Geographical Distribution and Cluster Detection of Childhood Leukemia in the Metropolitan Area of Guadalajara, Mexico

Alberto Tlacuilo-Parra1, Roberto Garibaldi-Covarrubias2, Hugo Romo-Rubio2,3, Leonardo Soto-Sumuano1,4, Carlos Fernando Ruiz-Chávez5, Mijail Suárez-Arredondo1, Fernando Sánchez-Zubieta3 and Sergio Gallegos-Castorena3

1Medical Research Division and 2Pediatric Hematology and Oncology Department, UMAE Hospital de Pediatría CMNO, Instituto Mexicano del Seguro Social (IMSS); 3Pediatric Hematology and Oncology Department, Nuevo Hospital Civil de Guadalajara; 4Information Systems Department CUCEA, Universidad de Guadalajara, Zapopan; 5Information Technology Consultant, Geospatial Data Specialist, Guadalajara, Jal. Mexico

ABSTRACT

Background: Acute leukemia is the most common cancer in childhood. Analyzing the spatial distribution of acute leukemia may generate the identification of risk factors. Objective: To study the incidence rate of acute leukemia, its geographic distribution, and cluster detection in the metropolitan area of Guadalajara, Mexico. Methods: We included children under 15 years of age diagnosed with acute leukemia during the period 2010-2014 in the metropolitan area of Guadalajara. Each case was geo-referenced to street level to latitude and longitude coordinates using Quantum Geographic Information System (QGIS). Spatial clusters were found in the location of the acute leukemia cases applying the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm with R statistical software. Results: A total of 269 cases of leukemia were registered, 227 (84%) were acute lymphoblastic leukemia and 42 (16%) acute myeloblastic leukemia. The mean age was 6 ± 4 years. The mean incidence of acute leukemia was 6.44 cases/100,000 inhabitants: El Salto 10.12/100,000, Guadalajara 7.55/100,000, and Tlaquepaque 6.74/100,000. The DBSCAN found three clusters, all located within the municipality of Guadalajara. Conclusions: The incidence of acute leukemia in our population is higher than that in Canada and the USA. We found three spatial clusters of childhood acute lymphoblastic leukemia in the municipality of Guadalajara, suggesting the presence of local predisposing factors.

(REV INVES CLIN. 2017;69:159-65)

Key words: Cancer. Childhood. Density based clustering. Disease mapping. Geographic information system. Leukemia. Spatial epidemiology.
INTRODUCTION

Leukemia is the most common type of childhood cancer, accounting for 25 to 35% of the incidence of all cancers among most populations. The commonest type of childhood leukemia is acute lymphoblastic leukemia (ALL), which occurs in approximately 80% of leukemia cases, followed by acute myeloid leukemia (AML) and chronic myeloid leukemia, and a few other categories. A worldwide cancer survey in 2012 estimated a higher incidence in developing countries than in developed ones: in Canada, the USA, and the UK, the incidence rates are 5.8, 4.6, and 4.0 per 100,000 inhabitants per year, respectively. It is expected that this disease will continue to increase because the population in these countries is younger and growing. It is anticipated that by 2030, developing countries will account for 70% of all childhood cancers, and the cancer-related mortality rate will be 5/100,000 children. A high incidence of ALL has been reported among the Hispanic pediatric population of the USA. The worldwide incidence of ALL varies from 2.0 to 3.5 cases per 100,000 inhabitants per year, whereas in Mexico there are more than 6/100,000 inhabitants per year.

Although the etiology of most childhood leukemia is unknown, several environmental factors have been associated with the disease, including ionizing radiation, non-ionizing radiation, chemicals, pesticides, infections, and parental exposure.

Spatial cluster detection is an important tool in cancer surveillance to identify areas of high risk and to generate subsequent hypotheses about cancer etiology. A spatial disease cluster is an area with an unusually elevated disease incidence rate. Reported studies have focused on childhood leukemia because of its relatively high incidence among children compared to other malignant diseases, its apparent tendency to cluster, and the public concern over locally elevated leukemia incidence.

