

This is the peer reviewed version of the following article:

Apparicio, Philippe, Marie-Soleil Cloutier, Véronique Chadillon-Farinacci, Johanne Charbonneau, and Gilles Delage. 2014. "Blood donation clusters in Québec, Canada (2003–2008): spatial variations according to sex and age." *Vox Sanguinis* 106 (4):297-306. doi: 10.1111/vox.12082.

which has been published in final form at <http://dx.doi.org/10.1111/vox.12082>. This article may be used for non-commercial purposes in accordance with [Wiley Terms and Conditions for Self-Archiving](#).

BLOOD DONATION CLUSTERS IN QUÉBEC, CANADA (2003-2008): SPATIAL VARIATIONS ACCORDING TO SEX AND AGE

Philippe Apparicio, Marie-Soleil Cloutier, Véronique Chadillon-Farinacci, Johanne Charbonneau, Gilles Delage

ABSTRACT

Background and objectives The detection of spatial clusters of blood donation rate is an important issue, especially for targeting spatial units with significantly low rates, where it could be possible to increase the numbers of donors. The objective of this study is to detect spatial clusters of high or low blood donation rate in Québec according to sex and age of the donors.

Materials and Methods Blood donation data were obtained from Héma-Québec over a period of five years. We aggregated these data for each of 101 *municipalités régionales de comté* (i.e. counties) for men, women and four age groups. To detect spatial high/low donation rate areas, we used the Kulldorff's scan statistics. Kappa coefficient was used to assess discordance between clusters obtained for the different groups (18-29, 30-39, 40-49, 50-59, 60-69 years old). T-Test analyses were conducted to identify significant associations between spatial clusters and socio-economic variables.

Results The results indicate the presence of several geographical areas with high or low blood donation rates for each group. The size, the location and the sociodemographic profiles of low/high clusters vary according to sex and age categories.

Conclusion The Kulldorff's scan statistics are an efficient tool to assess the blood donation performance across a country or even a specific region over a period of several years. In terms of strategic planning and monitoring, it can be used as a fully operational tool to target areas with significantly low rates (for all donors or specific demographic groups) in future blood donation campaigns.

Key words: blood donation, demography, spatial clustering, geographical information systems, Quebec.

INTRODUCTION

In Canada, as elsewhere around the globe, maintaining the supply of blood products is a crucial issue for the two blood donation agencies (Canadian Blood Services and Héma-Québec). For Héma-Québec, as well as other blood agencies in developed countries, the “blood donor loyalty and retention are major challenges, which must also cope with a stagnating number of donations per person” [1]. Following these preoccupations, it is hardly surprising that a particular attention is given worldwide to analyses on social, ethnic and demographic characteristics of first-time or repeat donors, and on factors influencing the decision to donate blood [2-8]. In addition, some authors focus on the forecast of demand and supply by using projected population data, in a context of ageing population and increase in blood products demand [9, 10].

By contrast, there have been few studies exploring the geographic aspects of blood donation. More specifically, spatial variations of blood donation rates and local factors impacting these rates have received little attention during the last decades. To fill this gap, some interesting studies, which integrate a geographic dimension, have recently been conducted in different contexts, e.g. Canada [11-13], Netherlands [14] and Germany [15]. In methodological terms, the exercise consists in the joint use of blood donation data provided by a blood agency and several variables extracted from a national census. By doing so, it is possible to aggregate the blood donations data at a given spatial level (e.g. municipalities, counties, or regions) and then calculate and map blood donation rates for a specific period of time. Starting from these joined datasets, the use of classic or spatial regression models let one to explore the relationships between rates and several predictors such as demographic, socio-economic, ethnic/racial/religious, and social capital (e.g. voter turnout) indicators. These studies have shown that, all things being equal, percentage of immigrants and unemployment rate are generally associated with lower donation rates while percentages of university graduates and voter turnout are associated with higher donation rates [11, 12, 14, 15]. For example, Bekkers et Veldhuizen [14] analysed the geographic variations of proportion of donors among the total population across municipalities of Netherlands in 2005 and Weidmann et al. [15] calculated blood donation rates for 1533 municipalities in south-west Germany in 2007. In Canada, Saberton et al. [11] calculated donation rates at the census tract level for 40 metropolitan areas across Canada whereas Cimaroli et al. [12] produced an interesting map of the percent of total donors in Toronto at a fine spatial level (i.e. dissemination areas, equivalent to the census block group in the US). Finally, Cloutier et al. [13] presented a map of blood donor rates for the administrative regions of Québec. All these previous works [11-15] have demonstrated that blood donation rate varies significantly across space, no matter the geographical scale retained. Moreover, two studies have demonstrated that this indicator is positively spatially autocorrelated [11, 15]. In practical terms, this means that in a given study area, close or adjacent spatial units share high values of blood rates whereas other ones present low values.

In this context, low/high blood donation rate spatial clusters detection becomes an interesting issue to explore, especially if these mapped clusters and the spatial units underneath are suggested as interventions targets: it makes it possible to try to increase the number of donors in significantly low rate areas, for example. This study's main objective is to detect geographical clusters of high or low blood donation rates in the Québec Province (Canada). We hypothesise that these rates may vary spatially, and according to donors' sex and age. Compared to previous studies, this work is original in two ways. Firstly, it includes a spatial-clustering analysis to identify low/high donation rate areas (i.e. clusters) using blood data over a five-year period (2003-2008). Secondly, it highlights the low/high cluster space variations for different groups (total population, men, women and four age categories) while, in previous studies [11, 12, 14, 15], blood donation rates for a given spatial level were calculated for the entire population and for one year only. It is worth noting that SaTScan (the software used for the spatial cluster analysis) has been employed in epidemiological studies to detect spatial and space-time disease clusters, such as different types of cancer –leukemia [16], breast [17, 18], colon [19], prostate [20], lung [21], and brain [22]–, infectious diseases [23, 24], low birthweight and infant mortality [25]. To date, it has not yet been used to detect space or space-time clusters of blood donation rates (no study is indexed in PubMed with 'SaTScan' and 'blood' as keywords).

