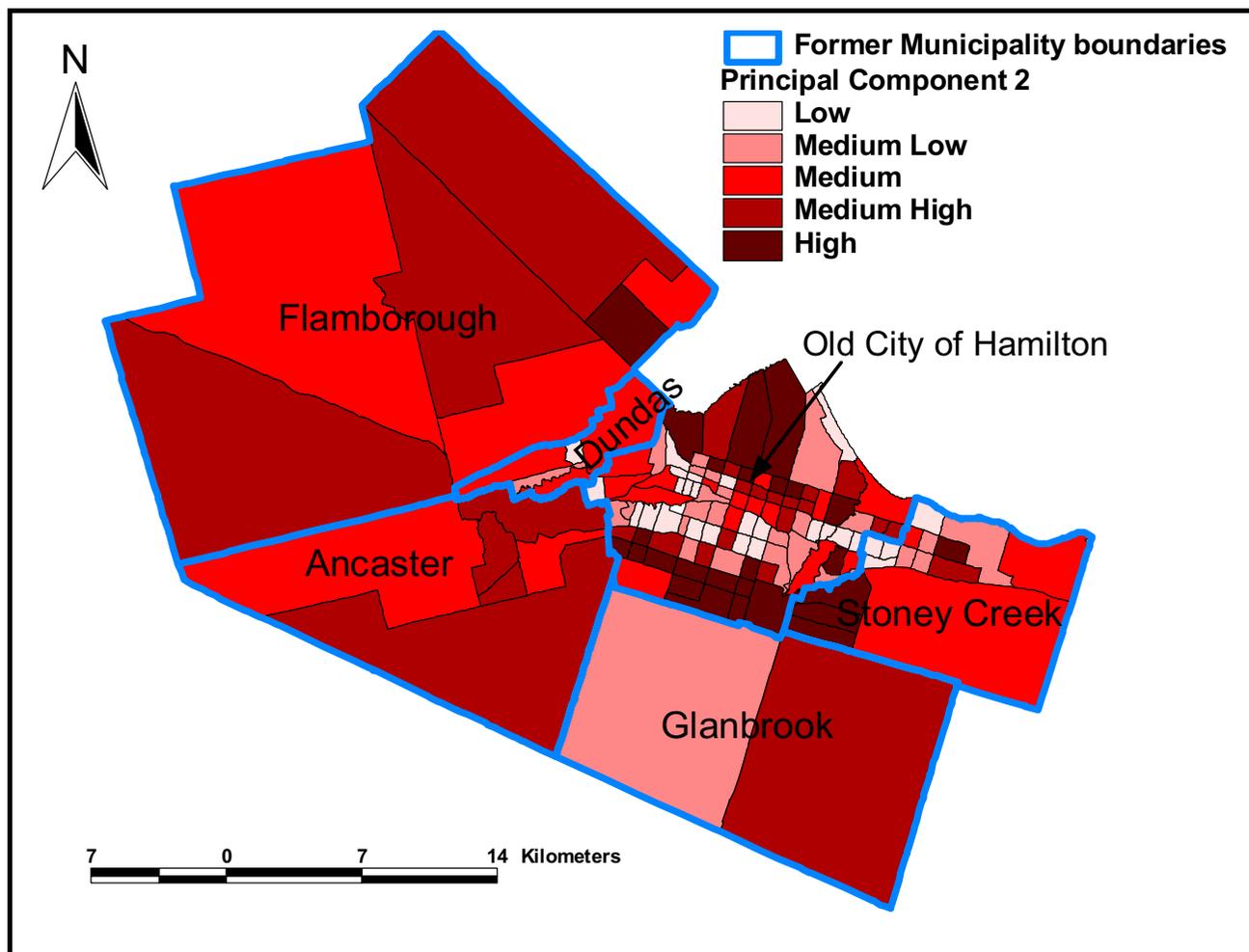
Figure 2
Principal component 2. Spatial distribution of second principal component extracted in the principal components analysis of socio-economic factors in Hamilton, Ontario, Canada (2004)

residents. It has high median income, dwelling values, and percentage of owner-occupied dwellings and a medium percentage of seniors. This neighbourhood type also has a low percentage of new immigrants, single-parent families, and individuals with less than grade 9 education.

Neighbourhood Type C represents a relatively high economic status neighbourhoods within a more urban environment. Unlike Neighbourhood Type B, an area similar in income and dwelling value levels, this NT has a relatively high percentage of visible minority groups. It consists of 17 census tracts and is characterized by high median income, and dwelling values. In addition, this NT has few seniors and persons living alone, but a medium percentage of individuals who cannot speak either English

or French. It also has high dwelling values and population density. Approximately 16.9% of Hamilton residents live in this type of neighbourhood.

