SPATIAL AUTOCORRELATION OF CANCER DEATHS IN ECUADOR

Mario Celleri Mujica* and Vivian Sistachs**
*Universidad Católica de Santiago de Guayaquil, Ecuador.
**University of Havana, Cuba.

ABSTRACT

Spatial autocorrelation (SA) techniques were applied to analyze the deaths caused by three types of cancer in Ecuador. One of the aims of these methods is to describe the scale over which spatial patterns of mortalities occur, which may provide suggestions concerning the agents bringing about the patterns. The analysis was carried out using data from the 23 provinces of the country. Applying global spatial autocorrelation techniques as well as local ones (Global Moran's I and Anselin's local I statistics), it was found that in certain regions of the country there is an unusual pattern of the geographical distribution of cancer mortality indicating that a similar geographic structure could be associated with some risk factor. The data used for the study correspond to deaths due to the types of cancer studied for the years 2012 to 2015 and are available from the Statistics Institute of Ecuador. The results should stimulate further research on the possible causes underlying these clusters and associations.

KEYWORDS: Cancer Death; Mortality Rate, Spatial Autocorrelation, Spatial Pattern, Spatial Analysis.

MSC:62P10

RESUMEN

Las técnicas de Autocorrelación Espacial (AS) han sido aplicadas para el análisis de causas de muerte producidas por tres tipos de cáncer en el Ecuador. Uno de los objetivos al utilizar estos métodos es describir cual es el tipo de patrón especial geográfico de la mortalidad que está ocurriendo, lo cual provee alguna hipótesis acerca de los agentes que brindan información sobre estos patrones. Este análisis fue llevado a cabo usando datos de las 23 provincias del Ecuador. Se aplicó las técnicas de autocorrelación espacial global como también las de tipo local (Índice Global de Moran y el estadístico Índice Local de Anselin), se encontró que ciertas provincias del país muestran un inusual patrón geográfico para la distribución de la mortalidad de cáncer indicando que una similar estructura geográfica podría estar asociada con algún factor de riesgo. Los datos usados para este estudio corresponden a los decesos ocurridos por cáncer en los años 2012 a 2015 y se encuentran disponibles en el Instituto Estadístico del Ecuador (INEC). Los resultados estimulan a futuras investigaciones sobre posibles causas subyacentes en estas agrupaciones y asociaciones.

PALABRAS CLAVES: Muerte por Cáncer, Tasa de Mortalidad, Autocorrelación Espacial, Patrón Espacial, Análisis Espacial.

1. INTRODUCTION

Cancer is a disease that over the years has expanded globally, creating a challenge for specialists since it is a very complex disease, which has affected a large number of people causing, in the majority of cases, the death of the individual. Currently in the country has been detected a high rate of mortality due to cancer; in the case of male gender refers to cancer of prostate which is the most common and, in the case of the female gender, the most frequent cancer is the cervix and breast.
The information referring to the mortality rate of said diseases is available in the INEC (National Institute of Statistics and Census, 2014).
The purpose of this study was: to explore the spatial patterns and clusters of the most common cancers in Ecuador using global and local spatial autocorrelation analyses (*i.e.*, by calculating the global Moran's *I* statistic and the Anselin local Moran's *I* statistic)

2. MATERIALS AND METHODS

As an example, this paper presents the situation of the prostate cancer mortality rate for 2012 in the provinces of Ecuador, that were recorded in the National Institute of Statistics and Census (INEC). Cancer data were obtained from the INEC, whose main goal is to define the incidence of cancer in the population of Ecuador. The death rate was calculated using standard procedures (Hennekens & Buring, 1987). The distribution of any

random phenomenon in a space (in this case the cancer mortality rate) can produce a pattern. According to Fischer & Wang (2011) the most commonly used measures of spatial autocorrelation in ecological, health, environmental and geological studies are Moran's I statistic. In this study, both the global Moran's I statistic and the Anselin local Moran's I statistic were applied to explore the spatial patterns of the rate of deaths due to cancer of prostate.