In Mexico, geographic information systems (GIS) have rarely been implemented to monitor the spatial dispersion of leukemia. We present the incidence rate, spatial distribution findings, and spatial clustering analysis of acute leukemia (AL) cases among children attending the three main public hospital facilities in the metropolitan area of Guadalajara, Jalisco, Mexico.

MATERIALS AND METHODS

Guadalajara (20° 41´N, 103° 21´ W, centroid: -103.33608 20.67627) is located in Jalisco, a state with an estimated population of 7.3 million inhabitants according to the latest census of 2010, in the Western region of Mexico. The metropolitan area of Guadalajara consists of eight municipalities: Guadalajara, El Salto, Tlajomulco de Zúñiga, Tlaquepaque, Tonalá, Zapopan, Juanacatlán, and Ixtlahuacan de los Membrillos. The last two municipalities were excluded because of the wide dispersion of the population in their territory.

Study population

Inclusion criteria for AL cases were: being a child under 16 years of age; living in the metropolitan area of Guadalajara; having been diagnosed with AL, confirmed by bone marrow aspiration and immunophenotypic characterization; and having been diagnosed and treated in a public hospital within the metropolitan area during the period 2010-2014.

Hospitals (National Health Service)

Medical attention for children with AL is provided by different health institutions; the public health sector treats an estimated 97.5% of the cases of AL that occur in Mexico City. In the metropolitan area of Guadalajara, children with AL receive services provided by the Unidad Médica de Alta Especialidad (UMAE) Hospital de Pediatría, from Instituto Mexicano del Seguro Social (IMSS) and by other public health services such as the Nuevo Hospital Civil de Guadalajara and the Hospital General de Occidente, which belongs to the Secretaría de Salud Jalisco (SSJ). At present, there is no official hospital-based registry of childhood leukemia cases.

Source of patient data

In each participating hospital, a pediatric hematologist or onco-hematologist identified the cases of pediatric AL from the hospital-based cancer registry. In addition, the patient’s record was reviewed to obtain the clinical variables, diagnosis, and street address.

Diagnosis

Once a diagnosis of AL was established, a pediatric hematologist or onco-hematologist confirmed the
The Agglomeration Geographic Basic Statistics (AGEB) of INEGI consists of the branch municipalities or delegations that make up the country. Its usefulness lies in allowing the formation of primary sampling units and the organization of statistical information. The AGEB has three main attributes: it is easily recognizable in the field to be bound by identifiable and enduring topographic features; it is generally homogeneous in terms of geographic, economic, and social characteristics; and the extension is such that it can be covered by a single person, and thereby used to quantify and characterize the urban population.

The research protocol was approved by the Ethics and Research Committee of the IMSS and the review board of the participating hospitals.

RESULTS

Demographic findings

During the study period 2010-2014, 269 childhood AL cases were treated among the three hospitals, distributed as follows: Nuevo Hospital Civil de Guadalajara 128 cases (47.5%); UMAE Hospital de Pediatría 127 cases (47.2%); and Hospital General de Occidente 14 cases (5.2%). Of the 269 cases, 138 were boys (51%) and 131 were girls (49%). Of the 269 patients, 227 (84%) had ALL, and 42 (16%) had AML.

Table 1 shows the clinical characteristics of patients among the participating hospitals.

Disease mapping

The overall incidence rate of AL in the metropolitan area of Guadalajara was 6.44/100,000 persons per diagnosis by bone marrow smear. Leukemia cases were grouped according to the morphological classification of the World Health Organization (WHO) and the International Classification of Childhood Cancer (based on the 2nd Edition of the International Classification of Diseases for Oncology).

Population

Data on the size of the baseline population was obtained from the 2010 census for Jalisco from the Instituto Nacional de Estadística y Geografía (INEGI). Guadalajara had a population of 1.4 million inhabitants, Zapopan 1.2 million, Tlaquepaque 608,000, Tonalá 478,000, Tlajomulco 416,000, and El Salto 138,000. All these areas together were designated as the metropolitan area of Guadalajara for this study. Twenty-nine per cent of the population was made up of children 14 years of age or younger.