Neighbourhood Type D depicts a "mature" urban area with a high percentage of seniors and owned dwellings. It has a relatively low percentage of low-income earners and a high percentage of individuals with less than grade 9 education. In addition, this neighbourhood type has few recent immigrants and internal migrants, low unemployment rate, high percentage of owner-occupied dwellings and medium percentage of persons not able to speak English or French. It is composed of 20 census tracts and 9.6% of Hamilton residents live in this type of neighbourhood.
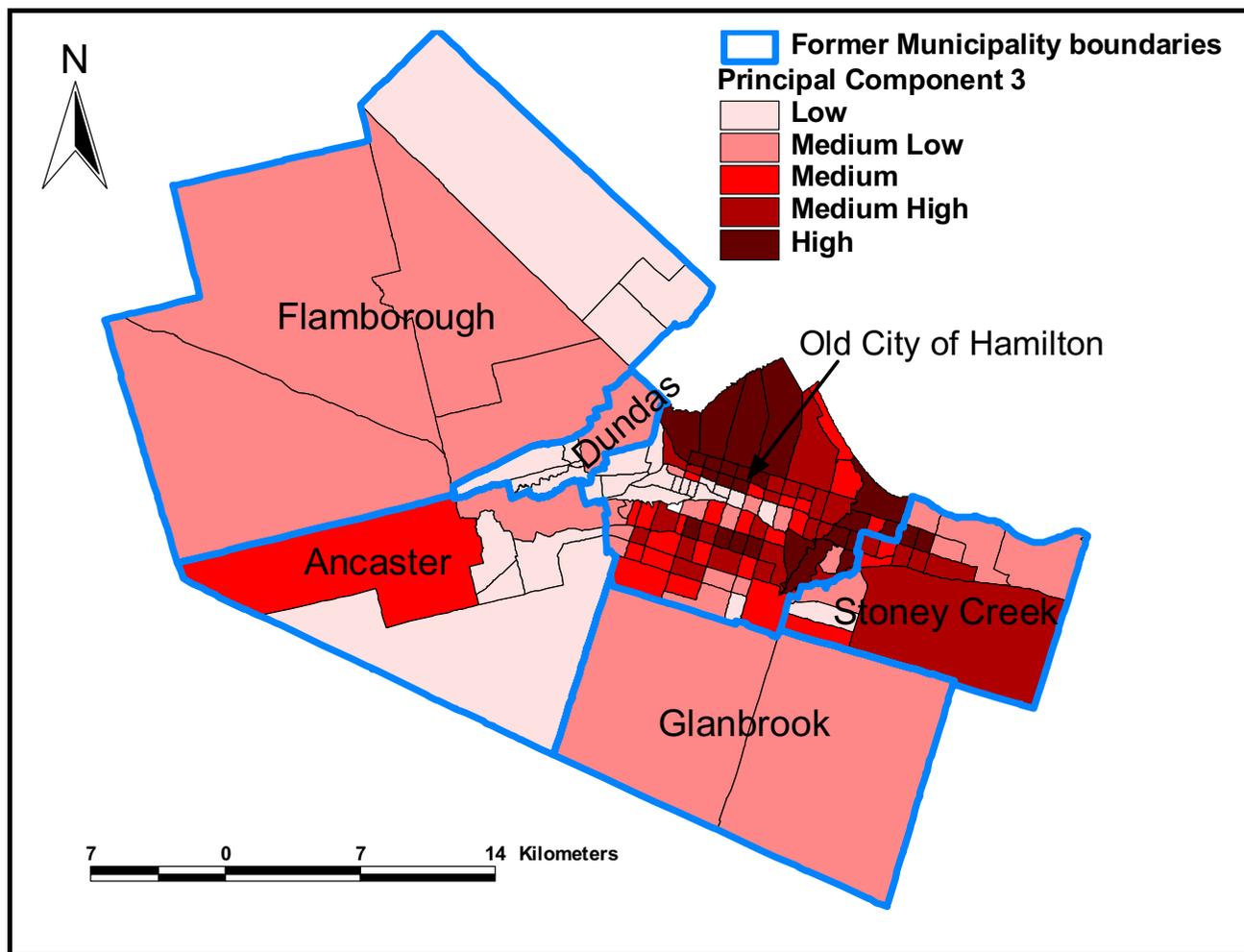
Figure 3
Principal component 3. Spatial distribution of third principal component extracted in the principal components analysis of socio-economic factors in Hamilton, Ontario, Canada (2004)

Neighbourhood Type E constitutes the inner city core and a few areas scattered in the inner ring around the core, and is comprised of 24 census tracts. It has a high prevalence of low-income earners, new immigrants, visible minority groups, and persons with less than grade 9 education. It also has many single-parent families, those receiving government transfer income and high unemployment rate. Approximately 17.7% of Hamiltonians live in these neighbourhoods. Note that the sum of the population percentages of the groups is not 100% because one census tract was not included in the analysis because of missing data.

