

## Original Article



# Geographical Pathology of Acute Lymphoblastic Leukemia in Iran with Evaluation of Incidence Trends of This Disease Using Joinpoint Regression Analysis

Sajjad Rahimi Pordanjani, PhD Candidate<sup>1</sup>; Amir Kavousi, PhD<sup>2</sup>; Babak Mirbagheri, PhD<sup>3</sup>; Abbas Shahsavani, PhD<sup>4,5</sup>; Koorosh Etemad, MD, PhD<sup>1\*</sup>

<sup>1</sup>Department of Epidemiology, School of Public Health and Safety, Shahid Beheshti University of Medical Sciences, Tehran, Iran

<sup>2</sup>Workplace Health Promotion Research Center, Department of Epidemiology, School of Public Health and Safety, Shahid Beheshti University of Medical Sciences, Tehran, Iran

<sup>3</sup>Center for Remote Sensing and GIS Research, Faculty of Earth Sciences, Shahid Beheshti University, Tehran, Iran

<sup>4</sup>Environmental and Occupational Hazards Control Research Center, Shahid Beheshti University of Medical Sciences, Tehran, Iran

<sup>5</sup>Department of Environmental Health Engineering, School of Public Health and Safety, Shahid Beheshti University of Medical Sciences, Tehran, Iran

## Abstract

**Background:** The aim of present study is to determine the spatial-temporal epidemiology of acute lymphoblastic leukemia (ALL) in Iranian children.

**Methods:** This ecological study was performed using an exploratory mixed design. The study population consists of 3769 children with ALL who were reported in the National Cancer Registry Program of Iran from 2006 to 2014. Data analysis was conducted using Mann-Whitney U, joinpoint regression analysis, Global Moran's I and Anselin Local Moran's I.

**Results:** The average annual incidence rate of ALL was 2.25 per 100 000 children under 15 years of age during the study period, which was 1.37 times higher in males. The average annual percentage change (AAPC) of the disease was 7.1%, which is higher than that of developed countries. The incidence of ALL was higher in spring and summer and its peak incidence was at the age of 2–5. Spatial autocorrelation of the ALL was 0.358 and significant ( $P$  value <0.001). The high-risk cluster of ALL was identified in Fars and Kohgiluyeh and Boyer-Ahmad provinces and the low-risk cluster in Kermanshah, Zanjan and Kurdistan provinces.

**Conclusion:** The incidence of ALL is on rise in Iranian children and appropriate healthcare services are required to prevent new cases of this disease in high-risk areas.

**Keywords:** Children, Ecological, Leukemia, Seasonal variation, Spatial autocorrelation

**Cite this article as:** Rahimi Pordanjani S, Kavousi A, Mirbagheri B, Shahsavani A, Etemad K. Geographical pathology of acute lymphoblastic leukemia in Iran with evaluation of incidence trends of this disease using joinpoint regression analysis. Arch Iran Med. 2021;24(3):224–232. doi: 10.34172/aim.2021.34.

Received: September 27, 2020, Accepted: January 28, 2021, ePublished: March 1, 2021

## Introduction

Blood cancers represent the fifth most prevalent type of cancer in the world, accounting for about 8% of all cancers.<sup>1</sup> Leukemia is the fourth leading cause of death among Iranian children aged 5 to 14 years.<sup>2</sup> The most prevalent cancer of children in the world and Iran is acute lymphoblastic leukemia (ALL).<sup>3,4</sup> This disease accounts for about 30% of all childhood cancers. The peak incidence age of ALL is between 2–5 years, and it is slightly more common in males.<sup>5</sup> The high incidence and prevalence of leukemia are associated with significant mortality, incurring high diagnostic and therapeutic costs in Iran.<sup>6</sup>

Leukemia is a multifactorial disease, but despite the remarkable progress of the medical sciences, the etiology of this disease remains unknown.<sup>7,8</sup> Environmental risk factors in each region can provide the ground for ALL occurrence

and clustering.<sup>9–12</sup> Researchers attribute variations in the incidence of ALL in different communities to a number of infectious, environmental, geographic, and genetic susceptibility risk factors.<sup>5,8–11</sup>

Researchers suspect the role of infectious pathogens, especially viruses, among the predisposing, enabling, and reinforcing factors that contribute to the occurrence of leukemia. This hypothesis was first called the infectiousness hypothesis, but if a seasonal trend for the diagnosis of leukemia is identified, the infectiousness of this disease is proved with greater certainty.<sup>12–15</sup> On the other hand, infectious pathogens can easily be transmitted from the patient to other healthy people. In the event of this transmission, the possibility of spatial clustering should be considered.<sup>16,17</sup>

The hypothesis regarding the role of infectious agents in

\*Corresponding Author: Koorosh Etemad MD, PhD; Associate Professor of Epidemiology, Department of Epidemiology, School of Public Health and Safety, Shahid Beheshti University of Medical Sciences, Velenjak St, Daneshjoo Blvd, Shahid Chamran Highway, Tehran, Postal code: 198353-5511, Iran. Tel: +98 21 22 43 20 40- 1, Email: etemadk@sbmu.ac.ir

the occurrence of leukemia is still unclear, but the detection of seasonal patterns.<sup>12,14</sup> and spatial clusters<sup>18,19</sup> can serve as a supporting factor that reinforce this hypothesis. Also, the clustering and spatio-temporal analysis of leukemia can help identify the environmental factors likely to be involved in the etiology of the disease.<sup>10,13</sup>

So far, few studies have investigated the temporal trend of childhood cancer incidence. On the other hand, spatio-temporal analyzes are important tools in the epidemiology, etiology and surveillance of childhood cancer, which can be highly effective in generating new hypotheses.<sup>20,21</sup>

Without using remote sensing and medical geographic information systems (GIS), future studies would lack the precision needed to predict the trends, distributions, causes, and methods of controlling and preventing diseases.<sup>22</sup>

Since the etiology of ALL is still unknown, and to the best of the authors' knowledge, no epidemiologic and geographical pathology study of ALL has been conducted in Iran so far, this study was designed to determine temporal trends in the incidence of ALL along with spatial autocorrelation and identification of high-risk and low-risk clusters to provide deeper and more accurate insights into the hypothesis about the impact of infectious pathogens and environmental risk factors on the disease cycle.

