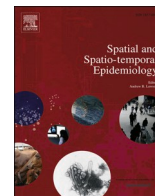




Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company's public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.



Spatial variation and hotspot detection of COVID-19 cases in Kazakhstan, 2020

Andrey Kuznetsov^{*}, Veronika Sadovskaya

National Scientific Center of Especially Dangerous Infections, Almaty, Kazakhstan

ARTICLE INFO

Keywords:
COVID-19
Spatial analysis
Cluster detection

ABSTRACT

Background: COVID-19 is the life-threatening infectious disease of zoonotic origin that has epidemic spread in Kazakhstan. The use of geoepidemiological techniques to detect territories of risk (hotspots) is essential for implementing control measures in the target area.

This study aims to conduct spatial analysis of the COVID-19 epidemic in Kazakhstan to increase understanding of the current features of the virus distribution and to explore its geographical patterns, especially its spatial clustering.

Methods: We used geographic information software (QGIS, GeoDa) to perform spatial analysis (Nearest Neighbour Analysis, Global Moran's I, Getis-Ord G_i^* , LISA) and to detect COVID-19 risk clusters in the entire territory of Kazakhstan.

Results: Clusters of COVID-19 cases were detected using the Getis-Ord G_i^* analysis (with first order Queen Continuity matrix) in two oblasts of Kazakhstan: Almaty (Ilyyskiy, Karasayskiy, Raiymbekskiy, Talgarskiy rayons and city of Almaty) and Aqmola (Arshalynskiy, Ereymengauskiy, Korgalzhynskiy and Shortandinskiy rayons). LISA defined four High-High clusters of COVID-19 cases in the Almaty oblast (Ilyyskiy, Karasayskiy and Talgarskiy rayons) and city of Almaty.

1. Introduction

The first cases of COVID-19 in Kazakhstan were registered on 13 March 2020, in the cities of Almaty and Nur Sultan. These cases were imported from Germany and Italy (Maukayeva, 2020). From March 20, confirmed cases of COVID-2019 began to be registered in other regions of the country (Maukayeva, 2020). The vast majority of symptomatic patients had mild disease manifestations and the proportion of moderate disease was around 10% (Semenova et al., 2020). According to the SEIR model, there will be 156 thousand hospitalized patients due to severe illness and 15.47 thousand deaths at the peak of an outbreak (Semenova et al., 2020).

2. Objective

This study aims to conduct spatial analysis of the COVID-19 epidemic in Kazakhstan to increase understanding of the current features of the virus distribution and to explore its geographical patterns, especially its spatial clustering.

3. Data and methods

3.1. Target area

Republic of Kazakhstan is a state in the Central Asia that lies between Russia in the west and north, China in the east, and Kyrgyzstan, Uzbekistan and Turkmenistan in the south. Kazakhstan consists of 14 oblasts (provinces) and three cities of republican subordination (Nur-Sultan, the capital; Almaty and Shymkent). Oblasts are further divided into 174 of rayons (districts).

3.2. Study design

To evaluate spatial patterns of the COVID-19 distribution in different administrative territories of Kazakhstan, we carried out a descriptive cross-sectional study, which involved 14 oblasts and three cities of republican subordination. GIS analysis of the COVID-19 cases database was conducted for the entire study area at the rayon level.

As a case we considered a patient positive for SARS-CoV-2 RNA

^{*} Corresponding author.

E-mail address: dr.a.n.kuznetsov@gmail.com (A. Kuznetsov).