MATERIAL AND METHODS

Blood donations, population and geographic data

The data on blood donations were obtained from Héma-Québec, which is the only organization responsible for blood supply in Québec. As described in a previous study [13], only allogeneic donations from people aged between 18 and 69 were retained (age being the first eligibility criteria), including attempted donations (donations not collected due to deferral or another reason) since, even if the blood collection did not take place, the would-be donor generally goes to a collection site with the intention to donate blood. The study period was between September 1st 2003 and August 31st 2008 (five years). The analysis is based on 101 *municipalités régionales de comté* (hereafter MRC – i.e. equivalents of U.S. counties). Population size varies among the MRCs, especially between urban and rural areas (mean population size in 2008 =76 336; median=32 256; standard deviation=198 971). Three spatial units which belong to the Nord-du-Québec administrative region were excluded due to their remoteness, low population density and low number of donations.

For each blood donation, age –calculated from the date of the donation–, gender and residential donor's postal code were available in the database, making it possible to compute the number of cases for the 101 MRCs by using geographic information systems (GIS). To calculate the donation rates, we used demographic data publicly available on the *Institut National de la Statistique du Québec* website at the MRCs spatial level for our five-year study period (2003 to 2008). Once fully structured into ArcGIS 10 and SAS 9.2 software, the database contains 1 428 849 blood donations given by people aged from 18 to 69 over the five-year study period.

Detection of spatial clusters method: Kulldorff's scan statistics

In order to identify areas with significantly low or high rates of blood donations, the SaTScan software (version 9.1.1) [26] was used to perform spatial analysis, assuming that the number of blood donations follows a Poisson distribution. This computer software detects location and size of spatial and space-time clusters while checking their statistical significance. In practical terms, with the spatial scan statistic, a circular or elliptic window scans the study area with different radius (circles) or shapes and angles (ellipses) that will never reach more than half the total population [27, 28]. Thereafter, for each location and size of the scanning window, it will test the alternative hypothesis which states that the annual number of blood donation per 1000 people is highest or lowest within the window compared to outside. According to Kulldorff [28], under the Poisson assumption, the likelihood function compares the observed and expected numbers of blood donations for each scanning window as follows:

$$LR = \left(\frac{d}{E[d]} \right)^d \left(\frac{D-d}{D-E[d]} \right)^{D-d}$$

where D is the total number of blood donations over the whole area, d is the number of blood donations within the scanning window, and $E[d]$ is the expected number of blood donations within the window under the null hypothesis. Without covariate adjustment, this expected number of blood donation is:

$$E[d] = p * D / P$$

where d and D are defined above, p and P are respectively the population within the scanning window and P the total population in the Québec Province. With covariate adjustment (i.e. sex and age groups), $E[d]$ is written as follows:

$$E[d] = \sum_i E[d_i] = \sum_i p_i * D_i / P_i$$

where d , D , p , P are defined for each sex and age category i .

The SaTScan program computes a p-value for each likelihood ratio by using Monte Carlo replications (usually 999 or 9999) [29]. Finally, the Kulldorff's scan statistics allows identifying two types of clusters: the most likely cluster and secondary clusters that are descending-ordered according to their significance. Output result for each cluster contains the number of spatial units (here MRCs) included in the cluster, the number of observed and expected blood donations, the rate inside the cluster (annual blood donations per 1000 inhabitants), the log likelihood ration and its p-value, and the relative risk (RR). The latter statistic is the ratio between "observed divided by the expected within the cluster divided by the observed divided by the expected outside the cluster" [29]:

$$RR = \frac{d / E[d]}{(D-d) / (E[D] - E[d])} = \frac{d / E[d]}{(D-d) / (D - E[d])}$$

Therefore, for a detected cluster, a RR value greater than 1 indicates a high annual blood donation rate area in comparison with the rest of the province; and a value lower than 1, a low annual blood donation rate area.

The analyses were conducted by setting the SaTScan parameters in this following way: a scan for areas with either high or low rates, an ellipse with a medium non-compactness penalty as a spatial window shape, 50% of the total population as the maximum spatial cluster size and 9 999 Monte-Carlo replications to test for statistical significance.

Because we want to verify if MRCs with high or low donation rate's location varies according to sex and age, eight spatial cluster analyses were carried out: 1) all the donors (18-69) adjusted by sex and age category; 2) men; 3) women; 4) 18-29 years-old group; 5) 30-39; 6) 40-59; 7) 50-59; 8) 60-69. For each analysis, three files are needed in SaTScan: a) blood donation count for the specific group; b) population count for the same group for each MRC from 2003 to 2008; c) MRC centroid's geographic coordinates. Concerning this last file, due to the fact that a large part of certain MRC are uninhabited (especially in rural and remote areas), the centroid location was spatially adjusted using the 2006 population ecumene map of Canada produced by Statistics Canada [30]: the original MRC boundaries were clipped to this map before computing the centroid x/y-coordinate values. It should be mentioned that space-time analysis were also performed. Because results were very similar to the purely spatial analysis, they are not reported here.

Statistical Methods

Two main statistical methods will be assessed here: a) a comparison of the obtained clusters between eight sex and age groups previously introduced and b) sociodemographic characteristic identification for high donation rate clusters.