**Materials and Methods**

**Study Design**

This ecological study was performed using an exploratory mixed design. As an exploratory research, it investigates the spatial pattern and trends of disease occurrence. The mixed design used in this study means that it investigates spatial patterns (multiple-group study) and temporal trends at the same time.<sup>23</sup> In the present study, the incidence of ALL

in different time periods and places was compared in Iran.

**Target and Study Populations**

The target population consisted of all children aged 0–14 years diagnosed with ALL in Iran and the study population comprised children with ALL who were reported in the National Cancer Registry Program from 2006 to 2014 and met the inclusion criteria. Inclusion criteria consisted of (1) access to patients' location information (2) diagnosis of acute lymphocytic leukemia, and (3) onset age of disease from the birth to 14 years of age.

**Data Sources**

Three types of data sources were consulted in this study. Data of ALL patients were derived from the list reported by the National Cancer Registry of Iran during 2006–2014. Data on healthy children aged 0–14 years in the Iranian provinces was obtained from the Statistical Center of Iran. Geographic coordinates of patients were obtained using Google Map and georeferenced layer of Iran provinces.

**Description of the Study Area**

A country in southwest Asia with an area of 1 648 195 km<sup>2</sup>, Iran sits in the Middle East. According to the latest information of World Bank in 2019, Iran's population is estimated at 82 913 906 people. Iran lies between 25° 3' and 39° 47' N in latitude and between 44° 5' and 63° 18' E in longitude. The geographical location of Iran is shown in Figure 1.

**Statistical Analysis**

Statistical analysis covered three main parts: (1) descriptive analysis of the epidemiological and demographic indicators of patients; (2) analysis of the temporal trends of incidence;

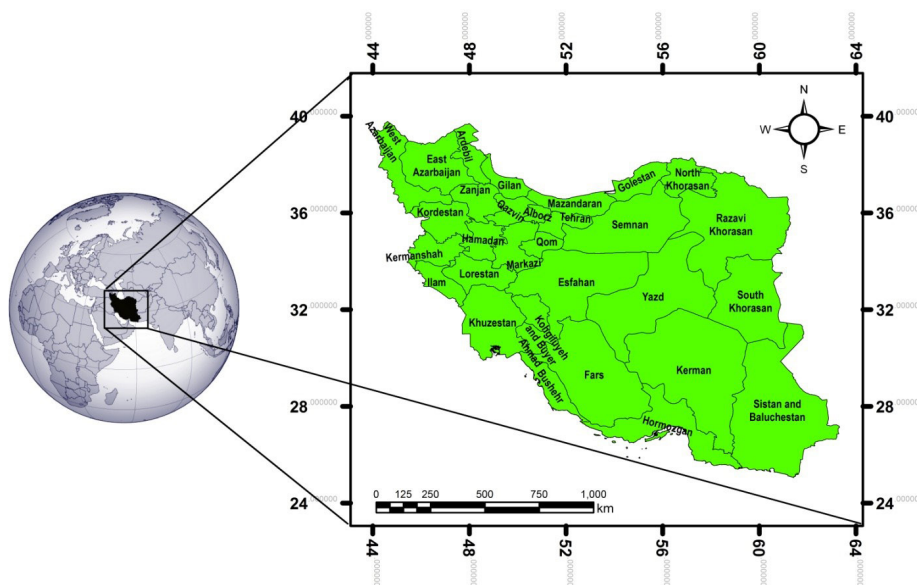


Figure 1. Geographical Location of the Study Area.

and (3) analysis of the spatial pattern of disease incidence and clustering. In the present study, all statistical tests were performed at 2-sided alpha ( $\alpha$ ) and significance level of 0.05.

#### Descriptive Analysis

In this study, central tendency and dispersion indices of ALL were obtained for the descriptive analysis of data. Then, the data was analyzed using appropriate statistical tests such as Mann-Whitney U using the SPSS software.

#### Temporal-Trend Analysis

In order to study the trends of ALL, the incidence rate was estimated separately for each year between 2006 and 2014, then trend variations of the disease were evaluated using joinpoint regression analysis as well as the version 4.7.0.0 of Joinpoint Regression Program.

Incidence rate is calculated by dividing the number of new cases of disease during a specified time interval by the average number of individuals in the population at risk in this specific time interval.

Joinpoint regression analysis is used to identify time trends and points when incidence rates change significantly. In the present study, the incidence rate (per 100 000 population) and frequency of ALL disease were the dependent variable and year, month and gender were the independent variables. Then, given that the dependent variable followed the Poisson distribution, we used the natural log-linear model to compute the annual percent change (APC) and average annual percentage change (AAPC). To calculate the number of time points at which the incidence rate changes drastically (joinpoints), the Grid Search method was utilized. The *P* value was calculated using Monte Carlo method with 4499 iterations. Bayesian information criterion (BIC) and sum of squared errors (SSE) were used to evaluate the precision of models and select the best model.<sup>24</sup>

Finally, in addition to identifying the time of joinpoints, the estimated APC-AAPC values as well as variations in regression line slopes along with the confidence interval (CI = 95%) were obtained for the annual trend of incidence between 2006 and 2014 for each gender. Also, to estimate the seasonal trend of ALL, the frequency of incident cases was calculated monthly and then the time of joinpoints, APC, AAPC values, and slope variations of the monthly trend along with the 95% CIs were calculated using joinpoint regression analysis.

#### Spatial Pattern and Cluster Analysis

In the next step, to evaluate the spatial pattern of ALL, the cumulative incidence rate (CIR) in each province of Iran was calculated in the period from 2006 to 2014 and then CIR was mapped in Iran using the ArcGIS software version 10.8.

CIR is also known as a “risk” in epidemiology. The CIR

is calculated by dividing the number of new cases of disease during a specified time interval by the total number of individuals in the population at risk at the outset of that time interval.

To assess the spatial autocorrelation of the disease, the Global Moran's I index was used. The range of this statistic is from -1 to +1 with values close to +1 indicating clustered distribution, values close to -1 indicating dispersion and zero denoting random distribution of the phenomenon under study. In our study, the null hypothesis stated that ALL distribution was randomly distributed in Iran. Finally, Anselin Local Moran's I index was employed to identify the location of high-risk and low-risk clusters of ALL in Iran. This index divides the polygons into 5 parts, including High-High, Low-Low, Low-High, High-Low, and not significant. In this context, the high-high (HH) suggests that a zone and its neighboring areas have a high incidence rate, which can be seen as the high-risk clusters of disease incidence or its hotspots. The low-low (LL), however, indicates that a zone and its surrounding areas have a low incidence of disease, therefore serving as low risk clusters of disease or coldspots.<sup>25</sup>

## Results

### Descriptive

In this study, 3769 ALL patients were enrolled after applying the inclusion and exclusion criteria. The average annual incidence rate of ALL was 2.25 per 100 000 children under 15 years of age.