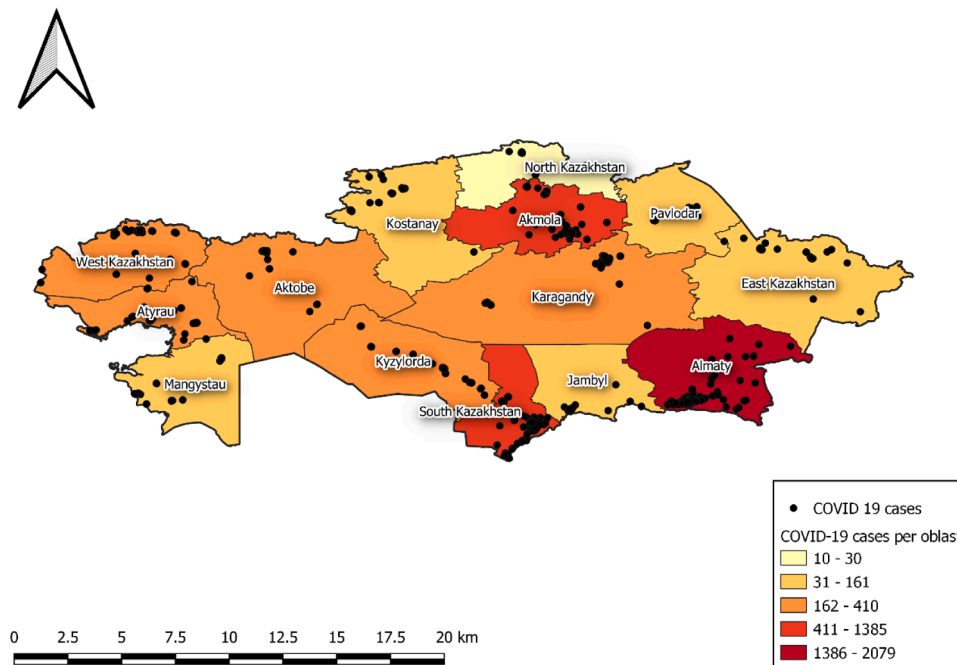


Fig. 1. Spatial distribution of COVID-19 cases in oblasts of Kazakhstan, March–July, 2020.

detected in clinical samples by RT PCR with a test system of any manufacturer.

3.3. Data source

Data about confirmed cases of COVID-19 were collected by residents of the CDC Central Asia Region Field Epidemiology Programme in June–July, 2020, from local health departments with subsequent taking coordinates of a case registration place.

Population data from 2019 in oblasts (provinces) and rayons (districts) of Kazakhstan were obtained from the Committee on Statistics, Ministry of National Economy of the Republic of Kazakhstan, <https://stat.gov.kz/>.

3.4. Dataset development

A geographic information system (GIS) database was constructed to evaluate spatial patterns of COVID-19 in the targeted area.

Excel spreadsheet with collected coordinates of COVID-19 cases registration was saved as a point shapefile with QGIS. The number of cases points were calculated per each rayon (district) to allow for spatial analyses.

3.5. Data analysis

3.5.1. Spatial statistical methods

- Conceptualisation of spatial relationship (Average Nearest Neighbour's analysis) was used to evaluate clustering in rayons. Every rayon is assumed to be a neighbour of other rayon, and the influence of rayons decreases with increased distance.
- Univariate exploratory analysis of the spatial data was performed for global spatial autocorrelation investigation of the COVID-19 cases number in rayons of Kazakhstan under the assumptions of normality and randomization. Global index of spatial autocorrelation (Moran I) was used to assess spatial dependencies across rayons with respect to COVID-19 cases presence. The Global Moran's I statistic is evaluated in terms of a null hypothesis that COVID-19 cases are spatially randomly distributed across rayons.

- High- and Low-Clustering (Getis-Ord G_i^* , LISA analyses). Rayons are deemed to be a cluster of either High-High (i.e. high number of COVID-19 cases per polygon surrounded by high numbers), Low-Low, High-Low, or Low-High values relative to neighbouring rayons. The null hypothesis states that there is no spatial autocorrelation or association of cases between districts.

We used a spatial contiguity (i.e. the property of sharing a common boundary or vertex) for rayons of Kazakhstan to assess unusual features in the connectivity distribution (Anselin, 1995). For this purpose, we used the first order Queen's contiguity method as a most suitable (Tsai et al., 2009). Based on the results of the connectivity distribution, we constructed a first order queen polygon contiguity weight file for rayons which share common boundaries and vertices. Weights file was utilized for subsequent local spatial calculations (Getis-Ord G_i^* , LISA).

3.5.2. Software

Data transformations were performed with the R free statistical environment, version 4.0.2.

Quantum GIS of version 3.10.08, an open source GIS application, was used to construct spatial databases and for data visualisation. The Global Moran's I, Getis-Ord G_i^* , Univariate Local Moran's I (LISA) analyses were conducted in GeoDa, version 1.14.0 (<http://geodacenter.asu.edu/>).

4. Results

4.1. Assessment whether the distribution of COVID-19 cases in the targeted area is random

We included into dataset 6165 confirmed cases of COVID-19 of 6691. We excluded 526 case points (7.9%) from the database because their coordinates went beyond the rayon of registration (incorrect coordinates).

Of 6165 cases included in the study, 3370 (54.7%) were men and 2560 were women; gender was unknown in 235 cases. Mean age of cases was 36.2 ± 16.6 years. Of them, 405 were children under the age of 14.

Among selected patients, 5115 cases (83%) were hospitalized, 32 (0.52%) died.