First, to explore the similarity of obtained clusters for each sex and age categories, Cohen's kappa coefficient were calculated. This coefficient is used to assess the degree of concordance between two nominal variables which have identical categories [31]. In our case, the variables have three (possible) categories: the MRC belongs to a high blood donation rate cluster, to a low blood donation rate cluster or the MRC does not belong to any cluster. Resulting clusters for all the donors will be compared with those for men, women and the six age categories: this will permit to evaluate how the low/high clusters obtained for each sex and age category differ from those obtained for the total population. According to the values of the Kappa Statistic, which varies between 0 and 1, the degree of concordance will be determined as follows: poor ($k < 0.2$), fair ($0.2 \leq k < 0.4$), moderate ($0.4 \leq k < 0.6$), good ($0.6 \leq k < 0.8$); excellent ($k \geq 0.8$), and perfect ($k=1$) [31, 32].

Secondly, six factors are retained to explain the MRC's membership to a low or high rate cluster. These "predictors" are quite common in the literature and refer to the blood collection sites spatial accessibility and to socio-economic and demographic characteristics [11-15]. Some of those variables are known to

have a negative relation to donation rates: 1) road network distance between the donor's place of residence and the collection site, 2) population density, 3) unemployment rate, 4) percentage of immigrants; while others have a positive relation to it: 5) average individual total income, and 6) percentage of university graduates. According to the 2006 census dictionary of Statistics Canada a "landed immigrant (permanent resident) is a person who has been granted the right to live in Canada permanently by immigration authorities" [33]. Except the distance predictor, all of them were extracted from the 2006 census of Statistics Canada. Since the number of MRCs is too low to apply a logistic regression ($n=101$), independent group t-test for each variable were computed to compare mean values of MRCs belonging to a high blood donation rate cluster to those not included into any cluster. This analysis was conducted in SAS by using the procedure PROC TTEST.

RESULTS

First, Table 1 presents descriptive statistics and rates per year, sex and age group. Except for 2003 and 2008 which were not full years, the proportion of blood donations per year is stable at around 20% (Table 1). Over half of the 1.4 million blood donations made to Héma-Québec over the 5-year period came from men (55% versus 45% for women). Regarding the donor's age, the highest rates were observed for the 40-49 and 50-59 age categories (respectively 60.61 and 68.08 donations per thousand people per year), whereas the older (60-69) and younger (18-29) donors have the lowest rates, followed by the 30-39 category (respectively 48.07, 49.28 and 39.98). These numbers are in line with the American and European literature showing that men aged between 40 and 60 are the most frequent donors, despite the fact that women are more numerous first-time donors [3, 8].

Blood donation clusters in Québec according to sex and age groups

As described previously, the spatial analysis using SaTScan identifies two types of clusters: the most likely cluster and the secondary clusters. In the present case, only two clusters were identified for each subgroup analysis, one for high blood donation rate and one for low rate, both significant at $p \leq 0.001$. Statistics for these clusters (total, sex and age categories) can be found in Table 2.

For the sex and age group adjusted total population analysis, the most likely low blood donation rate cluster has a relative risk of 0.56 compared to the rest of Québec. In other words, in this cluster, which groups two urban MRCs: Montréal and Laval (Figure 1a), the annual blood donation rate is 44 percent lower than in the rest of the Province. The secondary cluster has a significantly higher blood donation rate, and contains over half of all Québec's MRCs (61 MRCs included over 101 total: Figure 1a). In this secondary cluster, the annual blood donation rate is 53 percent higher than in the rest of the Province. Finally, 38 MRCs do not belong to any cluster. They are mainly located in three parts of Québec: 1) in the south, particularly in Montérégie, Estrie and Outaouais regions which border the New-York/Vermont/Maine States and the Ontario province; 2) in the north-west, including Abitibi-Témiscamingue's remote region and 3) in the north-east, including three other remote regions –Côte-Nord, Gaspésie-Îles-de-la-Madeleine and Bas-Saint-Laurent– (Figure 1.a).

The low rate most likely cluster including Montréal and Laval is repeatedly significant in all analyses except for the youngest age (18-29), where there are 18 MRCs included (Table 2; Figure 1d). High rate secondary clusters vary much more, both geographically and in terms of the number of MRCs included. For example, the total population high rate cluster contains 61 MRCs against 55 MRCs for the oldest age group (60-69) and 82 for the 30-39 age group. Moreover, in comparison with the total population secondary cluster, the relative risk for the 18-29 is lower (1.41 versus 1.53) and higher for the thirties (1.67 versus 1.53) (Table 2).

The values of the Kappa statistics, reported in Table 3, show that the degree of concordance between the total population clusters and the other subgroup clusters is moderate for four age categories (18-29, 30-39, 40-49, 50-59); good for the women and the oldest (60-69) category; and excellent for men, which means that men's high cluster geography is similar to the total population one. The comparison of the cluster maps allows an easy understanding of the differences between the obtained men, women and total population clusters (Figure 1a, b and c). Indeed, some of the MRCs which belong to remote regions of Bas-Saint-Laurent, Abitibi-Témiscamingue and Laurentides (labelled 01, 08 and 15 in Figure 1.a), are included in the secondary high blood donation rate area for women, whereas they are not for the total population secondary cluster. For the 18-29 year-old category, the most discordant aspect is certainly the significantly low blood donation rate area which contains 18 MRCs located in the south-east of Québec, in the administrative regions of Montérégie, Laurentides (close to Montréal) and Outaouais (close to Ottawa-Gatineau) (Figure 1.d). This might be explained in part by the student population, which would give on campuses in Montreal but has their permanent address elsewhere. Finally, the map for the oldest age category is quite similar to the total population one, except that none of the Bas-Saint-Laurent MRCs are located within the high rate cluster while those close to US border belong to it (Figure 1.a and h).