The mean and standard deviation of patients' age were  $5.90 \pm 3.68$  with median and interquartile range of 5. Among all patients, 1578 (42.1%) patients were females and 2182 (57.9%) were males, so the incidence of ALL in males was 1.37 times higher than females. The mean  $\pm$  standard deviation of age in females was  $5.81 \pm 3.61$ , also median and interquartile range of age were 5 years. On the other hand, the mean  $\pm$  standard deviation of age in males were  $5.97 \pm 3.73$ , median was 5 and interquartile range was 6 years. To compare the age in male and female children, given the abnormal and right-skewed distribution of patients' age, the Mann-Whitney U test was used. The results of this test revealed that there was no significant difference in age distribution between male and female patients (*P* = 0. 261).

### Temporal Trends

The results of temporal trends incidence of ALL by year, month, sex and age of patients, which were analyzed by Joinpoint regression, are shown in Table 1 and Figure 2.

Table 1 and Figure 2A display the crude incidence rate of ALL by year. According to the results, the variation trend of the incidence increased from 2006 to 2014 at an annual average rate of 7.1% but this increase was not statistically significant (95% CI = -10.5, 28.1). However, a joinpoint was observed in 2008 with the incidence rate changing

**Table 1.** More Details of Temporal Analysis with Joinpoint Regression Models Fitted to Crude Incidence Rate and New Cases of ALL

Relevant Graph in Figure 2	Time Period	Sex	APC			AAPC			BIC	SSE
			Point Estimate	95% CI	P Value	Point Estimate	95% CI	P Value		
A	2006–2008	Both	17.74	-54.1, 201.8	0.7	7.1	-10.5, 28.1	0.5	3.04	71.05
	2008–2014		3.79	-8.6, 17.9	0.5					
B	2006–2008	Male	11.45	-60.1, 211.2	0.8	5.7	-13.1, 28.7	0.6	2.72	51.85
	2008–2014	3.91	-10.1, 20.1	0.5						
	2006–2008	Female	27.98	-48.2, 216.2	0.5	9.2	-7.9, 29.5	0.3	2.07	26.89
	2008–2014		3.57	-7.6, 16.1	0.4					
C	Jan–Apr	Both	12.37	-2.4, 29.3	0.1	0.2	-3.3, 3.9	0.9	1.04	14.91
	Oct–Dec		-3.98	-6.7, -1.1	<0.001					
D	Birth–2 year	Both	141.89	71.9, 240.4	<0.001	2.4	-2.4, 7.4	0.3	1.09	51.11
	2–14 year		-11.30	-12.6, -10	<0.001					

APC, Annual Percent Change; AAPC, Average Annual Percentage Change; BIC, Bayesian information criterion; SSE, Sum of Squared Errors.

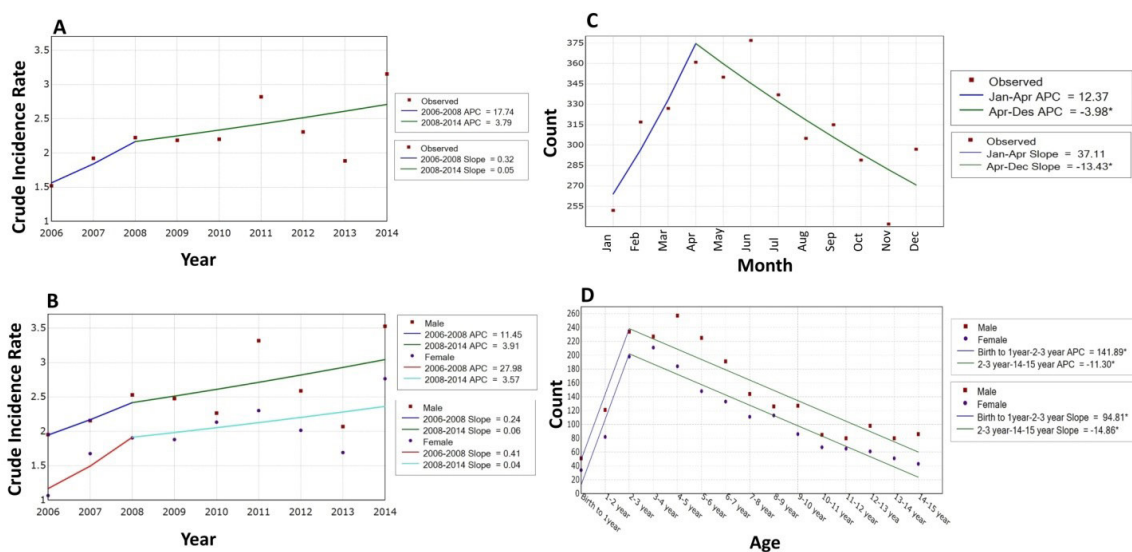
significantly at this point, so that the annual incidence rate increased more strongly between 2006 and 2008 (APC = 17.74%, 95% CI = -54.1, 201.8) but the incidence rate was mitigated between 2008 and 2014 (APC = 3.79%, 95% CI = -8.6, 17.9). Further information on this temporal trend is shown in Table 1.

Table 1 and Figure 2B show the results of the crude incidence rate of ALL by year and gender. Overall, the incidence rate of disease was higher in boys than in girls, but speed and average annual increase in incidence rates were higher in girls than in boys during the 2006–2014 period (AAPC = 9.2% vs. 5.7%). As can be seen, a joinpoint appears in 2008, so that the incidence APC of disease is higher in girls during the 2006 to 2008 period and slightly higher in boys during the 2008 to 2014 period. Overall, the results suggest that ALL incidence was higher in girls on average throughout the study period, especially in the early years of study, but by the end of study, the incidence rate of disease rose in boys. Further information on this

trend is displayed in Table 1.