Discussion

This study has used multivariate techniques to characterize neighbourhoods based on differences and/or similarities of their socioeconomic and demographic characteristics. The positive correlation between single-parenthood and low-income is consistent with observations from other studies that single-parents generally tend to spend more time in low-income neighbourhoods compared to childless couples and unattached individuals [63]. Moreover, it has also been reported that single-parenthood is common among socially disadvantaged groups and compounds social disadvantage [64]. In low socioeconomic neighbourhoods, people experience barriers in creating and benefiting from social capital, leading to social exclusion. The societal costs of social exclusion
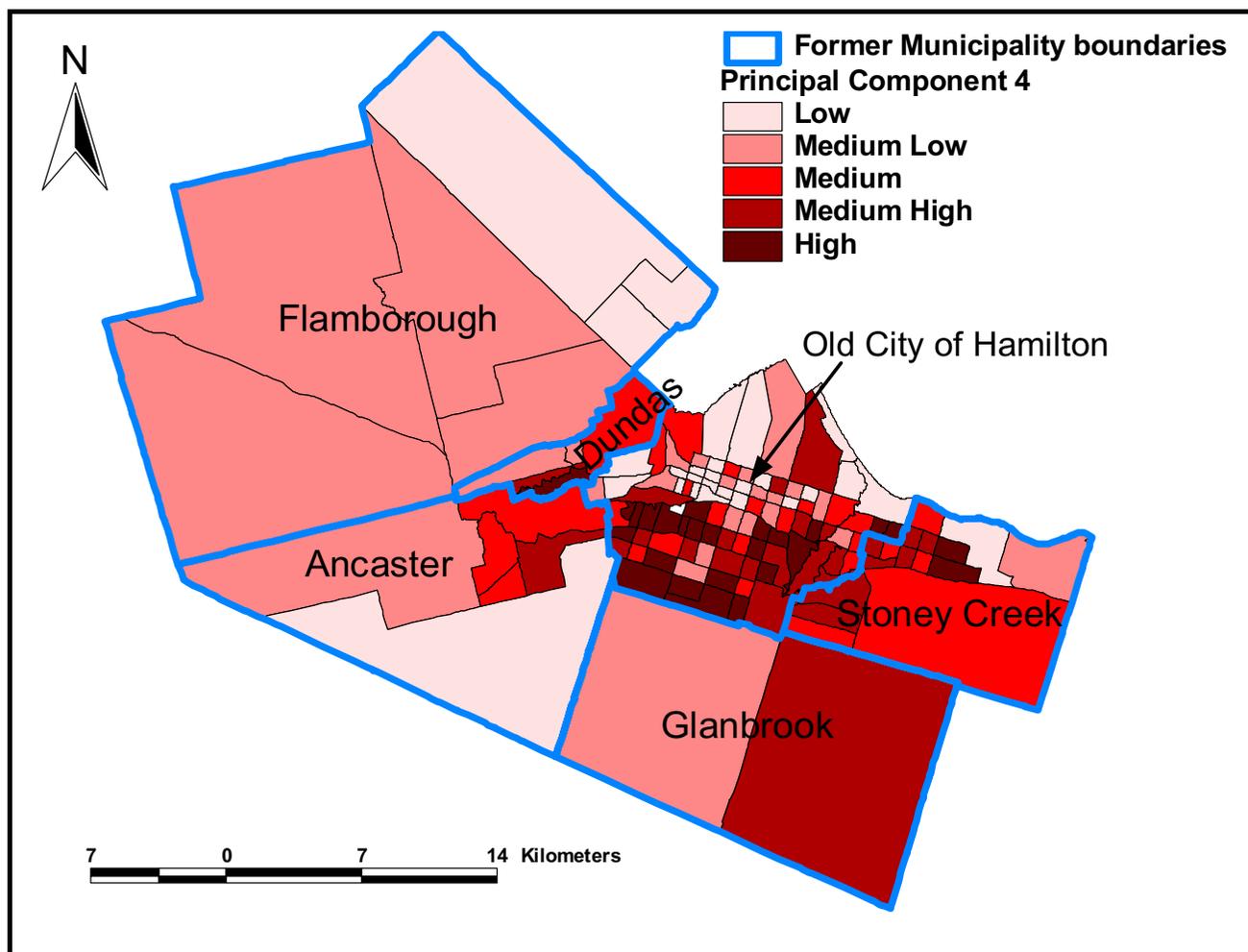
Figure 4
Principal component 4. Spatial distribution of fourth principal component extracted in the principal components analysis of socio-economic factors in Hamilton, Ontario, Canada (2004)

are lack of cohesion, higher crime rates, increased pressure on societal services and the stigma associated with particular neighbourhoods. Social exclusion is especially a problem in neighbourhoods with high unemployment rates, low-incomes, poor housing, etc., all of which combine to create a vicious cycle of poverty, low social capital and increased health risks [65].

The negative correlation between housing ownership and visible minority has been reported in other studies [66]. In addition, the observed positive correlation between visible minority and low-income has also been reported in other Canadian studies which reported that visible minority Canadians (people of colour) experience persistent income gap, above average levels of living on low-