Factors related to the presence of high blood donation rate clusters

Results of T-test conducted between MRCs which belong to a high cluster and MRCs which are not included into any cluster are presented in Table 4. First, only two factors are significantly associated with the total population: 1) the mean network distance between the donor's residence and his/her collection site and 2) the unemployment rate. This means that MRCs included in the high blood donation rate cluster have shorter distances between donor's residence and collection site and lower unemployment rates than MRCs not included in any cluster. Yet, the differences are not significant for income and higher-education indicators. Consequently, this last result is not consistent with those of recent studies based on classic or spatial regression [11, 14, 15].

Secondly, these associations vary according to sex and age group. Network distance to blood collection site and unemployment rate are still significantly lower for men, but not for women (Table 4) while the presence of immigrants and the average income are lower for MRCs included in the women high rate cluster.

The socioeconomic profile of the high rate clusters obtained for the age categories are also quite different. Network distance to blood collection sites is significantly lower only for 18-29 and 60-69 years old analyses. Similarly, unemployment rate is lower for the 18-29 and 60-69 groups but higher for the 30-39 group. Average income is lower in high donation rate areas for the 30-39, 40-49 and 50-59 analyses while high-education is significantly higher only for the youngest blood donor (the 18-29 age category). Finally, the proportion of immigrants is lower for the 30s, 40s and 50s groups but higher for the 18-29 age group. These two last results for the young donors can be related once again to the importance of blood drive on university or college campuses, where donors are highly educated and immigrants more present.

DISCUSSION

Our analyses have shown that spatial clusters with low or high blood donation rates vary significantly in space and according to sex and age in Québec. The most likely cluster reveals that Montréal and Laval always belong to a low blood donation rate area, regardless of the age and sex categories. This situation is not surprising when we compare the socio-economic and demographic profile of these two urban regions to the rest of the province (Table 5). Although distance to collection site is smaller, population density, unemployment rate (for Montreal only) and immigration are higher in those two cities. The strong presence of immigrants (two to three times more than the rest of the province) may partially explain the unique situation of these two cities. As reported by several authors [11, 12, 14, 15], immigrants tend to give less blood for religious, cultural or other reasons such as a misunderstanding of the process or safety of the system [34, 35]. Consequently, this variable is also negatively significant in five of the eight subgroup analyses (Table 4). The spatial concentration of the immigrant population in the Montreal region (in the low cluster area), and its very low number in the 60-69 age group might explain the non-significance of this variable for the total population, the men and the 60-69 analyses.

In addition, the size and the location of significantly high or low donation rate are quite different between the eight sub-group analyses performed. Of interest is the situation of younger donors (18-29), which seems to represent a unique sub-group considering the substantial size of their low incidence area (18 MRCs in Southern Québec, including the Montréal and Laval urban areas). As said earlier, these results might be explained by the “educational” character of young donors, who often give blood for the first time on college and university campuses [36]. If Héma-Québec wants to increase and retain its young clientele, a better understanding of what is going on in regions highlighted in our results could help to create future blood donation advertising campaigns that will be appealing to this population.

The presence of regional educational institutions as well as other public services in cities far from the Montreal metropolitan area might also explain the high rate cluster’s location, a result consistent with the “regional affiliation” hypothesis brought by Cloutier et al. [13, 37]. Complementary qualitative research has

led to hypothesize that social capital and place attachment in different in those regional “hubs”, influencing blood donation behaviour. For example, rural communities, which have better donation rates, refer to different motivation factors to explain their donation: doing their “civic duty”, get together with friends (since blood drive are mobile, they represent a “social event” once or twice a year in those villages). To the opposition, reasons given by urban donors included “have a better self-esteem: it feels good”, or “convenience of the blood drive: at work” [38].

Finally, Kulldorff's spatial scan statistic was proven to be an efficient tool to assess the blood donation performance across a country or even a specific region over a period of several years. However, other epidemiological approaches such as *Flexible scan statistics* [39] or *Bayesian hierarchical modelling* [40] could also be useful to explore spatial variations in blood donation rates. It would also be interesting to conduct future research at a smaller scale, especially in urban settings like Montréal and Laval (e.g. at the census tract level). This would permits to model blood donation rates using spatial regression techniques. Such work would enhance once again our understanding of donors and non-donors behaviours, a knowledge useful to guarantee the availability of blood products in time. In terms of strategic planning and monitoring, empirical results from the Kulldorff's method can be used by any blood agencies as a tool to target areas with significantly low rates (for all the donors or a specific group).