Results for the trend of new ALL cases by the month of diagnosis are shown in Table 1 and Figure 2C. According to the findings, this disease has seasonal variations, so that the highest incidence of disease was observed in spring, summer, winter, and autumn, respectively, and the highest and lowest number of patients were recorded in June (377 cases) and November (242 cases), respectively. In general, the curve of monthly incidence rate of the disease is pyramid-shaped and has a joinpoint in April. According to joinpoint regression analysis, there is a progressive increase in the incidence rate of disease from January to June (APC = 12.37%, 95% CI = -2.4, 29.3), which peaks in June. However, it is followed by a drastic and significant drop from June to December (APC = -3.98%, 95% CI = -6.7, -1.1). Further information on this trend is displayed in Table 1.

Table 1 and Figure 2D depict the results for the trend of new cases of ALL by age and sex. As can be seen, the



**Figure 2.** Temporal Analysis with Joinpoint Regression Models Fitted to Crude Incidence Rate and Count of ALL in Iranian Children During 2006–2014. A) Temporal trend of crude incidence rate of ALL during 2006–2014. B) Temporal trend of crude incidence rate of ALL by sex during 2006–2014. C) Temporal trend of new cases of ALL by month during 2006–2014. D) Temporal trend of new cases of ALL by Age during 2006–2014. \*Indicates that APC and Slope are significantly different from zero at  $\alpha = 0.05$ .

trend of incidence is parallel and identical for girls and boys, with a peak incidence between 2 and 5 years of age. Also, a joinpoint was found at the age of 2 to 3 years, with a significant increase in the trend of ALL incidence from the birth to 2 years of age (APC = 141.89%, 95% CI=71.9, 240.4), after which the trend of incidence dropped progressively and significantly by the age of 14 (APC = -11, 95% CI = -12.6, -10). Further information on this trend is displayed in Table 1.

### Spatial Autocorrelation and Cluster Identification

To analyze the spatial pattern of ALL, the CIR of disease was first calculated per 100 000 children under 15 years of age during the 2006-2014 period and then the spatial distribution map of the disease was produced (Figure 3). The three provinces with the highest CIR were Fars (CIR = 61.52, 95% CI = 57.13-67.15), Kohgiluyeh and Boyer-Ahmad (CIR = 54.8, 95% CI = 45.12-67.05), and Yazd (CIR = 46.53, 95% CI = 39.05-56.17) while the three provinces with the lowest CIR were Kermanshah (CIR = 1.83, 95% CI = 1.26-3.08), East Azerbaijan (CIR = 3.96, %95 CI = 3.01-6.13) and Zanjan (CIR = 4.75, 95% = 3.21-8.04).

The results of the Global Moran's I analysis are shown in Figure 4. As can be seen, the value of this index was estimated at 0.358, indicating a high degree of spatial autocorrelation as well as great tendency of ALL for clustering. The Z score = 3.71 and P value < 0.0001 indicate that the spatial autocorrelation of the ALL is significant.

With the high spatial autocorrelation and clustered spatial distribution of the disease known, we decided to identify hotspots and coldspots of the disease in Iran. "Hotspot" means a cluster of high incidence values while "cold-spot" means a cluster of low incidence values. Statistically, if an area has a high incidence rate and

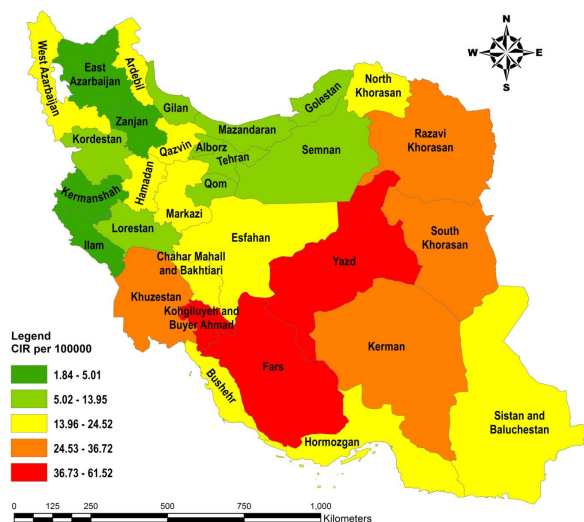


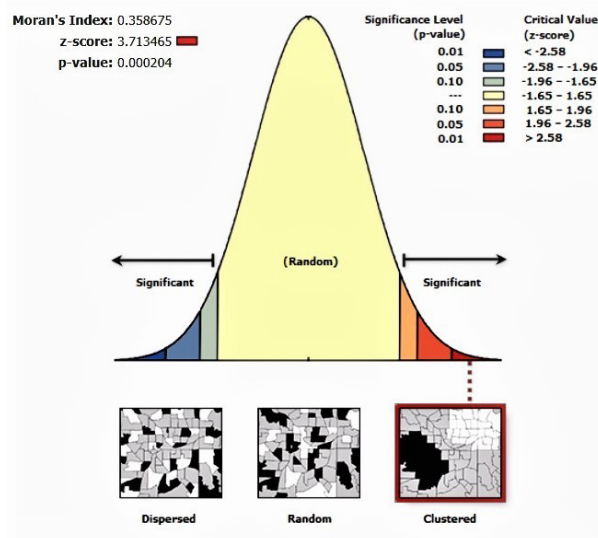
Figure 3. Cumulative Incidence Rate (CIR) of ALL During 2006-2014 in Iran.

is surrounded by high incidence rates, it is known as a hotspot. On the other hand, if the incidence rate is low in an area and its neighbors, it is known as a coldspot. To achieve this goal, Anselin Local Moran's I index was employed and the map was produced (Figure 5). According to the results, the provinces of Fars and Kohgiluyeh and Boyer-Ahmad constitute a hotspot, which exhibit spatial clusters of high values and high risk of the disease. On the other hand, Kermanshah, Zanjan and Kurdistan provinces were identified as coldspot, which indicate spatial clusters of low values and low risk of the disease. The geographic coordinates of the provincial centers identified as high-risk and low-risk ALL clusters are shown in Table 2.

### Discussion

The aim of the present study is to determine the temporal trend of ALL incidence along with spatial pattern assessment and identification of high-risk and low-risk clusters to gain deeper insights into the hypothesis of the impact of infectious pathogens and environmental risk factors on the incidence of ALL.

