MEDINFO 2021: One World, One Health – Global Partnership for Digital Innovation
P. Otero et al. (Eds.)

doi:10.3233/SHTI220260

COVID-19 Geographical Maps and Clinical Data Warehouse PREDIMED

Svetlana Artemova[a,b], Alban Caporossi[a,b,c], Christophe Cancé[b,c], Pierre-Ephrem Madiot[a], Benjamin Nemoz[a], Sylvie Larrat[a], Alexandre Moreau-Gaudry[a,b,c], Jean-Luc Bosson[a,b,c]

[a] *CHU Grenoble Alpes, F-38000, Grenoble, France,*
[b] *Clinical Investigation Center-Technological Innovation, Univ. Grenoble Alpes, INSERM CIC1406, CHU Grenoble Alpes, F-38000, Grenoble, France,*
[c] *TIMC, Univ. Grenoble Alpes, CNRS, VetAgro'Sup, CHU Grenoble Alpes, Grenoble INP, F-38000, Grenoble, France*

Abstract

PREDIMED, Clinical Data Warehouse of Grenoble Alps University Hospital, is currently participating in daily COVID-19 epidemic follow-up via spatial and chronological analysis of geographical maps. This monitoring is aimed for cluster detection and vulnerable population discovery. Our real-time geographical representations allow us to track the epidemic both inside and outside the hospital.

Keywords:

COVID-19, Data Warehousing, Geographic Mapping

Introduction

Daily monitoring of COVID-19 epidemic remains a significant challenge for hospitals and city administrations worldwide for over a year now. Important decisions are to be taken, according to the results of this tracking, varying from bed reallocations and personnel rotation in the hospital to full-agglomeration lockdown. Analyzing the data available at the local level is necessary to better understand the epidemic's dynamics and make predictions, reliable prognosis being the ultimate goal. The most relevant and simple data necessary for this monitoring are patients' addresses and laboratory tests' results including PCR tests. In Grenoble Alps University Hospital (CHUGA) this data is accessible via its Clinical Data Warehouse (CDW) PREDIMED [1], regrouping data produced by numerous hospital's professional software.

With PREDIMED data, 3 main use cases were identified by the decision makers. The first and the second use cases consist in daily monitoring of the origin (home addresses) of, respectively, people hospitalized with COVID-19 and people tested in the hospital laboratory for COVID-19, for the early detection of clusters and close surveillance of severe COVID-19 forms. The third use case represents detection of neighborhoods with the most emerging cases to adapt local prevention and information policies and define target populations for specific actions.

Methods

For the first use case of hospitalized COVID-19 patients, our daily cohort is based on the patients' list manually produced in the hospital for the regional authorities. This list is a quality-checked compilation of information sent up by all medical units.

For the second use case of people tested for COVID-19 in CHUGA we use hospital laboratory test results. To estimate the sampling distribution in each area, we take into account both negative and positive test results as getting tested suggests having symptoms.

For the first two use cases, to produce maps, we enrich the data by patients' addresses located via a geocoding service provided by the French government [2]. Geographical points are blurred in the visualization tool by adding random weights on latitude and longitude to avoid direct re-identification, while precise addresses are used in cluster detection algorithms. A prospective Poisson space-time scan statistic [3,4] implemented in SaTScan [3,5] is used to detect clusters of COVID-19 as Desjardins et al. did in early 2020 at the county level in the U.S. [6]. Clusters are computed at the communal and the sub-communal/administrative district level defined by the French National Institute of Statistics and Economic Studies (INSEE). For each geographic region, we consider the population according to the 2016 population census.

The maps from the two first cases are used for the third use case, combined with open data available for administrative districts in France [7] in terms of population density, demography descriptors (age groups, gender distribution), income, education, social vulnerability, etc.

Results

For the first use case maps as shown in Figure 1 were produced twice a week. As a result, we observed the origin of the hospitalized COVID-19 patients for the last four weeks, color-coded by the hospitalization date.

For the second use case maps as shown in Figure 2 were produced daily and updated five times a day. The following clustering parameters were used: analysis over a rolling period of 42 days, maximum reported diameter of 4 km, maximum time window covering 17% of the study period (i.e. 7 days), minimum cluster duration of 2 days, minimum number in a cluster of 3 confirmed cases. This space-time analysis helped to rapidly detect emerging clusters and declare them to regional authorities but also to inform Grenoble mayor at least twice a week.

For the third use case, data analysis was crucial to organize, in close collaboration with Grenoble City Hall, massive screenings on asymptomatic population in two Grenoble districts. Specific surveys were conducted on the residents to identify the most probable contamination sources. In Figure 3, per district, we color-code the indicator of the density of population demographically vulnerable to hospitalization with COVID-19: a scalar product between the probability of a person of a certain age group to be hospitalized and the population density per age group. This indicator matches well the real-life cases' distribution shown with dots. According to the maps' analysis, the key element explaining the risk of hospitalization is the population density, and not the individual risk related to age. On the other hand, social vulnerability plays an important role in the risk of developing severe forms. Finally, no impact of the education level was found.