REFERENCES

- 1 Hema-Quebec: Your first time, what was it like? , 2011. <http://www.hema-quebec.gc.ca/publications/communiqués/archives/2011/communiqués/capsules-mythes.en.html> (Last accessed 21-08-2012).
- 2 Ownby HE, Kong F, Watanabe K, Tu Y, Nass CC: Analysis of donor return behavior. *Transfusion*. 1999;**39**: 1128-35.
- 3 Zou S, Musavi F, Notari EP, Fang CT: Changing age distribution of the blood donor population in the United States. *Transfusion*. 2008;**48**: 251-7.
- 4 Crawford SO, Reich NG, An M-W, Brookmeyer R, Louis TA, Nelson KE, Notari EP, Trouern-Trend J, Zou S: Regional and temporal variation in American Red Cross blood donations, 1995 to 2005. *Transfusion*. 2008;**48**: 1576-83.
- 5 Gillespie TW, Hillyer CD: Blood donors and factors impacting the blood donation decision. *Transfusion Medicine Reviews*. 2002;**16**: 115-30.
- 6 Healy K: Embedded Altruism: Blood Collection Regimes and the European Union's Donor Population. *American Journal of Sociology*. 2000;**105**: 1633-57.
- 7 Hollingsworth B, Wildman J: What population factors influence the decision to donate blood? *Transfusion Medicine*. 2004;**14**: 9-12.
- 8 Misje AH, Bosnes V, Heier HE: Gender differences in presentation rates, deferrals and return behaviour among Norwegian blood donors. *Vox Sanguinis*. 2010;**98**: e241-e8.
- 9 Greinacher A, Fendrich K, Alpen U, Hoffmann W: Impact of demographic changes on the blood supply: Mecklenburg-West Pomerania as a model region for Europe. *Transfusion*. 2007;**47**: 395-401.
- 10 Drackley A, Newbold KB, Paez A, Heddle NM: Forecasting Ontario's blood supply and demand. *Transfusion*.**52**: 366-74.
- 11 Saberton PJ, Paez A, Newbold KB, Heddle NM: Geographical variations in the correlates of blood donor turnout rates: an investigation of Canadian metropolitan areas. *International journal of health geographics*. 2009;**8**: 56.
- 12 Cimaroli K, Paez A, Newbold KB, Heddle NM: Individual and contextual determinants of blood donation frequency with a focus on clinic accessibility: a case study of Toronto, Canada. *Health & place*. 2012;**18**: 424-33.
- 13 Cloutier MS, Apparicio P, Dube J, Charbonneau J, Delage G: Regional variation in the modeling of donation frequency: the case of Hema-Quebec, Canada. *Transfusion*. 2012.
- 14 Bekkers R, Veldhuizen I: Geographical differences in blood donation and philanthropy in The Netherlands – what role for social capital? *Tijdschrift voor Economische en Sociale Geografie*. 2008;**99**: 483-96.
- 15 Weidmann C, Schneider S, Litaker D, Weck E, Kluter H: A spatial regression analysis of German community characteristics associated with voluntary non-remunerated blood donor rates. *Vox sanguinis*. 2012;**102**: 47-54.
- 16 Demoury C, Goujon-Bellec S, Guyot-Goubin A, Hemon D, Clavel J: Spatial variations of childhood acute leukaemia in France, 1990-2006: global spatial heterogeneity and cluster detection at 'living-zone' level. *European journal of cancer prevention : the official journal of the European Cancer Prevention Organisation*. 2012;**21**: 367-74.
- 17 Kuldorff M, Feuer EJ, Miller BA, Freedman LS: Breast cancer clusters in the northeast United States: A geographic analysis. *American journal of epidemiology*. 1997;**146**: 161-70.
- 18 Tian N, Gaines Wilson J, Benjamin Zhan F: Female breast cancer mortality clusters within racial groups in the United States. *Health and Place*. 2010;**16**: 209-18.

- 19 DeChello LM, Sheehan TJ: Spatial analysis of colorectal cancer incidence and proportion of late-stage in Massachusetts residents: 1995-1998. *International journal of health geographics*. 2007;**6**: 20.
- 20 Klassen AC, Kulldorff M, Curriero F: Geographical clustering of prostate cancer grade and stage at diagnosis, before and after adjustment for risk factors. *International journal of health geographics*. 2005;**4**: 1.
- 21 Dietz NA, Sherman R, Mackinnon J, Fleming L, Arheart KL, Wohler B, Lee DJ: Toward the identification of communities with increased tobacco-associated cancer burden: Application of spatial modeling techniques. *Journal of carcinogenesis*. 2011;**10**: 22.
- 22 Kulldorff M, Athas WF, Feurer EJ, Miller BA, Key CR: Evaluating cluster alarms: a space-time scan statistic and brain cancer in Los Alamos, New Mexico. *American journal of public health*. 1998;**88**: 1377-80.
- 23 Cooper DL, Smith GE, Regan M, Large S, Groenewegen PP: Tracking the spatial diffusion of influenza and norovirus using telehealth data: a spatiotemporal analysis of syndromic data. *BMC medicine*. 2008;**6**: 16.
- 24 Huang SS, Yokoe DS, Stelling J, Placzek H, Kulldorff M, Kleinman K, O'Brien TF, Calderwood MS, Vostok J, Dunn J, Platt R: Automated detection of infectious disease outbreaks in hospitals: a retrospective cohort study. *PLoS medicine*. 2010;**7**: e1000238.
- 25 Grady SC, Enander H: Geographic analysis of low birthweight and infant mortality in Michigan using automated zoning methodology. *International journal of health geographics*. 2009;**8**: 10.
- 26 Kulldorff M: SaTScan™ v9.1.1: Software for the spatial and space-time scan statistics. 2009.
- 27 Kulldorff M, Huang L, Pickle L, Duczmal L: An elliptic spatial scan statistic. *Statistics in Medicine*. 2006;**25**: 3929-43.
- 28 Kulldorff M: Scan statistics for geographical disease surveillance: an overview; in Lawson AB, Kleinman K, (eds) *Spatial and syndromic surveillance for public health*. West Sussex, Wiley, 2005: p. 115-32.
- 29 Kulldorff M: *SaTScan™ User Guide for version 9.0*: <http://www.satscan.org>; 2010.
- 30 Weiss C, Cillis P, Rothwell N: The population ecumene of Canada: exploring the past and present. Ottawa, Geography division, Statistics Canada, Ministry of Industry, 2008: p. 31.
- 31 Landis JR, Koch GG: The measurement of observer agreement for categorical data. *Biometrics*. 1977;**33**: 159-74.
- 32 Pellett PE, Wright DJ, Engels EA, Ablashi DV, Dollard SC, Forghani B, Glynn SA, Goedert JJ, Jenkins FJ, Lee TH, Neipel F, Todd DS, Whitby D, Nemo GJ, Busch MP: Multicenter comparison of serologic assays and estimation of human herpesvirus 8 seroprevalence among US blood donors. *Transfusion*. 2003;**43**: 1260-8.
- 33 Minister of Industry: *2006 Census Dictionary*. Ottawa: Statistics Canada; 2010.
- 34 Tran N, Charbonneau J: Blood donation practices, motivations and beliefs in Montreal's Black communities: the modern gift under a new light. *Ethnicity & Health*. 2012;**3**: 11-22.
- 35 Murphy EL, Shaz B, Hillyer CD, Carey P, Custer BS, Hirschler N, Fang J, Schreiber GB: Minority and foreign-born representation among US blood donors: demographics and donation frequency for 2006. *Transfusion*. 2009;**49**: 2221-8.
- 36 Wang B, Schreiber GB, Glynn SA, Nass CC, Smith JW, Higgins MJ, Hutching ST, Wright DJ, McEntire RL, For The Retrovirus Epidemiology Donor Study ELM: Prevalence of transfusion-transmissible viral infections in first-time US blood donors by donation site. *Transfusion*. 2003;**43**: 705-12.
- 37 Cloutier M-S, Apparicio P, Charbonneau J: La géographie du don de sang au Québec: quelles variations régionales? *Cahiers de géographie du Québec*. 2011;**55**: 257-99.