The global Moran's I statistic was used to represent the degree of clustering and it was calculated as follows (Anselin, 1988):

$$I = \frac{n \sum_i \sum_j W_{ij}(X_i - \bar{X})(X_j - \bar{X})}{\sum_i \sum_j W_{ij} \sum_i (X_i - \bar{X})^2}$$

Where:

X_i = the mortality rate of cancer for the ith province of Ecuador
X_j = the mortality rate of cancer for the jth province of Ecuador
$\bar{X}$ = the mean mortality rate of cancer for all of the provinces of Ecuador
W_{ij} = a weight parameter for the pair of provinces I and j that represents proximity
n = the number of provinces of Ecuador.

The Anselin local Moran's I was also calculated for each province of the country to explore spatial clusters of similar mortality rate of cancer. The Anselin local Moran's I (I_i) for the ith province is calculated according to Anselin (1988):

$$I_i = \frac{(X_i - \bar{X}) \sum_j W_{ij} (X_j - \bar{X})}{S^2}$$

Where:

X_i = the mortality rate of cancer for the ith province of Ecuador
X_j = the mortality rate of cancer for the jth province of Ecuador
$\bar{X}$ = the mean mortality rate of cancer for all of the provinces of Ecuador
W_{ij} = a weight parameter for the pair of provinces I and j that represents proximity
S = the standard deviation of the mortality rate of cancer in the area of study.
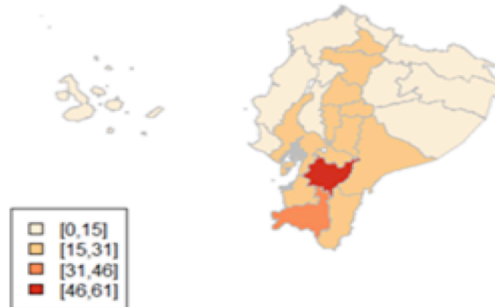
3. RESULTS

The contiguity type "queen" was used, from which the matrix showing the relationships represented in figure 1 was obtained.

Figure 1: Contiguity Provinces of Ecuador

This study explored the spatial patterns of the prostate cancer mortality rate in Ecuador using global and local spatial autocorrelation analyses. Figure 2 shows the mortality rate of prostate cancer in the provinces of Ecuador in the year 2012 for every 100,000 inhabitants.

Figure 2: Prostate cancer mortality rate 2012 . (per 100000 inhabitants)

Finally, Moran indexes, global and local, were calculated to determine the randomness of the variable studied.Next, one of the results is shown using the R package with the variable: "prostate cancer mortality rate".
This report shows that the mortality rate of prostate cancer in the provinces of Ecuador is not random, but rather presents a statistical significance of spatial autocorrelation.

```
Moran I test under randomisation

data:  datos2012$TMCP
weights: vecin

Moran I statistic standard deviate = 2.6992, p-value = 0.003476
alternative hypothesis: greater
sample estimates:
Moran I statistic      Expectation        Variance
0.24269504      -0.04545455        0.01139677
```

Figure 3 shows the values of significance of the Anselin local Moran's I for each province of Ecuador represented geographically. This index estimates a measure of spatial autocorrelation for each province, where it can see some
provinces exhibiting a statistically significant local spatial autocorrelation ($p\text{-}value$<0.05). This suggests that the number of deaths caused by prostate cancer is locally spatial autocorrelated.
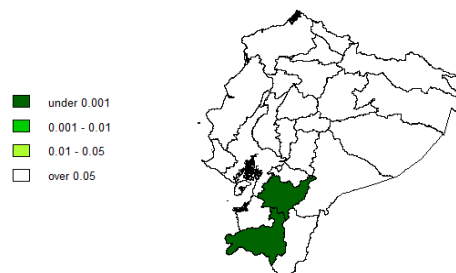
Figure 3: LISA significance map

Figure 4 shows the provinces of Ecuador that form statistically significant groupings. These groupings correspond to those in which the level of significance of the local Moran Index is high-high or low-low, of the adjacent provinces of Ecuador. Areas with *High-High* (hot spot) and *Low-Low* (cold spot) designations are those with positive spatial autocorrelation while areas with *Low-High* and *High-Low* designations belong to regions where the spatial autocorrelation is negative. There are three provinces of southern Ecuador marked as *High-High* (hot spot)*,* it means that these regions are positively spatial autocorrelated with their neighbors, that is, observations of prostate cancer mortality rate in those three areas are similar between their neighbors.

