A first insight about spatial dimension of COVID-19: analysis at municipality level

Josep-Maria Arauzo-Carod

Universitat Rovira i Virgili, Departament d'Economia (ECO-SOS), 43204, Reus, Catalonia, Spain
Address correspondence to Josep-Maria Arauzo-Carod, E-mail: josepmaria.arauzo@urv.cat

ABSTRACT

Background This paper is about spatial patterns of by corona virus disease-2019 (COVID-19).

Methods Using data for the first 21 weeks from municipalities in Catalonia, we analyse whether reported positive cases appear randomly or following some kind of spatial dependence. Global and local measures of spatial autocorrelation are used.

Results There are some clusters alongside Catalan municipalities that change over time.

Conclusions Use of spatial analysis techniques is suggested to identify spatial disease patterns and to provide spatially disaggregated public health policy recommendations.

Background

Corona virus disease-2019 (COVID-19) first appeared in December 2019 in Wuhan (China) and rapidly expanded all over the world, being the first example of a global pandemic in the 21st century threating dramatically public health and stopping economic activity worldwide.

Understanding spatial patterns of a disease like COVID-19 is of extreme importance to better allocate (scarce) public resources devoted to control the pandemic and to avoid its negative effects.[1] Identifying local clusters where the virus moves quickly is a key issue to implement firewalls by isolating these areas and concentrating public health facilities there.[2] Some of the lessons given by this pandemic are the social and mobility measures to be taken by public authorities.[3] To succeed, these measures need to be formulated upon data analysis[4] if public authorities want to minimize their negative economic impacts, as they may imply compulsory lockdowns and temporary firms' closures. Some of these risks were anticipated from theoretical exercises in which such situations were simulated,[5,6] although before COVID-19 there were no serious expectations of such a quick and spatially spread contagion. During previous pandemics, authorities of the most affected countries, as China, took preventive measures in order to control outbreaks, especially in dense urban areas with extensive public transport systems that could help to spread severe acute respiratory syndrome (SARS).[7] In spite of these preventions, there is evidence[8] showing that for Chinese case spatial autocorrelation of disease was relevant, and that population density triggered contagion process.[9]

This paper aims to contribute to empirical literature by analysing the first 21 weeks of the evolution of the pandemic in Catalonia, a region situated in North-East of Spain whose capital is Barcelona. Our approach is at local level because it is essential to identify in a precise way the places where the COVID-19 appears and how it moves to neighbouring areas. As Catalan Government started to collect data at local level from the very beginning of the spread of the disease, they got valuable information that could be used for the identification of local clusters of COVID-19. This strategy may allow public authorities to take quick measures to prevent the dissemination of the pandemic. As literature about spatial patterns of coronaviruses focus mainly in China, empirical evidence for European cities is still scarce, especially for rural areas. This lack of analysis is an important shortcoming, as urban and social structures that may shape spatial processes in Chinese cities do not necessary coincide with those in Europe or in other countries, so applied analyses using data from different areas may help to a better understanding of such spatial contagion processes.

Josep-Maria Arauzo-Carod, Professor and Director and Chair ECO-SOS

Methods

There are different potential data sources that can be used for public health purposes to understand the dissemination process of the pandemic. The most obvious ones are positive cases and number of deaths. Both measures have some advantages and disadvantages: number of positive cases allows to have a more immediate measure and to capture a higher number of cases but, at the same time, not all real cases are reported (i.e. there is no full availability of diagnostic tests); number of dead people seems to be more accurate but (i) some deaths caused by the virus are not properly identified, (ii) there is a time lag between the disease starts and the phenomenon (dead) is observed and (iii) only a fraction of people suffering from COVID-19 finally dead. According to these issues, we consider that is much more accurate to use the positive cases reported, even if there is an important number of positive cases not reported, but it is reasonable to assume they do not have a spatial bias.