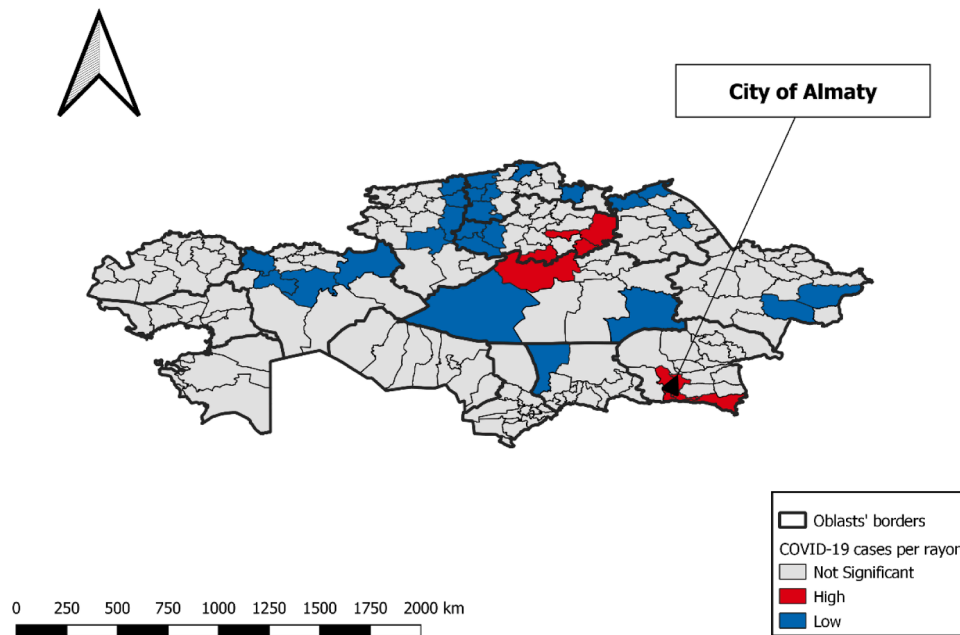


Fig. 2. Identification of COVID-19 significant clusters with Getis-Ord GI^* autocorrelation analysis.

Table 1
Getis-Ord GI^* statistics for several rayons of Kazakhstan.

Oblast	Rayon	Cases per rayon	GI^*	P -value
Almaty	City of Almaty	1187	0.077	0.023
Almaty	Ilyyskiy	109	0.052	0.001
Almaty	Karasayskiy	501	0.052	0.008
Almaty	Raiymbekskiy	19	0.022	0.043
Almaty	Talgarskiy	99	0.040	0.006
Aqmola	Arshalynskiy	7	0.042	0.022
Aqmola	Ereymentauskiy	5	0.023	0.049
Aqmola	Shortandinskiy	3	0.044	0.01

Travel history was collected in 5784 COVID-19 patients; most of cases (5511, 89.4%) were locally acquired. 246 cases had a history of air flight within 14 days before the disease onset, and 27 had a train trip.

A choropleth map of the distribution of cases (Fig. 1) is showing that the highest counts are in the south-eastern portion of the country (Almaty and South Kazakhstan oblasts) and in Aqmola oblast, where the state capital, city of Nur Sultan, is situated.

Using the Average Nearest Neighbour’s analysis, we detected clustering of 6165 cases’ points in oblasts of Kazakhstan (z -score: -142.7 ; $p = 0.001$). This indicates that COVID-19 cases in the study area are clustered.

The Moran’s I for COVID-19 cases per rayon in all oblasts of Kazakhstan is positive (0.104) and statistically significant (p -value = 0.022), so mainly cases’ points dispersed and distributed not randomly.