income and higher levels of unemployment [67]. The high negative correlation of low-income and housing ownership is not surprising and is in agreement with observations by Anderson and co-workers [68] who noted inadequate supply of affordable housing for low-income families and the increasing spatial segregation of some households by income, race, ethnicity, or social class into "unsafe neighbourhoods". Moreover, when affordable housing is not available to low-income households, family resources needed for food, medical or dental care, and other necessities are diverted to housing costs leading to the concept of "concentrated poverty" in certain neighbourhoods [68].
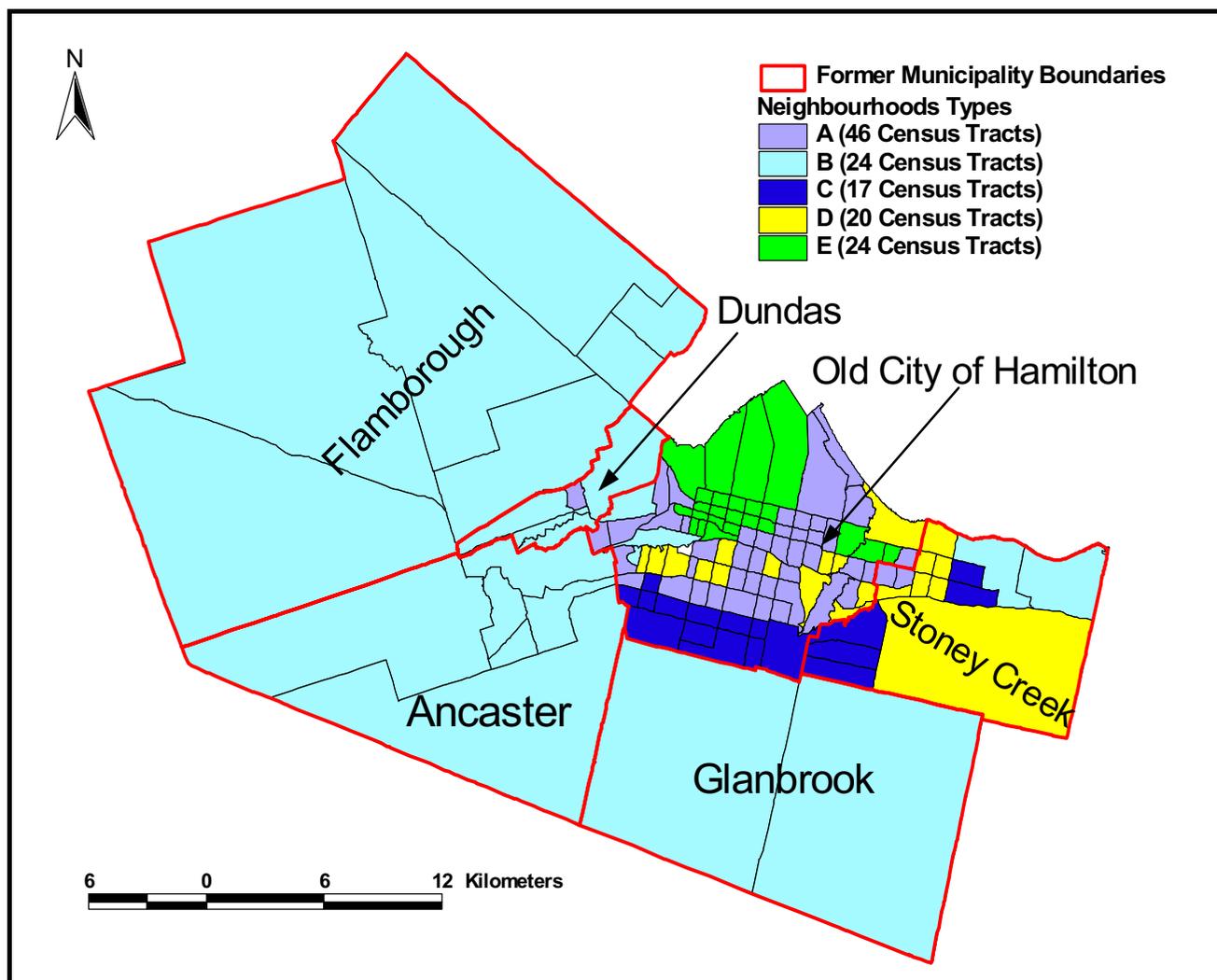
Figure 5
Neighbourhood types. Spatial distribution of identified neighbourhood types in Hamilton Ontario, Canada (2004)

The low uniqueness values of the PCA imply that the five PCs appropriately represent the socioeconomic and demographic variables included in the analysis. Uniqueness values higher than 0.6 are considered high [69]. The advantage of using either PCA or cluster analysis in this kind of study is that they allow incorporation of many variables in the characterization of neighbourhoods. Therefore, from a population health planning perspective, they provide a better understanding of neighbourhood characteristics compared to representations based on only one variable. This is because the health of a population is determined by several socioeconomic, demographic and health care service factors and therefore analyses that

incorporate only one variable would provide insufficient information for population health planning purposes.

Choice of the unit of analysis is critical in these kinds of analyses due to the modifiable areal unit problem (MAUP) since choice of a different and/or inappropriate unit could lead to quite different results [70,71]. As has been pointed out by Ross and coworkers (2004) [72], it is more meaningful to use 'naturally' defined neighbourhoods, rather than arbitrary geostatistical or political units since the distribution of population characteristics or health outcomes may not necessarily follow these arbitrary/political boundaries. Ross and co-workers compared

the performance of census tracts to more 'natural' neighbourhoods and found very similar results and concluded that census tracts, used as proxies of neighbourhoods in our study, are good proxies for natural neighbourhood boundaries [72].