- 38 Carrier É: Ville, village, banlieue : quelles différences quand vient le temps de donner du sang? Master thesis in Urban Studies; INRS-Centre Urbanisation Culture Société. Montréal, 2013.
- 39 Tango T, Takahashi K: A flexibly shaped spatial scan statistic for detecting clusters. *International Journal of Health Geographics*. 2005;**4**: (online).
- 40 Lawson AB, Biggeri AB, Boehning D, Lesaffre E, Viel JF, Clark A, Schlattmann P, Divino F: Disease mapping models: an empirical evaluation. *Statistics in Medicine*. 2000;**19**: 2217-41.

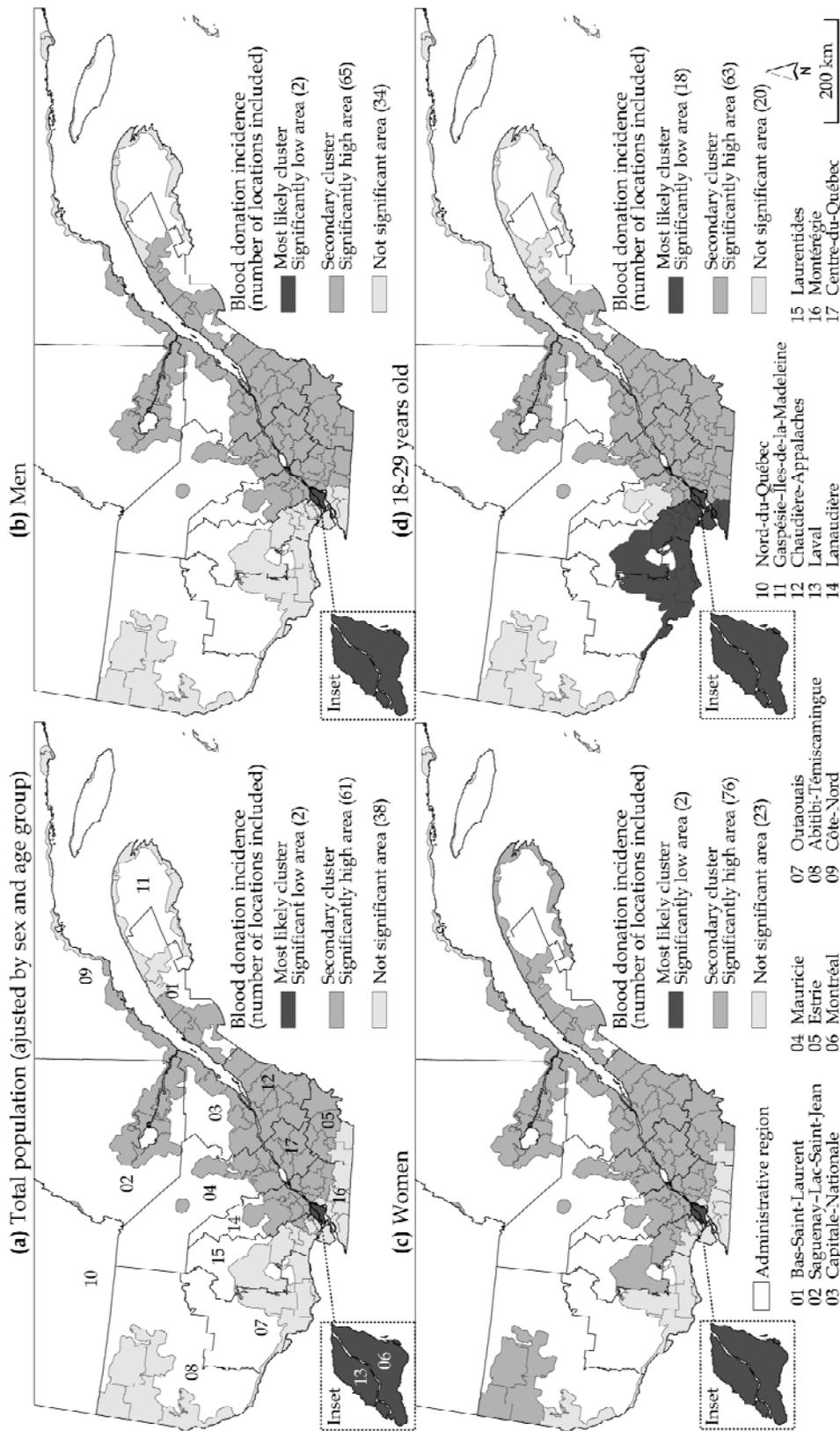


Figure 1. Blood donation clusters in Québec Province for the 2003-2008 period according to sex and age

