Geographic clustering of elderly people with above-norm anthropometric measurements and blood chemistry

Carlos Mena,1 Eduardo Fuentes,2,3 Yony Ormazábal,1 Iván Palomo2
1Geomatics Center, University of Talca; 2Platelet Research Laboratory, Department of Clinical Biochemistry and Immunohaematology, Faculty of Health Sciences, University of Talca; 3Scientific Multidisciplinary Nucleus, University of Talca, Chile

Abstract
The global percentage of people over 60 is strongly increasing and estimated to exceed 20% by 2015, which means that there will be an increase in many pathological conditions related to aging. Mapping of the location of aging people and identification of their needs can be extremely valuable from a social-economic point of view. Participants in this study were 148 randomly selected adults from Talca City, Chile aged 60-74 at baseline. Geographic information systems (GIS) analyses were performed using ArcGIS software through its module Spatial Autocorrelation. In this study, we demonstrated that elderly people show geographic clustering according to above-norm results of anthropometric measurements and blood chemistry. The spatial identifications found would facilitate exploring the impact of treatment programmes in communities where many aging people live, thereby improving their quality of life as well as reducing overall costs.

Introduction
The global percentage of people over 60 is currently 11% with an estimated climb to 21% by 2050 (Bloom et al., 2011). This means an increase in many pathological conditions related to aging (Vermeulen et al., 1998; Fjell and Walhovd, 2010; Bueno et al., 2014) and may partly explain the higher incidence of cardiovascular disease (CVD) in older adults (Kovacic et al., 2011; Raskob et al., 2014; Sarink et al., 2014). While aging is considered an independent risk factor for the development of CVD, older adults are also at high risk for developing other chronic diseases (Lopes et al., 2012). Mena et al. (2015, 2016) showed the importance of the role of environmental factors such as parks and markets in the development of cardiovascular risk. Mapping of the location of aging people and identification of their needs can therefore be of high value from the social-economic point of view.

Geographic Information Systems (GIS) are part of the more general and modern concept denominated geo-information, or geomatics, that integrates GIS with disciplines, such as global positioning system (GPS), remote sensing (RS), cartography, geodesy and similar approaches (Gonzalez, 2013). The application of GIS technology includes fields like forestry, agriculture, mining, urban planning, network management, geo-marketing, ecology, transportation, geology, land planning, public health, etc. (Tanser, 2002). In the field of public health, GIS have multiple potential applications in areas such as disease surveillance, risk analysis, health access and planning, community health profiling, which takes into account various settings like countryside, city, built environment and other neighbourhoods (Shaw, 2012). In addition, the GIS potential has been identified by experts in relation to improved understanding of complex health issues and to support the design and evaluation of effective programmes and strategies based on population data (Carroll et al., 2014). The purpose of this study was to examine a cohort of elderly people, controlling their body measurements and general blood chemistry values, and collate this information to find out if there is any degree of geographic clustering that could be useful from the point of view of health prevention.
5% and a power of 80%, a representative sample of 148 participants aged 60–74 years at baseline (98 women and 50 men) were randomly selected from elderly people in Talca City, Chile. The sample selected represented close to 1% of Talca’s population. Only persons free of clinical cardiovascular disease at baseline were enrolled. The variables included postal address, weight, height, waist circumference, body mass index (BMI), blood pressure and blood chemistry with special reference to total lipoproteins, total cholesterol, low-density lipoprotein cholesterol (LDL-C), high-density lipoprotein cholesterol (HDL-C), triglycerides, glycaemia and uric acid. Additional details on the design of study are provided elsewhere (Palomo et al., 2007).

Geographic information system application

A geo-referenced database was built in ArcGIS software version 10 (ESRI, Redlands, CA, USA). This database provided all relevant geo-information stored as digital layers representing roads, urban blocks, green areas (squares, avenues, parks, etc.). We also used SPOTMaps satellite-generated images (http://www.intelligence-airbusds.com/spotmaps/). All this information was organised as independent information layers and conformed a base map where every single participant was geo-located according to the home address given during the medical evaluation. GIS coverage was created using georeferenced points as participant-representative elements. This coverage was the main input data for the spatial autocorrelation analysis (Mena et al., 2015, 2016). Moran’s $I$ was calculated by means of ArcGIS’s module Analyzing Patterns.