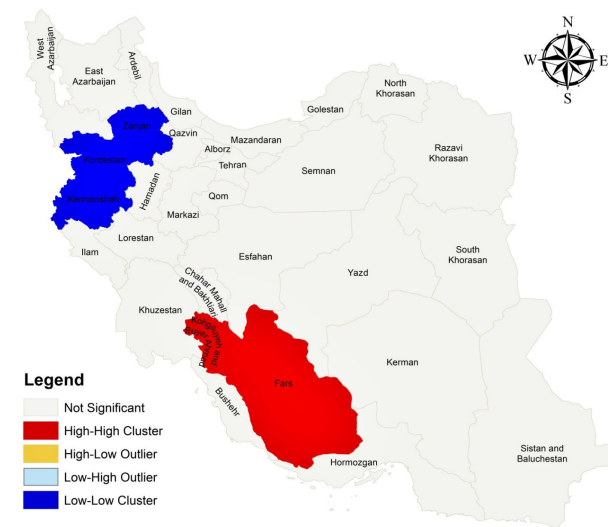
The average annual incidence rate of ALL during the 2006–2014 period was 2.25 per 100 000 children under the age of 15 years. The incidence of ALL in Iranian children appears to be lower than that of developed countries and analogous to that of developing countries. The incidence of ALL is more common in countries with a



Given the z-score of 3.7134657563, there is a less than 1% likelihood that this clustered pattern could be the result of random chance.

Global Moran's I Summary	
Moran's Index:	0.358675
Expected Index:	-0.033333
Variance:	0.011144
z-score:	3.713465
p-value:	0.000204

Figure 4. Spatial Autocorrelation of ALL Disease Using Global Moran's I Index.



**Figure 5.** Hotspots (Clusters of High Values) and Coldspots (Clusters of Low Values) of ALL in Iran Using Anselin Local Moran's I Index.

higher human development index (HDI).<sup>26</sup> Other studies have estimated the incidence rate of ALL to be between 3 and 7 per 100 000 children.<sup>20,27-29</sup>

In the present study, the mean and standard deviation of age were  $5.90 \pm 3.68$  with a median age of 5 years. There was no statistically significant difference in age distribution between male and female children. The results of the study regarding the patients' age are aligned with other studies in this field. The mean and standard deviation of patients' age were reported at  $6 \pm 4$  years by Parra et al,<sup>20</sup>  $5.5 \pm 0.92$  years by Mehrvar et al,<sup>29</sup> and 6.3 years by Sousa et al.<sup>28</sup>

The incidence of ALL new cases was higher in male (57.9%) than in female (42.1%) children, but the speed and the average annual increase in incidence were higher in female children during the 2006–2014 period (AAPC = 9.2% vs. 5.7%). These results are consistent with those reported in other studies according to which the incidence rate of ALL was higher in males than females.<sup>9,20,27,29</sup> According to American Cancer Society, 55% of new cases and 56% of ALL deaths are related to male children.<sup>30</sup>

In summary, as expected, the demographic data obtained from Iranian children with ALL resemble those of children with ALL in the rest of the world.

An analysis of temporal trends of ALL suggests that the incidence rate of disease saw an average increase of 7.1% between 2006 and 2014. These findings exhibit that in

Iran, as in many developing countries, the incidence rate of ALL is increasing with a greater intensity.

In 2019, Hubbard et al showed that the incidence and value of AAPC in ALL have been increasing in most parts of the world. According to these studies, the average increase in the incidence of ALL in countries with low HDI is greater than countries with very high HDI (AAPC = 4.04% vs. 0.44%). This study points out that the incidence of ALL usually rises by 1 to 2% each year, but this rate is significantly higher in West Asia (3.68%) and the Caribbean (4.4%).<sup>25</sup> It is predicted that the incidence of ALL rises in developing countries due to their younger population,<sup>31</sup> About 70% of cases are expected to emerge in developing countries by 2030.<sup>32</sup>

Results of new cases of disease by month indicate that ALL is more prevalent in spring and summer, which serves as high-risk season for the transmission of infectious pathogens. Many communicable diseases and some non-communicable diseases have a seasonal rhythm, which often reflects the role of infectious agents in the incidence of diseases.<sup>14</sup> Other studies have also considered the seasonal incidence of ALL as a way of lending support to the hypothesis that the disease is infectious.<sup>13,15</sup> On the other hand, joinpoint regression analysis revealed that the incidence of disease increased progressively from January to June, reaching its peak in June, and then declined significantly from June to December. Accordingly, it can be stated that the incidence of ALL is higher during warm seasons. This finding aligns with other studies in which summer and spring were identified as high risk and susceptible seasons for ALL.<sup>15,33,34</sup>

This finding can be attributed to greater and prolonged exposure to UV light during the spring and summer seasons. To date, very few studies have explored the association between ALL incidence rate and UV exposure. Masamich et al found a positive and significant correlation between leukemia death risk and UVB exposure in Japan.<sup>35</sup> In France, Coste et al demonstrated that with every 25 J/cm<sup>2</sup> increase of energy intensity in children under 5 years of age, the incidence of ALL rises significantly (standardized incidence rate = 1.09; 95% CI = 1.03–1.14). The model was well-fitted in young children and irradiation above 100 J/cm<sup>2</sup>. Under this condition, with a 25 J/cm<sup>2</sup> increase in energy intensity, a standardized incidence rate of 1.24 (95% CI = 1.14–1.36) was obtained. Finally, researchers

**Table 2.** Geographical Coordinates of Provincial Centers Identified as High Risk and Low Risk ALL Clusters

Cluster Type	District Name	District Center	Area (km <sup>2</sup> )	Latitude (°N)	Longitude (°E)	Altitude (m)	CIR*	95% CI**
High-High	Fars	Shiraz	122607	29.15	53.27	1519	61.52	57.13–67.15
	Kohgiluyeh and Buyer Ahmad	Yasuj	15504	30.77	50.83	1864	54.81	45.12–67.05
Low-Low	Kordestan	Sanandaj	29137	35.68	46.98	1463	8.85	6.02–12.17
	Zanjan	Zanjan	21773	36.49	48.38	1638	4.75	3.21–8.04
	Kermanshah	Kermanshah	24998	34.43	46.68	1374	1.83	1.26–3.08

\*Cumulative Incidence rate per 100 000 children under 15 years of age; \*\* 95% Confidence interval for cumulative incidence rate

suggested that greater exposure to UV may be associated with a higher incidence of ALL in different geographic areas, and suppression of the immune system by UV could be proposed as a biological hypothesis for justification of this finding.<sup>36</sup> Thus, the results of our study align with those reported by Coste et al. However, some studies found that the highest incidence of disease was in winter<sup>37</sup> while others provide little evidence of seasonality of the disease.<sup>14,38</sup>

These inconsistencies can be due to divergent ways of diagnosing and reporting the disease, differences in the incubation period, differences in the frequency and pattern of population referrals to medical centers, variances in cancer registry systems, different age-gender pyramid of countries, and discrepancy in the type of study design and method of analysis.