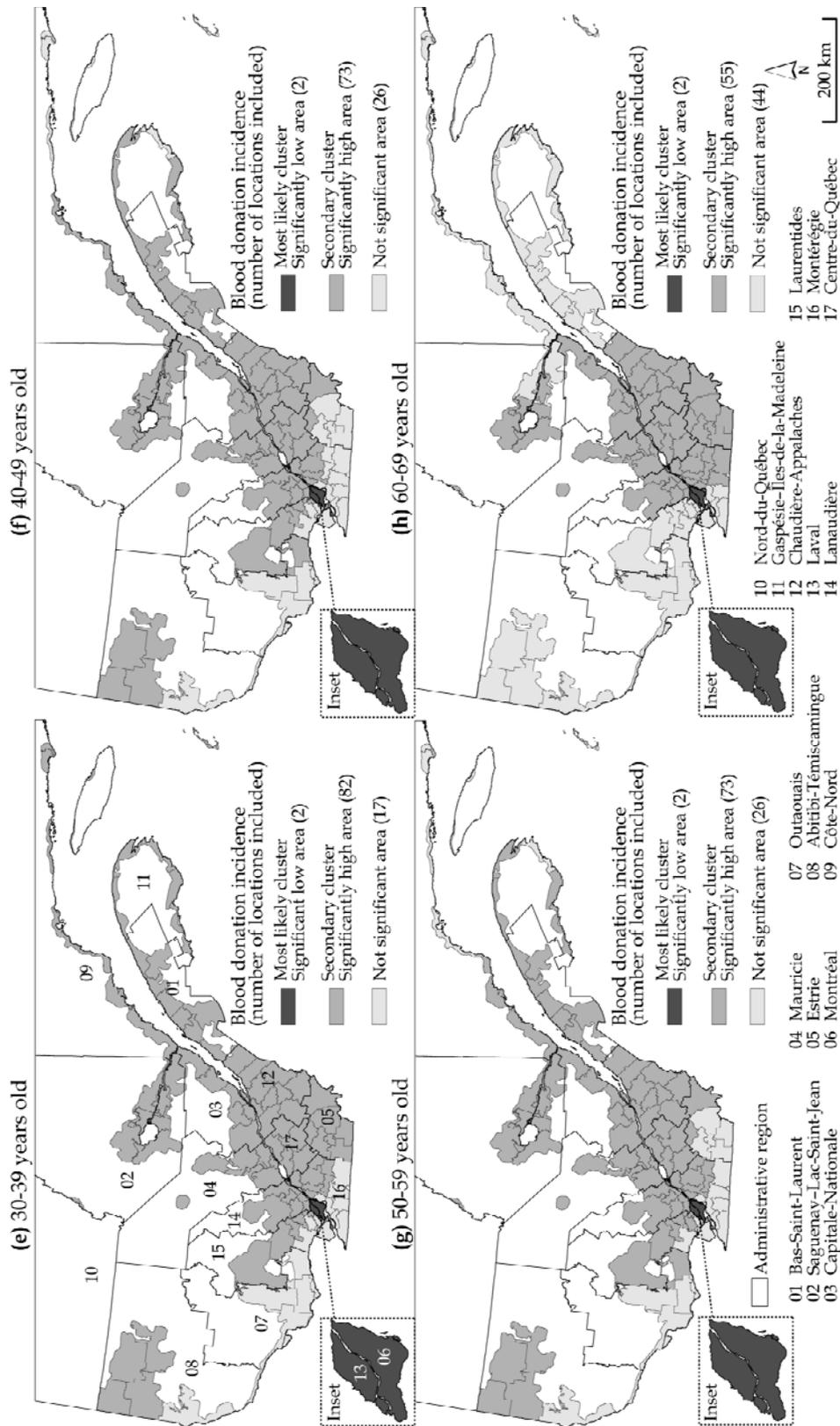


Figure 1. Blood donation clusters in Québec Province for the 2003-2008 period according to sex and age (continued)

Table 1. Blood donations per year, sex and age group

Blood donations				
	Number	Percentage	Average population during the study period	Average annual cases / 1000 people
Period (September 1, 2003 to August 31, 2008)				
All	1 428 849	100.00	5 299 912	53.90
2003	101 939	7.13	5 192 733	58.89
2004	287 032	20.09	5 236 161	54.82
2005	286 534	20.05	5 271 325	54.36
2006	275 556	19.29	5 305 973	51.93
2007	285 287	19.97	5 351 774	53.31
2008	192 501	13.47	5 405 075	53.42
Sex and age groups				
Men	788 064	55.15	2 662 064	59.18
Women	640 785	44.85	2 637 849	48.56
18-29	295 984	20.71	1 200 685	49.28
30-39	203 187	14.22	1 015 966	39.98
40-49	384 183	26.89	1 267 221	60.61
50-59	370 287	25.92	1 087 328	68.08
60-69	175 208	12.26	728 712	48.07