In this study, the results of the PCA were generally similar to those of cluster analysis since the distribution of areas identified as high risk by PC1 tended to follow similar spatial patterns as the high risk areas identified by cluster analysis. Both methods are therefore useful in identifying neighbourhood socioeconomic characteristics that would enhance health planning. However, as has been pointed out by Luginaah and co-workers [49], interpretation of the results of PCA is difficult due to its subjective nature and the fact that as many maps as number of principal components have to be produced. This makes cluster analysis methodology better for these purposes. Moreover, cluster analysis allows computation of statistics for each of the clusters (neighbourhood types) making the methodology much more objective than PCA.

Similar to the pattern seen in other industrial North American cities [28,73-75] most of the high risk neighbourhoods in this study (i.e. with high percentage of low-income earners, low educational attainment, etc), were located in the downtown core with the risk decreasing towards the suburban environments. The observed diverse neighbourhood socioeconomic characteristics may imply great variability in the health needs of the different population subgroups living in the different neighbourhoods since the conditions in which people live strongly influence their health. Health inequalities are produced by the clustering of several of these socioeconomic risk factors [76]. Therefore, populations living in different neighbourhood types differ in the type and number of socioeconomic risk factors to which they are exposed [77,78]. Although it is obvious that neighbourhood type E has the lowest socioeconomic status and highest risk while neighbourhood type B has the highest status and lowest risk, the intent of this study was not merely to classify the neighbourhoods based on economic status. Rather, this study was intended to generate neighbourhood socioeconomic information on which needs-based health planning and service delivery can be based. There is benefit in targeting improvement strategies to materially and socially deprived groups [79].

Future directions and potential applications
The current study is the first of a series of projects designed to investigate neighbourhood health inequalities and provide information to foster health planning with a view to reducing health inequities. The identified neighbourhood clusters will be used, in subsequent studies, as units of analyses in investigating equity in health status, access

and utilization of health services. Additionally, the identified of neighbourhood characteristics are expected to provide useful information on which health planning decisions will be based in order to:

1) Identify population health needs at the neighbourhood level

2) Assess health service utilization patterns across neighbourhoods and compare these with neighbourhood population characteristics and needs

3) Create geographic boundaries for the integrated delivery of social and community health care services

4) Allow for the development of strategies tailored and responsive to the unique characteristics and needs of each neighbourhood.

5) Enhance the use of empirical data for local advocacy for marginalized and under-served neighbourhoods and other populations in need.

Incorporation of the differences in neighbourhood socioeconomic characteristics in population health planning decisions such as decisions on funding allocation to community health agencies will help ensure that health planning strategies are best tailored to address the unique needs of each population. This is because a "one-size-fits-all" planning approach is neither efficient nor practical due to the different socioeconomic and demographic characteristics of the different neighbourhoods. It is expected that inclusion of neighbourhood socioeconomic and demographic characteristics in population health planning will provide health planners with more evidence to guide needs-based decisions that would be more appropriate for the socioeconomically diverse neighbourhoods. Therefore, it is hoped that the results of these analyses will be useful in ensuring that planning is tailored to the unique needs of the different neighbourhood population groups. For instance, neighbourhood types A and D have very similar median incomes and therefore if income was the only variable used to characterize the neighbourhoods, they would be treated similarly. However, the rest of the characteristics of these neighbourhood types are different. For example, neighbourhood type D has a much lower percentage of new immigrants, visible minority population and single-parent families than neighbourhood type A. Moreover, there are significantly more owner occupied dwellings in neighbourhood type D than A. The implication is that these neighbourhoods have potentially different challenges and health needs. If only median income was used (as is most often done) to classify the neighbourhoods, the two NTs would inevitably erroneously be treated as similar. Planning strategies based on

Table 4: Summary statistics of socio-economic and demographic features of identified neighbourhood types in Hamilton, 2004. Data source: Statistics Canada, 2001 census. NT = Neighbourhood type. The sum of the populations from the 5 neighbourhood types is not 490,268 due to missing data in one census tract.

Variable	Means					
	Hamilton	NT A	NT B	NT C	NT D	NT E
Persons with <grade 9 education (%)	10.3	11.3	4.7	9.2	13.7	16.9
New immigrants (%)	3.2	3.4	0.9	2.7	1.2	6.6
Visible minority (%)	10.9	9.4	4.2	14.4	5.6	20.7
Aboriginal persons (%)	2.2	2.6	1.0	1.6	3.0	4.1
Median income ($)	22927	22262	30351	25958	22837	16250
Government transfer income (%)	12.1	15.9	7.3	8.3	13.4	23.2
Low income (%)	19.8	21.7	5.6	12.7	10.7	42.4
Persons not speaking English or French (%)	1.8	1.6	0.33	1.6	1.6	4.4
Unemployment rate (%)	6.4	5.7	2.2	4.2	3.5	10.1
Average dwelling value ($)	166783	125271	224342	173801	162157	89174
Owner-occupied dwellings (%)	65.2	61.9	87.4	81.8	86.7	34.7
Population under 20 years old (%)	26.1	23.5	27.5	33.6	21.2	23.6
Population 65 years or older (%)	14.3	17.2	13.2	7.2	21.0	14.3
Single-parent families (%)	16.6	19.9	9.5	15.4	10.7	25.7
Married (%)	51.7	47.3	60.9	58.9	58.5	35.7
Live alone (%)	10.3	13.3	5.4	2.9	8.7	21.1
Internal migrants (%)	10.6	9.2	14.5	7.7	6.9	12.2
Population Density (No. of persons per Km^2)	438.9	3627.4	357.1	2543.2	2352.1	5706.0
Population (count)	490268	160438	112875	82976	47137	86683

such single variable analysis may not be appropriate since the strategies would not be tailored to the unique characteristics and therefore needs of the NTs.