According to age-related results, ALL incidence may be significantly higher in the early years of life, meaning that the prevalence of microorganisms and pathogens is also significantly higher, but as individuals age and their immune system is strengthened, the disease incidence rate also declines. ALL can have an infectious origin, so that its incidence may be due to the abnormal body response to common infections early in life.<sup>15,39</sup> The highest incidence was observed between 2 and 5 years of age. In the studies by Hjalgrim et al,<sup>40</sup> Dohner et al,<sup>41</sup> Bahoush et al,<sup>27</sup> and Alrudainy et al,<sup>42</sup> the peak incidence of the disease was reported to be at the age of 2–5 years.

In addition to Z score and P value, Global Moran's I = 0.358 indicated that spatial autocorrelation and tendency for spatial clustering were high in ALL disease. In other words, areas with geographical proximity had a comparable incidence rate.

This is a reinforcing factor for the association of environmental risk factors, especially infectious ones, with the incidence of ALL.<sup>10,13</sup> In this regard, the results of most studies align with our study, so that a significant clustering of ALL was observed in other geographical areas, which serves as a factor to support the involvement of environmental factors, especially infections in the incidence of the disease.<sup>10,20,43</sup> In the study by Nyari et al, Global Moran's I index for spatial clustering of ALL was found to be 0.133 and significant. The researchers hypothesized an environmental etiology for the incidence of ALL, which is not yet known. They posited Chernobyl catastrophe as one reason for this clustering.<sup>43</sup> Some pathogens (HBV, HCV, HCG, HTLV-1) can be effective in leukemia due to their ability to cause acute and chronic infections in mononuclear blood leukocytes, especially lymphocytes (lymphotropic property).<sup>37,44</sup> Stephen's study suggested that the spatio-temporal pattern of abnormal leukemia incidence has been consistent with the involvement of infectious agents.<sup>45</sup>

A 2017 study by Kreis et al revealed that ETV6-RUNX1 (TEL-AML1) fusion with OR adjusted = 2.54 was the

cause of ALL clustering.<sup>46</sup> FLT3 ITD gene mutation occur in ALL patients with poor prognosis.<sup>47</sup>

Spatial distribution of the disease in Iran exhibited low incidence in northern and western provinces of Iran, high incidence rate in southern and eastern provinces and moderate incidence in central provinces. The results revealed that Fars and Kohgluyeh and Boyer Ahmad provinces were the hotspot and Kermanshah, Zanjan and Kurdistan provinces were the coldspot of ALL. The residence of children can be a reasonable substitute for environmental and local exposure because children spend a great deal of their time at home and are less likely to migrate.<sup>8,48</sup> Thus, it can be posited that the incidence of the disease is inversely correlated with latitude increase and is directly related to longitude increase. One reason for the high incidence rate of disease at low latitudes is the greater direct and intense exposure to UV radiation, which has been identified as a risk factor for ALL in other studies.<sup>35,36</sup> However, it should be noted that other climate variables such as temperature, precipitation, humidity, etc can be linked to latitude and longitude, and this relationship may be due to other climate variables and confounding factors.

In conclusion, the incidence of ALL in Iranian children is lower than in developed countries, but in Iran, as in many developing countries, the incidence of new cases is rapidly on the rise. In addition, the average annual incidence of ALL is higher in girls than boys. ALL is more likely to spread in spring and summer and warmer months, which may be due to longer exposure to UV as a carcinogen in ALL. The hotspot of disease was detected in Fars and Kohgluyeh and Boyer Ahmad and coldspot in Kermanshah, Zanjan and Kurdistan provinces. The high incidence of ALL in the early years of life, seasonal variations, specific spatial distribution of the disease, the severity of spatial autocorrelation, and the emergence of spatial clusters are evidence that support the impact of environmental risk factors and infectious pathogens on the disease cycle. It can be argued that the incidence of ALL is due to the synergistic interaction between environmental, infectious, geographical and genetic risk factors.

#### Authors' Contribution

SRP: Investigation, Conceptualization, acquisition and analysis of data, Writing - review & editing, Methodology, Resources. AK: Conceptualization, Methodology, Formal analysis, review & editing. BM: Software, Formal analysis, Visualization, Writing - review & editing. ASH: Conceptualization, Methodology, Writing - review & editing. KE: Investigation, Conceptualization, Validation, Writing - review & editing, Methodology, Project administration, acquisition and analysis of data, Supervision.

#### Conflict of Interest Disclosures

The authors have no conflicts of interest to declare.

#### Ethical Statement

This research project has been approved in Shahid Beheshti University of Medical Sciences with the code of ethics IR.SBMU.PHNS.REC.1399.117.

**Funding**

This work was supported by Shahid Beheshti University of Medical Sciences (SBUMS).

**Acknowledgements**

We would like to thank the Deputy of Research and Technology of Shahid Beheshti University of Medical Sciences, the Ministry of Health and Medical Education of Iran, the Statistical Center of Iran, the Remote Sensing and GIS Research Center of Shahid Beheshti University whose assistance was crucial to the implementation of this study. Receiving support from the Center of Excellence in Analysis of Spatio-Temporal Correlated Data at Tarbiat Modares University is acknowledged.