Table 2. Spatial cluster analysis according to sex and age groups

	Total	Sex		Age group				
	18-69 ^a	Men	Women	18-29	30-39	40-49	50-59	60-69
Most likely cluster								
# MRCs included	2	2	2	18	2	2	2	2
Observed cases	270 656	137 830	132 826	110 445	41 659	65 666	59 933	27 184
Expected cases	418 504	232 781	192 518	136 285	67 930	107 046	98 017	47 443
Annual cases / 1000	34.86	35.04	33.51	39.94	24.52	37.18	41.63	27.54
Relative risk	0.56	0.51	0.61	0.70	0.51	0.53	0.54	0.49
Log likelihood ratio	40 211.2	30 408.9	14 202.3	4 607.2	8 302.0	12 236.4	11 149.7	6 647.9
P value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Secondary cluster								
# MRCs included	61	65	76	63	82	73	73	55
Observed cases	861 333	490 895	375 066	170 652	121 368	226 372	223 319	105 750
Expected	711 912	393 432	315 748	145 250	95 522	183 144	183 249	86 463
Annual cases / 1000	65.21	73.84	57.69	57.90	50.80	74.91	82.97	58.79
Relative risk	1.53	1.66	1.45	1.41	1.67	1.57	1.55	1.56
Log likelihood ratio	31 468.0	2 4352.7	11 028.3	4 373.9	6 603.5	9 763.6	8 729.0	4 273.5
P value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

^a Adjusted for sex and age group.

Table 3. Degree of concordance between obtained clusters for the total population and those for different sex and age groups

	Sex		Age group				
	Male	Female	18-29	30-39	40-49	50-59	60-69
Simple Kappa coefficient	0.84	0.63	0.56	0.49	0.49	0.49	0.73
Degree of concordance between sub-group and total population	Excellent	Good	Moderate	Moderate	Moderate	Moderate	Good

^a The values of kappa coefficient are calculated with three categories according to the inclusion of the spatial unit into: (1) the most likely cluster, (2) a secondary cluster; (3) no cluster.

Table 4. T-test between MRCs included into a high blood donation rate cluster and those not included in any cluster according to six socio-demographic predictors (significant value in bold).

	Total population				Men			
	Mean		Difference		Mean		Difference	
	High cluster	Not incl. in cluster	Diff	P value	High cluster	Not incl. in cluster	Diff	P value
Mean distance (log) ^a	3.06	3.71	-0.65	0.000	3.08	3.73	-0.64	0.001
Population density (log) ^b	3.59	3.15	0.44	0.143	3.49	3.29	0.20	0.519
Unemployment rate	7.14	10.55	-3.41	0.003	7.33	10.60	-3.27	0.009
Immigrants (%)	2.08	2.59	-0.50	0.266	2.10	2.61	-0.51	0.274
Average income (1000\$) ^c	28.88	29.48	-0.60	0.549	28.61	30.06	-1.45	0.182
Higher educated people (%) ^d	16.46	15.86	0.60	0.625	16.30	16.10	0.20	0.873
	Women				18-29 years old			
Mean distance (log) ^a	3.22	3.59	-0.37	0.162	3.04	4.29	-1.25	<.0001
Population density (log) ^b	3.34	3.69	-0.35	0.407	3.57	2.62	0.95	0.007
Unemployment rate	8.29	8.97	-0.68	0.658	6.95	14.20	-7.24	<.0001
Immigrants (%)	1.80	3.84	-2.03	0.011	2.26	1.04	1.22	0.000
Average income (1000\$) ^c	28.46	31.25	-2.79	0.008	28.80	28.17	0.62	0.556
Higher educated people (%) ^d	16.01	16.95	-0.94	0.501	16.49	13.81	2.68	0.007
	30-39 years old				40-49 years old			
Mean distance (log) ^b	3.35	3.10	0.25	0.061	3.24	3.49	-0.24	0.308
Population density (log) ^c	3.31	3.96	-0.65	0.093	3.27	3.86	-0.59	0.073
Unemployment rate	8.82	6.66	2.16	0.021	8.42	8.54	-0.12	0.931
Immigrants (%)	1.78	4.67	-2.89	0.004	1.70	3.91	-2.21	0.002
Average income (1000\$) ^d	28.69	31.10	-2.40	0.044	28.57	30.62	-2.05	0.045
Higher educated people (%) ^e	15.85	18.04	-2.19	0.162	15.87	17.24	-1.37	0.307
	50-59 years old				60-69 years old			
Mean distance (log) ^a	3.24	3.49	-0.24	0.308	2.99	3.70	-0.72	<.0001
Population density (log) ^b	3.27	3.86	-0.59	0.129	3.68	3.10	0.57	0.049
Unemployment rate	8.42	8.54	-0.12	0.931	6.55	10.82	-4.27	<.0001
Immigrants (%)	1.70	3.91	-2.21	0.002	2.33	2.21	0.12	0.788
Average income (1000\$) ^c	28.57	30.62	-2.05	0.045	28.97	29.28	-0.32	0.740
Higher educated people (%) ^d	15.87	17.24	-1.37	0.307	16.68	15.67	1.01	0.396

^b Network distance between the donor's place of residence and the collection site (km). ^c People per square kilometre. ^d Total income in 2005 of population 15 years and over. ^e Percentage of the population 25 to 64 years with an university certificate, diploma or degree.

Table 5. Characteristics of the two spatial units belonging to the significantly low area

	Montréal	Laval	Province of Québec
Mean distance (in km) ^a	12.8	13.7	19.28
People per square kilometre	3 821	1 537	5.56
Unemployment rate	8.80	5.50	7.00
Immigrants (%)	30.73	20.18	11.45
Average individual total income (1000\$)	32.95	33.70	26.21
Higher educated people (%) ^b	39.51	27.74	26.20

^a Network distance between the donor's place of residence and the collection site. ^b Percentage of the population 25 to 64 years with an university certificate, diploma or degree. Source: 2006 census of Statistics Canada.