Conclusion

In this study, we have used multivariate techniques to identify unique neighbourhood characteristics and classify the neighbourhoods into groups with similar characteristics. Since the identified neighbourhood types are homogeneous with respect to the broad determinants of health, they offer potentially excellent opportunities for health planners and service providers to understand the characteristics and potential health needs of the different neighbourhoods and therefore better plan for them. Through continuous monitoring of health information across these neighbourhoods, health planners, service providers and policy makers could better make decisions based on knowledge of the local communities.

Authors' contributions

AO was involved in the study design, execution and writing up of the draft and final copies of the manuscript. RW conceived the need for the study, was involved in the study design and preparation of the manuscript. ME, SB, BH, JE, and TA participated in the study design, guiding implementation and preparation of the draft of the manuscript.

References

1. Kindig D, Stoddart G: **What is population health?** *Am J Public Health* 2003, **93**:380-383.
2. Yu IT, Wong TW, Wong SL: **Small area variations of cancer mortality in Hong Kong – the roles of health care and socio-economic status.** *Neoplasma* 2004, **51**:144-149.
3. Singh GK, Miller BA: **Health, life expectancy, and mortality patterns among immigrant populations in the United States.** *Can J Public Health* 2004, **95**:114-121.
4. Paulander J, Axelsson P, Lindhe J: **Association between level of education and oral health status in 35-, 50-, 65- and 75-year-olds.** *J Clin Periodontol* 2003, **30**:697-704.
5. Epstein AM: **Health care in America – still too separate, not yet equal.** *N Engl J Med* 2004, **351**:603-605.
6. Hnizdo E, Sullivan PA, Bang KM, Wagner G: **Association between chronic obstructive pulmonary disease and employment by industry and occupation in the US population: a study of data from the Third National Health and Nutrition Examination Survey.** *Am J Epidemiol* 2002, **156**:738-746.
7. Bartosinska M, Ejsmont J: **[Health condition of employees exposed to noise – extra auditory health effects] Stan zdrowia pracownikow narazonych na halas – skutki zdrowotne pozasluchowe.** *Wiad Lek* 2002, **55**:20-25.
8. Blau G, Tatum DS, McCoy K, Dobria L, Ward-Cook K: **Job loss, human capital job feature, and work condition job feature as distinct job insecurity constructs.** *J Allied Health* 2004, **33**:31-41.
9. Garbarino S, De Carli F, Mascialino B, Beelke M, Nobili L, Squarcia S, Penco MA, Ferrillo F: **Sleepiness in a population of Italian shift-work policemen.** *J Hum Ergol (Tokyo)* 2001, **30**:211-216.
10. Shields M: **Shift work and health.** *Health Rep* 2002, **13**:11-33.