**References**

- Koochi F, Ghoncheh M, Salehiniya H. Epidemiology and trend of incidence of the Leukemia in Iran. *J Isfahan Med School*. 2016;33(361):2055-61.
- Amini H, Seifi M, Niazi-Esfyani S, Yunesian M. Spatial epidemiology and pattern analysis of childhood cancers in Tehran, Iran. *Journal of Advances in Environmental Health Research*. 2014;2(1):30-7.
- Koochi F, Salehiniya H, Shamlou R, Eslami S, Ghoghj Z, Kor Y, et al. Leukemia in Iran: epidemiology and morphology trends. *Asian Pac J Cancer Prev*. 2015;16(17):7759-63. doi: 10.7314/apjcp.2015.16.17.7759.
- Daneshbod Y, Esmailbeig M. The influence of Bcl-2 and myeloid antigen expression on response to therapy in childhood acute lymphoblastic leukemia. *Arch Iran Med*. 2011;14(3):170-4.
- Seghatoleslam A, Monabbati A, Bozorg-Ghalati F, Nikseresht M, Bordbar M-R, Rahvar M, et al. Expression of UBE2Q2, a putative member of the ubiquitin-conjugating enzyme family in pediatric acute lymphoblastic leukemia. *Arch Iran Med*. 2012;15(6):352-5.
- Ehsani M, Shahgholi E, Hayati H, Kebriaeezadeh A, Nikfar S, Mehrvar A, et al. Cost-analysis of treatment of pediatrics acute lymphoblastic leukemia based on ALL-BFM protocol. *Int J Pediat*. 2018;65:S113-S.
- Cárceles-Álvarez A, Ortega-García JA, López-Hernández FA, Orozco-Llamas M, Espinosa-López B, Tobarra-Sánchez E, et al. Spatial clustering of childhood leukaemia with the integration of the Paediatric Environmental History. *Environ Res*. 2017;156:605-12. doi: 10.1016/j.envres.2017.04.019.
- Konstantinoudis G, Kreis C, Ammann RA, Niggli F, Kuehni CE, Spycher BD, et al. Spatial clustering of childhood leukaemia in Switzerland: A nationwide study. *Int J Cancer*. 2017;141(7):1324-32. doi: 10.1002/ijc.30832.
- Erdmann F, Li T, Luta G, Giddings BM, Alvarado GT, Steliarova-Foucher E, et al. Incidence of childhood cancer in Costa Rica, 2000–2014: an international perspective. *Cancer Epidemiol*. 2018;56:21-30. doi: 10.1016/j.canep.2018.07.004.
- McNally RJ, Alexander FE, Vincent TJ, Murphy MF. Spatial clustering of childhood cancer in Great Britain during the period 1969–1993. *Int J Cancer*. 2009;124(4):932-6. doi: 10.1002/ijc.23965.
- Nazari S, Ebrahimi M, Abdollah Gorji F, Abadi A, Fahimzad A. Association between serum levels of MASP-2 and neutropenic febrile attacks in children with leukemia. *Arch Iran Med*. 2012;15(10):625-8.
- Karimi M, Yarmohammadi H. Seasonal variations in the onset of childhood leukemia/lymphoma: April 1996 to March 2000, Shiraz, Iran. *Hematol Oncol*. 2003;21(2):51-5. doi: 10.1002/hon.702.
- Schmiedel S, Blettner M, Kaatsch P, Schüz J. Spatial clustering and space-time clusters of leukemia among children in Germany, 1987–2007. *Eur J Epidemiol*. 2010;25(9):627-33. doi: 10.1007/s10654-010-9488-7.
- Gao F, Nordin P, Krantz I, Chia K-S, Machin D. Variation in the seasonal diagnosis of acute lymphoblastic leukemia: evidence from Singapore, the United States, and Sweden. *Am J Epidemiol*. 2005;162(8):753-63. doi: 10.1093/aje/kwi272.
- Ross J, Severson R, Swensen A, Pollock B, Gurney J, Robison L. Seasonal variations in the diagnosis of childhood cancer in the United States. *Br J Cancer*. 1999;81(3):549-53. doi: 10.1038/sj.bjc.6690729.
- Sagara Y, Iwanaga M, Morita M, Sagara Y, Nakamura H, Hirayama H, et al. Fine-scale geographic clustering pattern of human T-cell leukemia virus type 1 infection among blood donors in Kyushu-Okinawa, Japan. *J Med Virol*. 2018;90(10):1658-65. doi: 10.1002/jmv.25239.
- Greaves M. Infection, immune responses and the aetiology of childhood leukaemia. *Nat Rev Cancer*. 2006;6(3):193-203. doi: 10.1038/nrc1816.
- Birch JM, Alexander F, Blair V, Eden O, Taylor G, McNally R. Space-time clustering patterns in childhood leukaemia support a role for infection. *Br J Cancer*. 2000;82(9):1571-6. doi: 10.1054/bjoc.1999.1072.
- McNally R, Alexander F, Birch J. Space-time clustering analyses of childhood acute lymphoblastic leukaemia by immunophenotype. *Br J Cancer*. 2002;87(5):513-5. doi: 10.1038/sj.bjc.6600498.
- Tlacuilo-Parra A, Garibaldi-Covarrubias R, Romo-Rubio H, Soto-Summano L, Ruiz-Chávez CF, Suárez-Arredondo M, et al. Geographical distribution and cluster detection of childhood leukemia in the metropolitan area of Guadalajara, Mexico. *Rev Invest Clin*. 2017;69(3):159-165. doi: 10.24875/ric.17002131.
- Ye X, Torabi M, Lix LM, Mahmud SM. Time and spatial trends in lymphoid leukemia and lymphoma incidence and survival among children and adolescents in Manitoba, Canada: 1984–2013. *PLoS One*. 2017;12(4):e0175701. doi: 10.1371/journal.pone.0175701.
- Sandul Y, Kranti Suresh V, Ashish U. Geographic Information System Applications in Public Health: Advancing Health Research. In: Panagiotis M, Evangelos G, Nikolaos M, Denis Y, editors. *Effective Methods for Modern Healthcare Service Quality and Evaluation*. Hershey, PA, USA: IGI Global; 2016. p. 137-66.
- Rothman KJ, Greenland S, Lash TL. *Modern epidemiology: Wolters Kluwer Health*. Philadelphia: Lippincott Williams & Wilkins; 2008.
- Dragomirescu I, Llorca J, Gómez-Acebo I, Dierssen-Sotos T. A joint point regression analysis of trends in mortality due to osteoporosis in Spain. *Sci Rep*. 2019;9(1):4264. doi: 10.1038/s41598-019-40806-0.
- Nicoletti A, Rascunà C, Boumediene F, Vasta R, Cicero CE, Lo Fermo S, et al. Incidence of multiple sclerosis in the province of Catania. A geo-epidemiological study. *Environ Res*. 2020;182:109022. doi: 10.1016/j.envres.2019.109022.
- Hubbard AK, Spector LG, Fortuna G, Marcotte EL, Poynter JN. Trends in International Incidence of Pediatric Cancers in Children Under 5 Years of Age: 1988–2012. *JNCI Cancer Spectr*. 2019;3(1):pkz007. doi: 10.1093/jncics/pkz007.
- Bahoush G, Nojoomi M. Frequency of Cytogenetic Findings and its Effect on the Outcome of Pediatric Acute Lymphoblastic Leukemia. *Med Arch*. 2019;73(5):311-5. doi: 10.5455/medarh.2019.73.311-315.
- Sousa DWLd, Ferreira FVdA, Félix FHC, Lopes MVdO. Acute lymphoblastic leukemia in children and adolescents: prognostic factors and analysis of survival. *Rev Bras Hematol Hemoter*. 2015;37(4):223-9. doi: 10.1016/j.bjhh.2015.03.009.
- Mehrvar A, Faranoush M, Asl AAH, Tashvighi M, Fazeli MA, Mehrvar N, et al. Epidemiological features of childhood acute leukemia at MAHAK's Pediatric Cancer Treatment and Research Center (MPCTRC), Tehran, Iran. *Basic Clin Cancer*