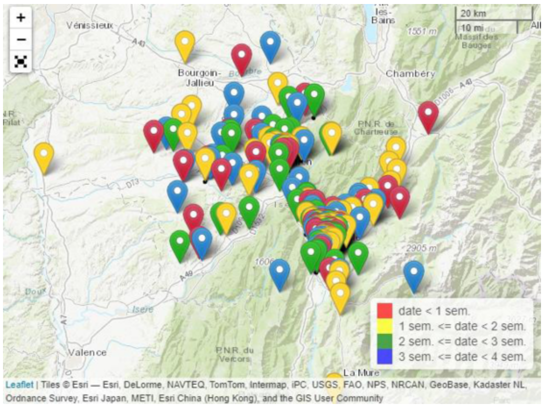

Figure 1. First use case: COVID-19 hospitalizations for 4 weeks preceding April 19th, 2021.

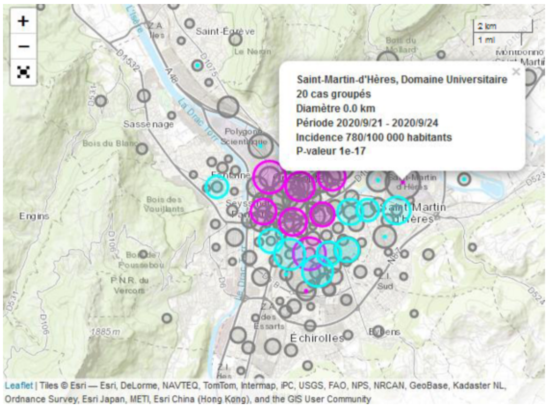

Figure 2. Second use case: episode of the 2nd COVID-19 wave, data for September 25th 2020, a cluster detected on campus.

All the maps produced were used for the intra-hospital analysis as well as for the regional authorities.

Conclusions

Using PREDIMED allowed for reliable real-time access to information on the COVID-19 epidemic's progress for the hospital's crisis unit and decision makers. This spatial and chronological analysis of hospital data was extremely useful both for the hospital management and the local authorities and allowed for fruitful collaborations between them. Two massive population screenings were organized based on the analysis of PREDIMED data combined with the social and demographic characteristics of the population. PREDIMED now has participated in developing epidemic's prevention and screening strategies.

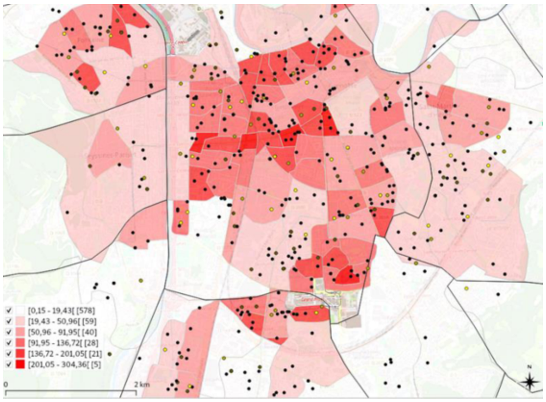

Figure 3. Third use case: population vulnerable to hospitalization.

This work will be pursued for the nosocomial COVID-19 detection, spreading via patients but also through caregivers. Another upcoming work is spatial analysis of variants of SARS-CoV-2, the virus responsible for the COVID-19. Automatic detection of COVID-19 patients instead of manual gathering of information and including air pollution data into analysis is work in progress as well.

Current work could easily be adapted for the seasonal epidemics such as influenza.

Acknowledgements

The authors thank the Numeric Services Department of CHUGA, the SIVIC team and the City of Grenoble.

References

[1] S. Artemova et al., «PREDIMED: Clinical Data Warehouse of Grenoble Alpes University Hospital.,» *Studies in health technology and informatics,* vol. 264, p. 1421–1422, 2019.
[2] French National Address Base, https://adresse.data.gouv.fr
[3] Kulldorff M, Athas WF, Feurer EJ, Miller BA, Key CR. Evaluating cluster alarms: a space-time scan statistic and brain cancer in Los Alamos, New Mexico. Am J Public Health. 1998 Sep; 88(9):1377-80.
[4] Kulldorff, M. (2001), Prospective time periodic geographical disease surveillance using a scan statistic. Journal of the Royal Statistical Society: Series A (Statistics in Society), 164: 61-72. https://doi.org/10.1111/1467-985X.00186
[5] https://www.satscan.org/
[6] Desjardins, M. R., Hohl, A., Delmelle, E. M. Rapid surveillance of COVID-19 in the United States using a prospective space-time scan statistic: Detecting and evaluating emerging clusters. Applied Geography (2020), 118, 102202.
[7] https://www.insee.fr/fr/information/2017372

Address for correspondence

Svetlana Artemova (SArtemova@chu-grenoble.fr)