4.2. Hot spot analysis

Clusters of COVID-19 cases were detected using the Getis-Ord GI^*

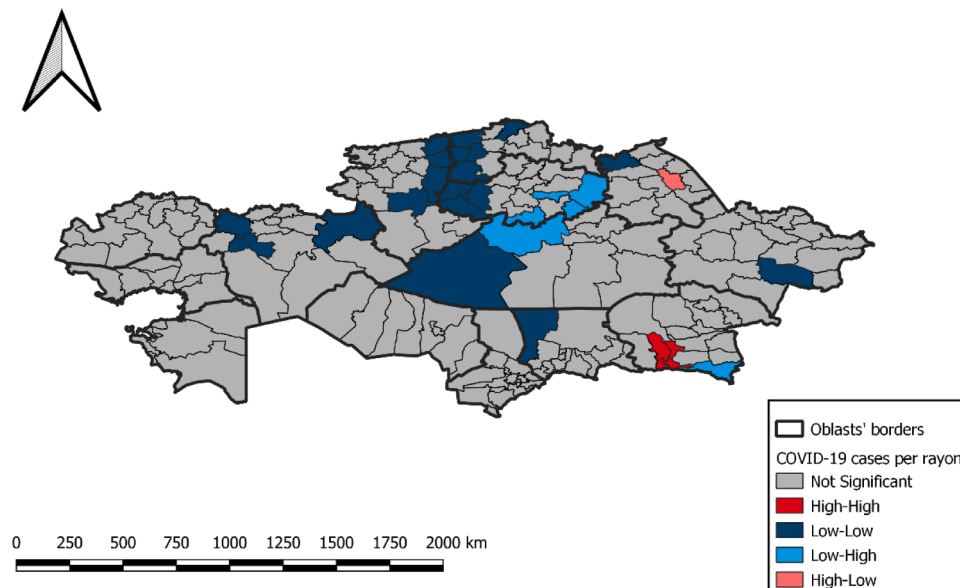


Fig. 3. COVID-19 cases significant clusters identified using LISA in Kazakhstan, March–July 2020.

Table 2
LISA statistics for several rayons of Kazakhstan.

Oblast	Rayon	Cases per rayon	LISA clusters	P-value
Almaty	City of Almaty	1187	11.082	0.023
Almaty	Ilyyskiy	109	1.149	0.001
Almaty	Karasayskiy	501	5.606	0.008
Almaty	Talgarskiy	99	0.715	0.006

analysis (with first order Queen continuity matrix) in two oblasts of Kazakhstan: Almaty (Ilyyskiy, Karasayskiy, Raiymbekskiy, Talgarskiy rayons and Almaty city) and Aqmola (Arshalynskiy, Ereymentauskiy, and Shortandinskiy rayons) (Fig. 2, Table 1).

LISA defined four high-high clusters of COVID-19 cases in the Almaty oblast (Ilyyskiy, Karasayskiy and Talgarskiy rayons) and city of Almaty. There's one high-low outlier (Fig. 3, Table 2).

In addition, we calculated a population density for each rayon of the target area and used the variable to measure its influence on the spread of COVID-19 with Bivariate Local Moran's I statistics. We identified four high-high COVID-19 clusters in the same rayons of Almaty oblasts as above that allows us to consider that high population density is a factor of risk for COVID-19 spread.

5. Discussion

Our study provides a comprehensive description of the spatial distribution of laboratory-confirmed COVID-19 cases, presents the spatial pattern of the epidemic in Kazakhstan, and tests the cluster characteristics of patients infected with SARS-CoV-2. Overall, we detected at least four hotspots of COVID-19 in Almaty oblast.

The Getis-Ord G_i^* statistic was used in this study to recognize hotspots of COVID-19 cases based on the case registration coordinates. With

this method, the confirmed cases shown local spatial aggregation, with hotspot areas found mostly in the southeast area of Kazakhstan (Almaty oblast and city of Almaty). While both the Getis-Ord G_i^* and LISA methods detected the same clusters in Almaty oblast, clusters were only detected using the first method in Aqmola ($n=3$) oblast.

This epidemiological trend may be related to population factors, such as local population density, that was confirmed by the Bivariate Local Moran's I analysis, as well as improper isolation measures, and other factors. As we do not explore these factors in this paper, further work should be conducted to study the influences of relevant factors.

Acknowledgement

This paper and the research behind it would not have been possible without the exceptional support of Dilyara Nabirova from the CDC Central Asia Region Field Epidemiology Programme.

Amber Dimer, Health Scientist at Centers for Disease Control and Prevention, have also looked over our manuscript and made valuable remarks that allowed us to improve the article.

References

- Anselin, L., 1995. Local indicators of spatial association—LISA. *Geogr. Anal.* 27, 93–115. <https://doi.org/10.1111/j.1538-4632.1995.tb00338.x>.
- Maukayeva, S., 2020. Epidemiologic character of Covid-19 in Kazakhstan: a preliminary report. *North Clin. Istanbul* 7, 210–213. <https://doi.org/10.14744/nci.2020.62443>.
- Semenova, Y, Glushkova, N, Pivina, L, Khismetova, Z, Zhunussov, Y, Sandybaev, M, Ivankov, A., 2020. Epidemiological characteristics and forecast of COVID-19 outbreak in the republic of Kazakhstan. *J. Korean Med. Sci.* 35, 1–12. <https://doi.org/10.3346/JKMS.2020.35.E227>.
- Tsai, PJ, Lin, ML, Chu, CM, Perng, CH., 2009. Spatial autocorrelation analysis of health care hotspots in Taiwan in 2006. *BMC Public Health* 9, 1–13. <https://doi.org/10.1186/1471-2458-9-464>.