- Res. 2015;7(1):9-15.
30. Siegel RL, Miller KD, Jemal A. Cancer statistics, 2019. *CA Cancer J Clin.* 2019;69(1):7-34. doi: 10.3322/caac.21551.
  31. Ferlay J, Soerjomataram I, Ervik M, Dikshit R, Eser S, Mathers C, et al. GLOBOCAN 2012 v1. 0, cancer incidence and mortality worldwide: IARC CancerBase No. 11. Lyon, France: International Agency for Research on Cancer; 2013.
  32. Pérez-Cuevas R, Doubova SV, Zapata-Tarres M, Flores-Hernández S, Frazier L, Rodríguez-Galindo C, et al. Scaling up cancer care for children without medical insurance in developing countries: the case of Mexico. *Pediatr Blood Cancer.* 2013;60(2):196-203. doi: 10.1002/pbc.24265.
  33. Gilman E, Sorahan T, Lancashire R, Lawrence G, Cheng K. Seasonality in the presentation of acute lymphoid leukaemia. *Br J Cancer.* 1998;77(4):677-8. doi: 10.1038/bjc.1998.109.
  34. Badrinath P, Day N, Stockton D. Seasonality in the diagnosis of acute lymphocytic leukaemia. *Br J Cancer.* 1997;75(11):1711-3. doi: 10.1038/bjc.1997.292.
  35. Uehara M, Takahashi K, Hoshuyama T, Pan G, Feng Y. Geographical correlation between ambient UVB level and mortality risk of leukemia in Japan. *Environ Res.* 2003;92(2):78-84. doi: 10.1016/s0013-9351(03)00034-3
  36. Coste A, Goujon S, Boniol M, Marquant F, Faure L, Doré J-F, et al. Residential exposure to solar ultraviolet radiation and incidence of childhood hematological malignancies in France. *Cancer Causes Control.* 2015;26(9):1339-49. doi: 10.1007/s10552-015-0629-x.
  37. Timonen T. A hypothesis concerning deficiency of sunlight, cold temperature, and influenza epidemics associated with the onset of acute lymphoblastic leukemia in northern Finland. *Ann Hematol.* 1999;78(9):408-14. doi: 10.1007/s002770050539.
  38. Gao F, Chia K-S, Machin D. On the evidence for seasonal variation in the onset of acute lymphoblastic leukemia (ALL). *Leuk Res.* 2007;31(10):1327-38. doi: 10.1016/j.leukres.2007.03.003.
  39. Greaves M. A natural history for pediatric acute leukemia. *Blood.* 1993;82(4):1043-51.
  40. Hjalgrim LL, Rostgaard K, Schmiegelow K, Söderhäll S, Kolmannskog S, Vettenranta K, et al. Age-and sex-specific incidence of childhood leukemia by immunophenotype in the Nordic countries. *J Natl Cancer Inst.* 2003;95(20):1539-44. doi: 10.1093/jnci/djg064.
  41. Döhner H, Weisdorf DJ, Bloomfield CD. Acute myeloid leukemia. *N Engl J Med.* 2015;373(12):1136-52. doi: 10.1056/NEJMra1406184.
  42. Alrudainy LA, Salih HM, Aldorky M. Incidence and pattern of childhood leukaemia in Basrah, Iraq during 2003-2007. *Iran J Blood Cancer.* 2009;2(1):11-7.
  43. Nyari T, Ottóffy G, Bartyik K, Thurzó L, Solymosi N, Cserni G, et al. Spatial clustering of childhood acute lymphoblastic leukaemia in Hungary. *Pathol Oncol Res.* 2013;19(2):297-302. doi: 10.1007/s12253-012-9582-0.
  44. Bagheri K, Yaghobi R, Karimi MH, Mirzaei M, Ramzi M. Determination of Serologic and Molecular Prevalence of Hepatitis Type B, C, and G Infections in Patients with Hematological Malignancy in the South of Fars Province, Iran. *Journal of Fasa University of Medical Sciences.* 2011;1(3):168-74.
  45. Francis SS, Selvin S, Yang W, Buffler PA, Wiemels JL. Unusual space-time patterning of the Fallon, Nevada leukemia cluster: Evidence of an infectious etiology. *Chem Biol Interact.* 2012;196(3):102-9. doi: 10.1016/j.cbi.2011.02.019.
  46. Kreis C, Lupatsch JE, Niggli F, Egger M, Kuehni CE, Spycher BD, et al. Space-time clustering of childhood leukemia: evidence of an association with ETV6-RUNX1 (TEL-AML1) fusion. *PLoS One.* 2017;12(1):e0170020. doi: 10.1371/journal.pone.0170020.
  47. Nasiri N, Shaikhy M, Zaker F, Hosseini S, Moosavi SA, Marjani AJ. Detection and biological characteristic of FLT3 gene mutations in children with acute leukemia. *Archives of Iranian medicine.* 2014;17(4):258-261.
  48. Jacquez GM. Current practices in the spatial analysis of cancer: flies in the ointment. *Int J Health Geogr.* 2004;3(1):22. doi: 10.1186/1476-072X-3-22.