Spatial correlation statistics

Moran’s $I$

Global statistics can help identify spatial association in a large group of values registered for the spatial location of individuals (Fotheringham and Brunsdon, 1999). We used Moran’s $I$ to analyse the correlation among observations of neighbouring participants in order to detect whether the spatial distribution was clustered, dispersed or random. Moran’s $I$ was calculated according to Chen et al. (2015):

\[
I = \frac{n \cdot \sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{(\sum_{i=1}^{n} \sum_{j=1}^{n} w_{ij}) \cdot \sum_{i=1}^{n} (x_i - \bar{x})^2}, \quad i \neq j
\]

where $n$ is the number of participants, $x$ the average value for some medical measurement, $x_i$ and $x_j$ the medical measurement for participant $i$ and $j$, respectively, and $w_{ij}$ the spatial weight between participants $i$ and $j$ (Ferguson and Korfmacher, 1997). The analyses were performed using ArcGIS software through its module Analyzing Patterns according to Ferguson and Korfmacher (1997). The tool calculates the value of Moran’s $I$ and also a z-score and P value, which are both useful to evaluate the significance of the index. Z test ($z$-score) is usually used as a hypothesis test to determine whether a spatial clustering exists and its calculation is made by the formula:

![Figure 1. Getis-Ord Gi* statistic result for weight values. The figure clearly shows a hotspot for the weight variable located in the southwestern part of the city centre, an area associated with medium-low income level, where the mean registered weight is 76.65 kg. A small coldspot is also seen to the northeast (the old city), where the mean weight value is 56.75 kg.](image)
Local indicators of spatial association statistics

Local indicators of spatial association (LISA) statistics can help to identify the local spatial association between a value registered for a participant spatially located and its neighbours, given a specified distance from the observation (Roger et al., 2012). We used the $G_i^*$ local statistic (Ord and Getis, 1995) to analyse the spatial correlation among observations of neighbouring participants in order to identify statistically significant spatial clusters of high values (hotspots) and low values (coldspots). Getis-Ord $G_i^*$ statistic was calculated by formula:

$$G_i^* = \frac{\sum_{j=1}^{n} w_{ij} x_j - \bar{X} \sum_{j=1}^{n} w_{ij}}{\sqrt{\sum_{j=1}^{n} w_{ij}^2 - (\sum_{j=1}^{n} w_{ij})^2}}$$

$$G_i^*$$

where $x_i$ is the attribute value for participant $j$, $w_{ij}$ the spatial weight between participant $i$ and $j$, while $n$ corresponds to the total number of participants. $\bar{X}$ and $S$ were calculated by equations 4 and 5 below:

$$\bar{X} = \frac{\sum_{j=1}^{n} x_j}{n}$$

Eq 4.

$$S = \frac{\sum_{j=1}^{n} x_j^2}{n} - (\bar{X})^2$$

Eq 5.

The analyses were applied using ArcGIS software through its module Mapping Cluster to those medical measurements that show clustered condition by Moran’s $I$. The $G_i^*$ statistic is a z-score so no further calculations were needed. Then z-score and p-value were used to evaluate where participants with either high or low values clustered spatially. A statistically significant hotspot or coldspot is identified when a participant has a high (or low) value and is surrounded by other participants with similar (high or low) values as well. A participant with a positive z-score corresponds to a spatial clustering of high values and participant with negative z-

### Table 1. Characteristics of the study population (n=148).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Mean (SD)</th>
<th>Norm</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age* (years)</td>
<td>66 (4)</td>
<td></td>
</tr>
<tr>
<td>Height (cm)</td>
<td>160 (20)</td>
<td></td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>72.6 (13.6)</td>
<td></td>
</tr>
<tr>
<td><strong>BMI (kg/m²)</strong></td>
<td>29.1 (5.5)</td>
<td></td>
</tr>
<tr>
<td>Waist circumference (cm)</td>
<td>94.6 (11.8)</td>
<td></td>
</tr>
<tr>
<td>Blood pressure* (mm Hg)</td>
<td>144.6 (22.4)</td>
<td>&lt;120</td>
</tr>
<tr>
<td>Uric acid (mg/mL)</td>
<td>5.2 (1.6)</td>
<td>3.5-7.0</td>
</tr>
<tr>
<td>Glycaemia (mg/dL)</td>
<td>105.6 (27.1)</td>
<td>&lt;100</td>
</tr>
<tr>
<td>Total cholesterol (mg/dL)</td>
<td>208.9 (38.4)</td>
<td>&gt;200</td>
</tr>
<tr>
<td>LDL-C (mg/dL)</td>
<td>121.7 (39.7)</td>
<td>&lt;100</td>
</tr>
<tr>
<td>HDL-C (mg/dL)</td>
<td>52.6 (14.8)</td>
<td>&gt;60</td>
</tr>
<tr>
<td>Triglycerides (mg/dL)</td>
<td>161.1 (87.1)</td>
<td>≤150</td>
</tr>
</tbody>
</table>