This dataset has been created by the Health Department of Catalan Government. Data are collected on a daily basis and are provided for the 947 municipalities in Catalonia. Data collection started on 25 February 2020, and our sample ranges from this date to 19 July 2020. To simplify the analysis and to avoid data bias due to lack of daily updating, we grouped cases into weeks, ranging from 25 February 2020 to the week starting on 13 July 2020, and we focussed on municipality differences, as our main interest is on the spatial patterns of the spread of the disease. Departing from previous approaches and taking into account the type of data available for the Catalan case, we use global (i.e. Moran's I) and local spatial autocorrelation (LISA) measures, because they allow to identify both general clustering processes as well as local agglomerations in terms of the disease. Concretely, spatial autocorrelation tools[14] indicate whether values of a given variable (i.e. positive cases of COVID-19) are correlated across nearby locations (see Supplementary material for details).

Results

Our dataset comprises 124 157 positive cases reported in Catalonia between 25 February and 19 July 2020. The spread of the disease is closely related with the distribution of the population, as more populated cities are those where the number of positive reported cases is higher. This association is easily demonstrated with a simple correlation analysis among weekly number of cases and total population in 2019 (see Supplementary Table S1). As we want to isolate the data from this urban size effect, we have (re)calculated number of positive cases in relative terms to the total population at local level (see Supplementary Table S2) and now correlation disappears at all, which indicates that there are other issues, apart from the size of the population, that are related to the number of positive cases of COVID-19. Then, number of cases is higher in big cities just because number of inhabitants there, not because the intensity of the disease is stronger in bigger urban areas. This is a relevant point as it is commonly assumed that, as potential social interactions are higher in densely populated urban areas, hazard rates should be also higher there.

The aforementioned linkage between urban size and number of positive cases is also shown in Fig. 1 were total positive cases and total population are displayed. Concretely, it can be observed that most of positive cases concentrate in metropolitan areas of Barcelona (where most of Catalan population live) and Lleida (due to a serious outbreak from June to July). As we guess that spatial distribution of positive cases is not directly related to population in absolute terms but to other variables, we also show previous data but in relative terms (i.e. positive cases per 100 000 inhabitants). It can now be observed that the spatial pattern changes completely and the positive cases have an erratic distribution without a clear pattern, although they are higher in more dense areas and also in some inland rural municipalities around Lleida.

A weekly approach on the disease shows (Fig. 2) that (i) territorial patterns differ considerably and that (ii) there are several waves across time focusing in different areas, especially from late June. These results have important policy implications, as public administrations should be able to react quickly and to provide healthcare facilities and specific social isolation measures (e.g. lockdown) at specific areas, but not on a general basis in all regions.

We have calculated Moran's I on a weekly basis (see Supplementary Table S3) and results show that there is no global spatial autocorrelation for the positive cases at local level (no matter if these are measured in absolute terms or relative to local population), except from end June due to the outbreak in Lleida. Our results differ from those for Chinese provinces[11] during current COVID-19 outbreak where strong spatial autocorrelation levels were reported, but as they were obtained using huge spatial areas this approach may suffer from a scale effect. On the contrary, when comparing our results with other studies using data for smaller spatial units,[12] then both are much more similar, suggesting a related trend.