Spatial data analysis

Each case was geo-referenced to the street level to latitude and longitude coordinates using Quantum Geographical Information System (QGIS), by matching the addresses with imagery and other documentary sources such as printed maps. Spatial analysis was performed in the AL cases dataset to find clusters according to their location using the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm on R statistical software using the flexible procedures for clustering package.
other populations from Latin America, such as El Salvador, Argentina, and Brazil do not have higher incidences of ALL than do populations of Caucasian origin. This high incidence may have been due to genetic or environmental factors, including ionizing radiation, non-ionizing radiation, pesticides, parental exposure, and infections.

Historically, maps have been the primary source for communicating spatial information; the most famous example of disease mapping is John Snow’s mapping of cholera cases in 1854 in London. Geographic location of disease cases can be a crucial first step in identifying and eventually preventing the source of disease outbreaks. The GIS may be used to manage the modeling and mapping of disease, to develop a new hypothesis in a geographic context, to analyze and predict future disease risks, and to undertake location/allocation analysis on the distribution of services and resources.

Figure 2 shows the geographic distribution of childhood AL and three significant density-based spatial clusters for ALL cases within the municipality of Guadalajara, where the incidence rate was 7.55/100,000 inhabitants per year. Cluster 1 was composed of 23 patients, cluster 2 of 31, and cluster 3 of 33. Because of the dispersion of the AML cases and the small number of cases, no spatial clusters were found.

### DISCUSSION

In Mexico City, as in this study of Guadalajara, the overall incidence of childhood leukemia is higher than that reported for other countries. For cities in Canada, the USA, and the UK, the incidence rates are 5.8, 4.6, and 4.0 per 100,000 inhabitants per year, respectively, whereas for the whole metropolitan area of Guadalajara, it was 6.4 cases per 100,000 inhabitants per year. This is similar to data published by IMSS for the years 1996-2002, determining the figure to be 5.5/100,000 and more recently for the years 2006-2007, with the figure of 5.7/100,000. It is surprising that the high incidence reported for children from Mexico City and Costa Rica, and for Hispanic children in Florida and Los Angeles in the USA is of lymphoid leukemia. This finding is interesting because...
that things that are closer to each other are more alike than things that are further apart. Tobler's law has been used as the basis for analysis of disease, as people who live next to each other tend to be exposed to the same factors that promote certain diseases; hence, the analysis of disease patterns and the hazards of disease clusters can help to determine the causality for disease outbreaks. Some of the environmental exposures that have been putatively linked to the etiology of AL may exhibit geographic variation. These include ionizing radiation, electromagnetic fields, chemical exposure, contaminated drinking water, and infections.

If any of these exposures do play a role in etiology, then the distribution of cases of childhood AL may be predicted to exhibit spatial heterogeneity. This is still a matter of debate; while studies from the UK, Greece, Hong Kong, and Europe have found statistically significant clustering among cases of childhood AL, studies from England, Sweden, Germany, and France showed no evidence of spatial clusters of childhood AL. In our study, we found three clusters of ALL in the metropolitan area of Guadalajara over a period of four years.

A number of statistical methods have been developed to test for spatial clustering; Alexander and Boyle evaluated the empirical performance of a number of methods for detecting spatial clustering, including those developed by Black, et al., Potthoff and Whittinghill, Cuzick and Edwards, Openshaw, Besag and Newell, Diggle and Chetwynd, Anderson and Titterington, and Oliver, et al. In this study, we decided to use the DBSCAN algorithm because it can identify clusters of arbitrary shapes based on density by the minimum number of points and the maximum distance of a point neighborhood, while performing the entire process in a reasonable period of time.
helpful comments on the manuscript draft that led to improvement of this article.

CONFLICT OF INTEREST

Data used in this study by the Hospital General de Occidente were kindly shared from Seguro Popular. Use of these data does not imply that Seguro Popular either agrees or disagrees with the analyses, interpretations or conclusions obtained.

REFERENCES