BMI, body mass index; LDL-C, low-density lipoprotein cholesterol; HDL-C, high-density lipoprotein cholesterol. *Age range of the study subjects=60-70 years; °systolic pressure.

### Table 2. Results of statistical analysis (n=148).

<table>
<thead>
<tr>
<th>Variable</th>
<th>Moran’s $I$</th>
<th>P</th>
<th>Z-score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age* (years)</td>
<td>0.082</td>
<td>0.145</td>
<td>1.456</td>
</tr>
<tr>
<td>Height (cm)</td>
<td>0.026</td>
<td>0.479</td>
<td>0.724</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>0.118</td>
<td>0.040</td>
<td>2.051</td>
</tr>
<tr>
<td><strong>BMI (kg/m²)</strong></td>
<td>0.197</td>
<td>0.001</td>
<td>3.396</td>
</tr>
<tr>
<td>Waist circumference (cm)</td>
<td>0.163</td>
<td>0.005</td>
<td>2.800</td>
</tr>
<tr>
<td>Blood pressure* (mm Hg)</td>
<td>0.043</td>
<td>0.405</td>
<td>0.832</td>
</tr>
<tr>
<td>Uric acid (mg/mL)</td>
<td>-0.036</td>
<td>0.620</td>
<td>-0.495</td>
</tr>
<tr>
<td>Glycaemia (mg/dL)</td>
<td>-0.005</td>
<td>0.974</td>
<td>0.031</td>
</tr>
<tr>
<td>Total cholesterol (mg/dL)</td>
<td>-0.035</td>
<td>0.638</td>
<td>-0.470</td>
</tr>
<tr>
<td>LDL-C (mg/dL)</td>
<td>-0.092</td>
<td>0.159</td>
<td>-1.405</td>
</tr>
<tr>
<td>HDL-C (mg/dL)</td>
<td>-0.018</td>
<td>0.849</td>
<td>0.031</td>
</tr>
<tr>
<td>Triglycerides (mg/dL)</td>
<td>0.038</td>
<td>0.451</td>
<td>0.754</td>
</tr>
</tbody>
</table>

BMI, body mass index; LDL-C, low-density lipoprotein cholesterol; HDL-C, high-density lipoprotein cholesterol. *Age range of the study subjects=60-70 years; °systolic pressure.

### Table 3. Spatial autocorrelation (cluster) values.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Low cluster</th>
<th>High cluster</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weight (kg)</td>
<td>53.4±4.1</td>
<td>101.9±13.4</td>
</tr>
<tr>
<td><strong>BMI (kg/m²)</strong></td>
<td>16.7±9.4</td>
<td>38.37±3.2</td>
</tr>
</tbody>
</table>

BMI, body mass index. Values are expressed as mean±standard deviation.
Figure 2. Getis-Ord Gi* statistic result for waist values. Two hotspots referring to the waist circumference variable are located in the southwestern part of the city centre (same area as that with weight hotspot) but also to the northwest (both areas are characterised by medium-low income levels). Mean value for these places is 99.34 cm. A small coldspot is also seen in the northeast (the old city), where the mean value is 88.25 cm.

Figure 3. Getis-Ord Gi* statistic result for body mass index (BMI) values. A BMI hotspot is located southwest of the city centre in the same area of hotspots registered for the weight and waist circumference variables. The BMI mean value for this area is 254.65 kg/m². Coldspots are located in the north and eastern part of the city, spatially coincident with an area of high-income where the mean value of BMI is 178.68 kg/m².
score corresponds to a spatial clustering of low values. The higher (or lower) the z-score, the more intense the clustering, while z-scores near zero show no apparent spatial clustering (Getis and Aldstadt, 2004).