Figure 4: LISA map (hot and cold spot)
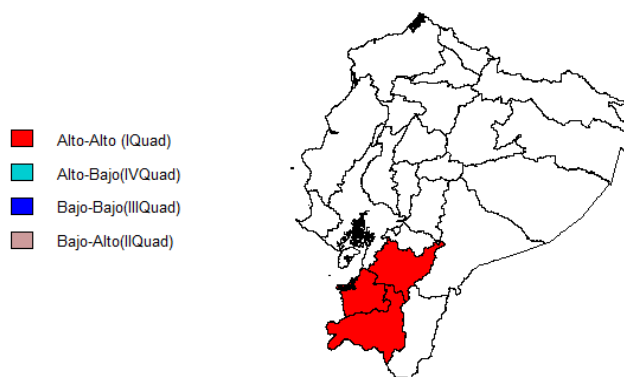
4. CONCLUSIONS

This study is the first to explore the spatial incidence patterns of the most common cancers in Ecuador using spatial autocorrelation analyses. Furthermore, it is the first to examine whether the incidence of the most common cancers in Ecuador is spatially correlated at the province level using both global and local regression methods. Overall, the results suggest that there is statistically significant spatial autocorrelation among the most common cancers in Ecuador.

REFERENCES

[1] AL-AHMADI, K., & AL-ZAHRANI, A. (2013). Spatial Autocorrelation of Cancer Incidence in Saudi Arabia. **International Journal of Enviromental Research and Public Health**, *10*, 7207-7228.

[2] ANSELIN, L. (1988). **Spatial Econometrics: Methods and Models.** Springer Science, Berlin.

[3] BANERJEE, S. (2016). Spatial Data Analysis. **Annual Review of Public Health**, *37*, 47-60.

[4] BARONIO, A., VIVANCO, A., & RABANAL, C. (2012). **Una Introducción a la Econometría Espacial: Dependencia y Heterogeneidad. Researchgate. net.**

[5] BIVAND, R., PEBESMA, E., & GÓMEZ-RUBIO, V. (2008). **Applied Spatial Data Analysis with R**. Springer, Berlin.

[6] CELEMÍN, J. P. (2009). **Autocorrelación espacial e indicadores locales de asociación espacial: Importancia, estructura y aplicación**. Mar del Plata: Grupo de estudios sobre Población y Territorio, *18*, 11-31.

[7] FISCHER, M., & WANG, J. (2011). **Spatial Data Analysis: Models, Methods and Techniques**. Heidelberg: Springer.

[8] GIRALDO, H. (s.f.). **Introducción a la Geoestadística.** Bogotá: Universidad Nacional de Colombia.

[9] HAINING, R. (2004). **Spatial Data Analysis: Theory and Practice.** Cambridge: Cambridge University Press.

[10] OCAÑA-RIOLA, R. (2010). **Common errors in disease mapping**. **Geospatial Health**, *4(2)*, 139-154 .

[11] OCAÑA-RIOLA, R., & SANCHÉZ-CANTALEJO, C. (s.f.). **Epidemiología y análisis espacial**. Información Estadística y Cartográfica de Anadalucía, *2*, 146-153.

[12] OCAÑA-RIOLA, R., MONTAÑO-REMACHA, C., & MAYORAL-CORTÉS, J. M. (2016). Geographical and Temporal Variations in Female Breast Cancer Mortality in the Municipalities of Andalusia (Southern Spain). **International Journal of Enviromental Research and Public Health,** *13***,** 1162-1179.

[13] PÉREZ, N. (2014). **Minería de datos espaciales: Detección de Agregaciones Espaciales de Riesgo de morir por cáncer de mama y próstata en España y su asociación con la Contaminación Industrial**. Madrid: Universidad Complutense de Madrid.

[14] WALLER, L., & GOTWAY, C. (2004). **Applied Spatial Statistics for Public Health Data**. New Jersey: Wiley-Interscience.