As this lack of global spatial autocorrelation may suggest a spatial random pattern of positive cases, but it could also hide the existence of several processes of local spatial autocorrelation in specific areas, we have calculated a LISA index using both measures of positive cases (in absolute and in relative terms).
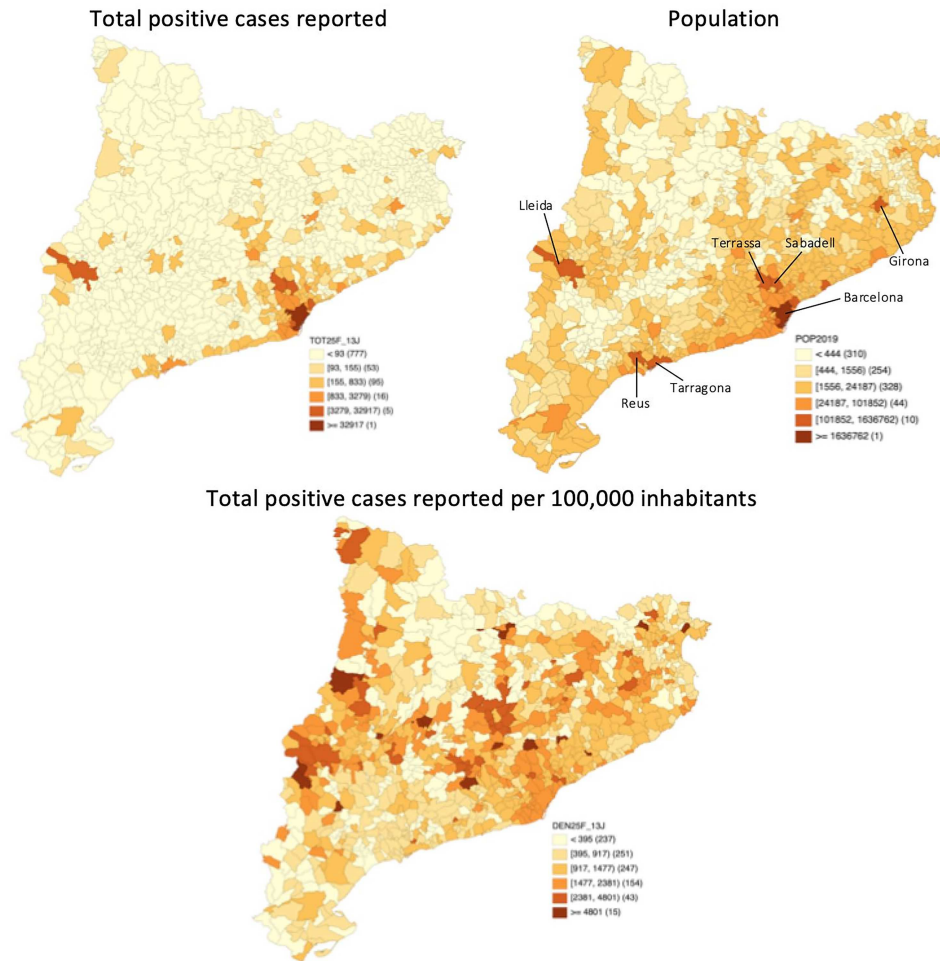
Fig. 1 Spatial linkages between positive cases of COVID-19 and population. Note: population refers to total inhabitants in 2019; density of positive cases is calculated as number of case per 100 000 inhabitants.