Statistical analysis

Means±SDs were determined using SPSS version 17.0 (SPSS Inc., Chicago, IL, USA). The GIS analyses were performed using ArcGIS software through its module Spatial Autocorrelation. P<0.05 were considered significant.

Results

Characterisation of the study population

Table 1 shows the individual characteristics together with normal values seen in the population at large. As can be seen, the majority of the study subjects, showed above normal results of body measurements and blood chemistry as follows: BMI (79%), blood pressure (86%), glycaemia (50%), total cholesterol (62%) and triglycerides (45%).

Geographic information system and spatial autocorrelation analysis

Moran’s I was calculated based on the geo-referenced ArcGIS database. The results of this statistical analysis are shown on Table 2. A significant positive spatial autocorrelation (cluster: P<0.05) was reported for weight, waist circumference and BMI, the latter being the most clustered variable. The spatial autocorrelation (clustering tendency) according to weight and BMI stood out. As shown in Table 3, people were broadly grouped around low and high values with each of the two variables clustering at low and high levels. Getis-Ord Gi* statistic was calculated in order to identify statistically significant spatial clusters of high values (hotspots) and low values (coldspots) among neighbouring participants. The calculated variables were weight, waist circumference and BMI. The results of this local statistic are shown in Figures 1-3.

Discussion

In this study, we have demonstrated for the first time that elderly people have a tendency to cluster according to some anthropometric measurements and blood chemistry levels that are above the norm. Thus, clustering as shown by GIS could help improving the quality of life by highlighting the need for exploration of the impact of treatment programmes in communities where aging people congregate. GIS is a relatively new concept in the geo-oriented sciences arising from the integration of cartography and information technologies (Bergquist and Tanner, 2012). The relational database model offered allows the linking of disparate types of geo-data by means of GIS software facilitating data management in this area of research (Wieczorek and Delmerico, 2009).

The future challenges of health-GIS integration mainly concern understanding and treatment of health problems that arise in different geographical areas (Fradelos et al., 2014). In this article, the integration of health-related data across different georeferenced areas in a city brought out novel insights how the health care system could be more involved by considering the residential distribution of the mainly aging population stratum with above-norm anthropometric measurements and blood chemistry. However, it is necessary to resolve methodological issues related to the number of cases of various ailments occurring in a geographic area as the numbers can be small and the people with health needs can be unevenly distributed. However, a better picture of this can be accomplished with efficient training in GIS concepts, practices and application besides a development focus on geomatics and health care integration.

In this study, we observed that when there was an increase or decrease in such body measurements as weight, BMI and waist circumference, elderly people were spatially autocorrelated (clustered) within limited areas, such as areas of 700-m radius. Spatial autocorrelation calculation is considered a useful technique for cluster mapping of regional health care problems, since it helps to analyse the geospatial correlation between aspects concerned with health care events (Tsai et al., 2009; Varga et al., 2013). The theory of spatial autocorrelation is an important part of geographical analysis and among the various approaches, Moran’s I is one of the most commonly used statistics due to its ability to explicitly identify spatial outliers (Sugumaran et al., 2009; Chen et al., 2015). Moran’s I is a generalisation of Pearson’s correlation coefficient, where an assessment of the correlation of a variable and its spatial location (cross-product between location similarity and attribute similarity) is performed (Zhao et al., 2013). The integration of spatial autocorrelation analyses, GIS and existing health data has provided public health officials with an efficient informatics tool that can be used to detect and study the prevalence of statistically significant hotspots with respect to many different medical issues (Toan et al., 2013), such as infectious disease (Bhunia et al., 2013), cancer incidence (Bhunia et al., 2013), intellectual disability (Goli et al., 2014), cardiac disease (Loughnan et al., 2008) and BMI (Duncan et al., 2012). In addition, results of Getis-Ord Gi* statistic show spatial clustering of medical measurement registered for elderly people, specifically weight, waist circumference and BMI, where hotspots and coldspots were shown to be related to areas characterised by different income levels.

Conclusions

The results presented here support the potential for developing a GIS-related health care approach in tasks where the impact of cluster evaluation of human health issues can be enabled when map analysis is included. In this context, it is very important to create geo-databases associated with health care providers at the local, even micro level.

References