11. Dearry A: **Impacts of our built environment on public health.** *Environ Health Perspect* 2004, **112**:A600-A601.
12. Kasmel A, Helasoja V, Lipand A, Prattala R, Klumbiene J, Pudule I: **Association between health behaviour and self-reported health in Estonia, Finland, Latvia and Lithuania.** *Eur J Public Health* 2004, **14**:32-36.
13. Ndubani P, Bond V, Liljestrom R, Hojer B: **Understanding young men's sexual health and prospects for sexual behaviour change in rural Zambia.** *Scand J Public Health* 2003, **31**:291-296.
14. Cree MW, Juby AG, Carriere KC: **Effect of home care service levels on health outcomes in hip fracture patients in Alberta.** *Health* 2004, **7**:49-53.
15. Gerrard N, Kulig J, Nowatzki N: **What doesn't kill you makes you stronger: determinants of stress resiliency in rural people of Saskatchewan, Canada.** *J Rural Health* 2004, **20**:59-66.
16. Gil KM, Anthony KK, Carson JW, Redding-Lallinger R, Daeschner CW, Ware RE: **Daily coping practice predicts treatment effects in children with sickle cell disease.** *J Pediatr Psychol* 2001, **26**:163-173.
17. Ellaway A, Macintyre S: **Does where you live predict health related behaviours?: a case study in Glasgow.** *Health Bull* 1996, **54**:443-446.
18. Macintyre S, Ellaway A: **Neighbourhood cohesion and health in socially contrasting neighbourhoods: implications for the social exclusion and public health agendas.** *Health Bull* 2000, **58**:450-456.
19. Krueger PM, Bond Huie SA, Rogers RG, Hummer RA: **Neighbourhoods and homicide mortality: an analysis of race/ethnic differences.** *J Epidemiol Community Health* 2004, **58**:223-230.
20. Macintyre S, Maciver S, Sooman A: **Area, class and health: should we be focussing on places or people?** *Journal of Social Policy* 1993, **22**:213-234.
21. Massey DS: **American Apartheid: Segregation and the Making of the Underclass.** *The American Journal of Sociology* 1990, **96**:329-357.
22. Massey DS, Eggers ML: **The Ecology of Inequality: Minorities and the Concentration of Poverty, 1970–1980.** *The American Journal of Sociology* 1990, **95**:1153-1188.
23. Massey DS, Denton NA: **Suburbanization and Segregation in U.S. Metropolitan Areas.** *The American Journal of Sociology* 1988, **94**:592-626.
24. Myles J, Picot G, Pyper W: **Neighbourhood inequality in Canadian cities.** Ottawa, Ontario, Canada: Statistics Canada, Analytical Studies Branch; 2000.
25. Kirby JB, Kaneda T: **Neighborhood socioeconomic disadvantage and access to health care.** *J Health Soc Behav* 2005, **46**:15-31.
26. Marmot M: **Social determinants of health inequalities.** *Lancet* 2005, **365**:1099-1104.
27. Arntzen A, Nybo Andersen AM: **Social determinants for infant mortality in the Nordic countries, 1980–2001.** *Scand J Public Health* 2004, **32**:381-389.
28. Jargowsky PA: **Poverty and Place: Ghettos, Barrios and the American city.** New York, USA: Rusell Sage Foundation; 1997.
29. Denton NA, Massey DS: **Patterns of neighborhood transition in a multiethnic world: U.S. metropolitan areas, 1970–1980.** *Demography* 1991, **28**:41-63.
30. Massey DS, Denton NA: **Hypersegregation in U.S. metropolitan areas: black and Hispanic segregation along five dimensions.** *Demography* 1989, **26**:373-391.
31. Massey DS: **The age of extremes: concentrated affluence and poverty in the twenty-first century.** *Demography* 1996, **33**:395-412.
32. MacLachlan I, Sawada R: **Measures of inequality and social polarization in Canadian metropolitan areas.** *The Canadian Geographer* 1997, **41**:377-397.
33. Lee KK: **Urban Poverty in Canada.** Ottawa, Ontario, Canada: Canadian Council on Social Development; 2000.
34. Henderson C, Diez Roux AV, Jacobs DRJ, Kiefe CI, West D, Williams DR: **Neighbourhood characteristics, individual level socioeconomic factors, and depressive symptoms in young adults: the CARDIA study.** *J Epidemiol Community Health* 2005, **59**:322-328.
35. Stafford M, Cummins S, Macintyre S, Ellaway A, Marmot M: **Gender differences in the associations between health and neighbourhood environment.** *Soc Sci Med* 2005, **60**:1681-1692.
36. Drukker M, Kaplan C, van Os J: **Residential instability in socioeconomically deprived neighbourhoods, good or bad?** *Health Place* 2005, **11**:121-129.
37. Diez-Roux AV, Borrell LN, Haan M, Jackson SA, Schultz R: **Neighbourhood environments and mortality in an elderly cohort: results from the cardiovascular health study.** *J Epidemiol Community Health* 2004, **58**:917-923.
38. Stafford M, Martikainen P, Lahelma E, Marmot M: **Neighbourhoods and self rated health: a comparison of public sector employees in London and Helsinki.** *J Epidemiol Community Health* 2004, **58**:772-778.
39. Schneiders J, Drukker M, van der Ende J, Verhulst FC, van Os J, Nicolson NA: **Neighbourhood socioeconomic disadvantage and behavioural problems from late childhood into early adolescence.** *J Epidemiol Community Health* 2003, **57**:699-703.
40. Stafford M, Marmot M: **Neighbourhood deprivation and health: does it affect us all equally?** *Int J Epidemiol* 2003, **32**:357-366.
41. Winkleby MA, Cubbin C: **Influence of individual and neighbourhood socioeconomic status on mortality among black, Mexican-American, and white women and men in the United States.** *J Epidemiol Community Health* 2003, **57**:444-452.
42. Sundquist J, Malmstrom M, Johansson SE: **Cardiovascular risk factors and the neighbourhood environment: a multilevel analysis.** *Int J Epidemiol* 1999, **28**:841-845.
43. Blake BJ, Bentov L: **Geographical mapping of unmarried teen births and selected sociodemographic variables.** *Public Health Nurs* 2001, **18**:33-39.
44. Pettit KLS, Kingsley GT, Coulton CJ: **Neighborhoods and Health: Building Evidence for Local Policy.** Washington DC, USA: The Urban Institute submitted to The Office of the Assistant Secretary for Planning and Evaluation U.S. Department of Health and Human Services; 2003.
45. Odoi A, Martin SW, Michel P, Middleton D, Holt J, Wilson J: **Investigation of clusters of giardiasis using GIS and a spatial scan statistic.** *Int J Health Geogr* 2004, **3**:11.
46. Lorant V, Thomas I, Deliege D, Tonglet R: **Deprivation and mortality: the implications of spatial autocorrelation for health resources allocation.** *Soc Sci Med* 2001, **53**:1711-9.
47. Malmstrom M, Sundquist J, Johansson SE: **Neighborhood environment and selfreported health status: a multilevel analysis.** *Am J Public Health* 1999, **89**:1181-6.
48. Pampalon R, Raymond G: **A deprivation index for health and welfare planning in Quebec.** *Chronic Dis Can* 2000, **21**:104-13.
49. Luginaah I, Jerrett M, Elliott S, Eyles J, Parizeau K, Birch S, Abernathy T, Veenstra G, Hutchinson B, Giovis C: **Health profiles of Hamilton: Spatial characterisation of neighbourhoods for health investigations.** *GeoJournal* 2001, **53**:135-147.
50. Diez-Roux AV, Kiefe CI, Jacobs DR Jr, Haan M, Jackson SA, Nieto FJ, Paton CC, Schulz R: **Area characteristics and individual-level socioeconomic position indicators in three population-based epidemiologic studies.** *Ann Epidemiol* 2001, **11**:395-405.
51. Hamilton District Health Council: **Hamilton Health System Fact Book.** Hamilton, Ontario, Canada: Hamilton District Health Council; 2004.
52. Statistics Canada: **2001 Census Dictionary.** Ottawa, Ontario, Canada. Ministry of Industry; 2003.
53. Frenette M, Picot G, Sceviour R: **How long do people live in low-income neighbourhoods? Evidence for Toronto, Montreal and Vancouver.** Ottawa, Canada: Statistics Canada; 2004.
54. **Population Health Approach: What Determines Health?** [http://www.phacaspc.gc.ca/ph-sp/phdd/determinants/index.html#What]
55. Statistics Canada: **Health Region Peer Groups.** Ottawa, Ontario, Canada. Statistics Canada; 2002.
56. Stata Corporation: **STATA Version 8.2.** 4905 Lakeway Drive, College Station, Texas, USA; 2004.
57. Everitt BS, Landau S, Leese M: **Cluster Analysis.** Oxford, United Kingdom: Oxford University Press; 2001.
58. Norman GR, Streiner DL: **Biostatistics: The Bare Essentials.** B C Decker Inc; 2000.
59. Dunteman GH: **Principal components analysis.** In *Quantitative Applications in the Social Sciences Series, No. 69* Thousand Oaks, California, USA: Sage Publications; 1989.
60. Kaiser HF: **The varimax criterion for analytic rotation in factor analysis.** *Psychometrika* 1958, **23**:187-200.