Figure 3 shows local spatial autocorrelation data for the whole period. Concretely, red areas show high–high spatial autocorrelation (i.e. municipalities with high levels of positive cases surrounded by municipalities with high levels of positive cases), dark blue areas show low–low spatial autocorrelation (i.e. municipalities with low levels of positive cases surrounded by municipalities with low levels of positive cases), light red areas show high–low spatial autocorrelation (i.e. municipalities with high levels of positive cases surrounded by municipalities with low levels of positive cases), light blue areas show low–high spatial autocorrelation (i.e. municipalities with low levels of positive cases surrounded by municipalities with high levels of positive cases), and light grey areas indicate not significant spatial autocorrelation results.
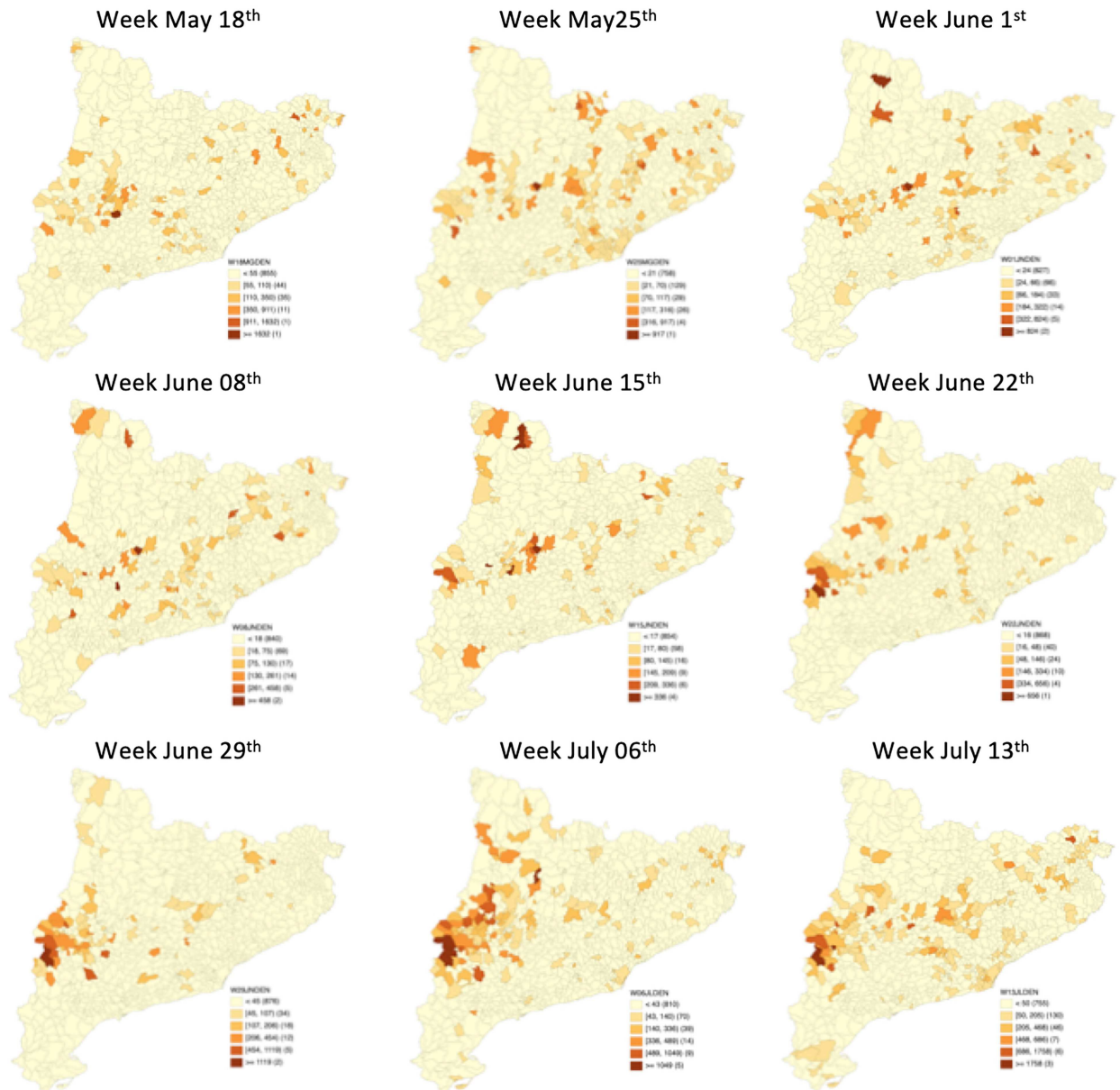
Results shown in Fig. 3 are quite different depending on the measure, which implies that to analyse positive cases in absolute or in relative terms would lead to quite different conclusions and, consequently, to different policy measures in terms of public health. The most obvious option seems to take the absolute measure (i.e. number of cases). According to that measure, there is a clear cluster of positive cases of COVID-19 in an around the metropolitan area of Barcelona, the region where most of people live and work. If