61. Caliński T, Harabasz J: **A dendrite method for cluster analysis.** *Communications in Statistics* 1974, **3:**1-27.
62. ESRI: **ArcView GIS. Version 3.2.** Redlands, California, USA: Environmental Systems Research Institute Inc; 1999.
63. Frenette M, Picot G, Sceviour R: **How long do people live in low-income neighbourhoods? Evidence for Toronto, Montreal and Vancouver.** Ottawa, Canada: Statistics Canada; 2004.
64. Martin SP: **Delayed marriage and childbearing: implications and measurement of diverging trends in family timing.** Maryland, USA: Department of Sociology and Maryland Population Research Center, University of Maryland, College Park; 2002.
65. **What is Social Exclusion?** [http://www.socialexclusion.gov.uk/page.asp?id=213]
66. Krivo LJ, Kaufman RL: **Housing and wealth inequality: racial-ethnic differences in home equity in the United States.** *Demography* 2004, **41:**585-605.
67. Galabuzi GE: **Canada's creeping economic apartheid.** Toronto, Ontario, Canada: CSJ Foundation for Research and Education; 2001.
68. Anderson LM, Charles JS, Fullilove MT, Scrimshaw SC, Fielding JE, Normand J: **Providing affordable family housing and reducing residential segregation by income. A systematic review.** *Am J Prev Med* 2003, **24:**47-67.
69. Stata Corporation: **Cluster Analysis.** College Station, Texas, USA: Stata Press; 2003.
70. Openshaw: **The modifiable areal unit problem. CATMOG No. 38.** Norwich: Geo Books; 1984.
71. Steel GG, Holt D: **Rules for random aggregation.** *Environ Planning A* 1996, **28:**957-978.
72. Ross NA, Tremblay SS, Graham K: **Neighbourhood influences on health in Montréal, Canada.** *Soc Sci Med* 2004, **59:**1485-1484.
73. Allard SW: **Access to social services: The Changing Urban Geography of Poverty and Service Provision.** Washington DC, USA: The Brookings Institute; 2004.
74. Wilson WJ: **The Truly Disadvantaged.** Chicago, USA: University of Chicago Press; 1987.
75. Wilson WJ: **When Work Disappears: The World of the New Urban Poor.** New York, USA: Alfred A Knopf; 1996.
76. Shaw M, Dorling D, Gordon D, Davey SG: **The widening gap: health inequalities and policy in Britain.** Bristol, United Kingdom: The Policy Press; 1999.
77. Shaw M, Dorling D, Davey SG: **Poverty, social exclusion and minorities.** In *Social Determinants of Health edition* Edited by: Marmot MG, Wilkinson RG. Oxford: Oxford University Press; 1999.
78. Kunst AE, Groenhof F, Mackenbach JP, Health EW: **Occupational class and cause specific mortality in middle aged men in 11 European countries: comparison of population based studies. EU Working Group on Socioeconomic Inequalities in Health.** *BMJ* 1998, **316:**1636-1642.
79. Martinez J, Pampalon R, Hamel D: **Deprivation and stroke mortality in Quebec. Chronic Diseases in Canada.** Ottawa, Ontario, Canada: Population and Public Health Branch, Health Canada; 2003.