this was true, appropriate measures should be to concentrate public efforts and resources in that area. Nevertheless, according to the relative measure (i.e. density of positive cases) the previous Barcelona's cluster is much smaller in size and is not the only one, as some other clusters appear around small and medium municipalities (in counties such as Noguera, Segarra, Bages and Conca d'Òdena) and at the south side of the city of Lleida (Baix Segrià). In this sense, Catalan Government was (partially) aware of these circumstances and decided to organize a police block to the Conca d'Òdena that prohibited
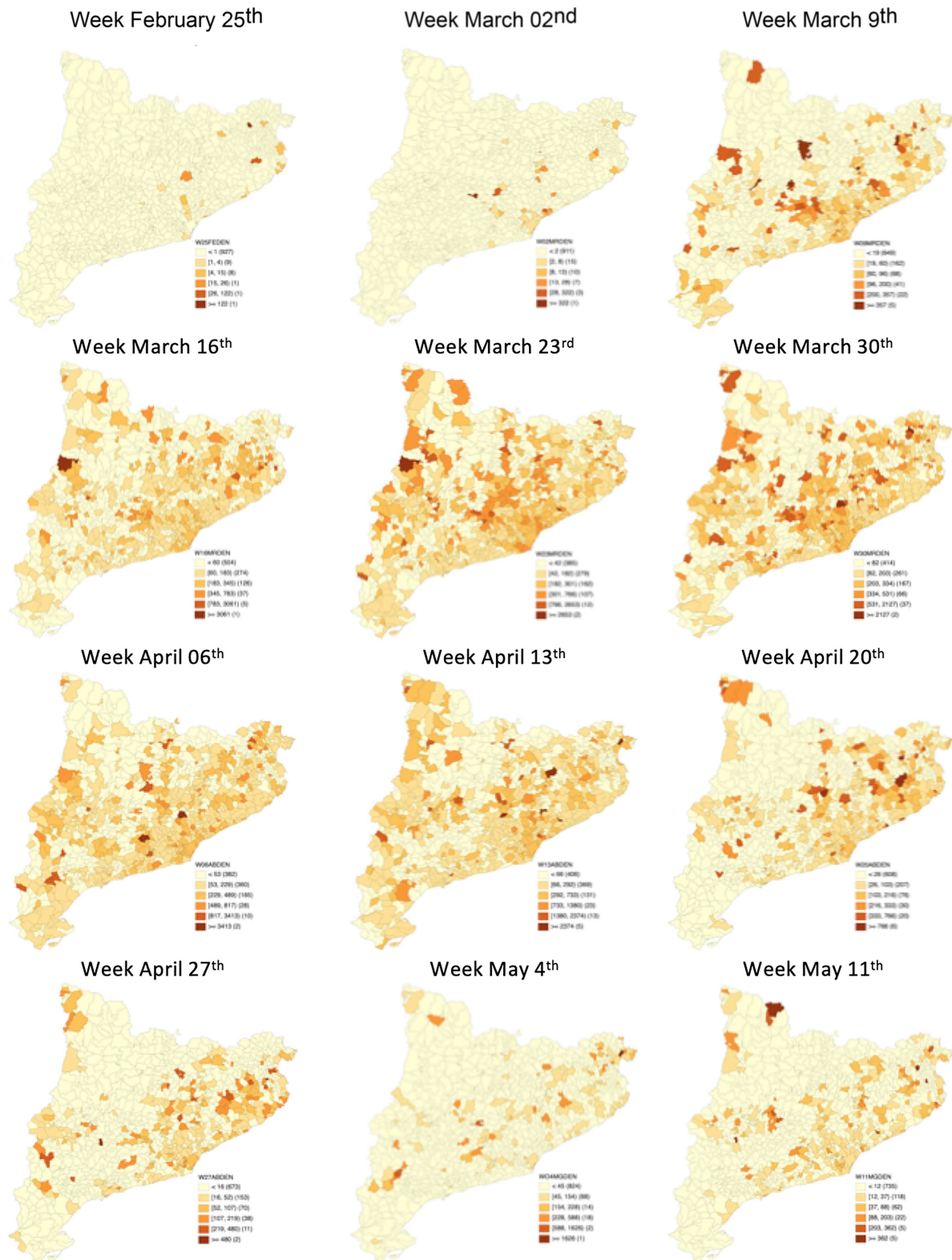
Source: own elaboration.

Fig. 2 Total positive case of COVID-19 reported per 100 000 inhabitants (weekly). Note: density of positive cases is calculated as number of cases per 100 000 inhabitants (according to population data of 2019).

Week May 18th   Week May25th   Week June 1st

Week June 08th   Week June 15th   Week June 22th

Week June 29th   Week July 06th   Week July 13th

Source: own elaboration.

Fig. 2b continued

entries and exits into the area during 3 weeks in March, but no special measures were taken in other clusterized areas as Baix Segrià and Noguera until July. This is a key issue, as the Government finally understood that public health measures had to be decided at a very disaggregated spatial level, such as municipalities and even neighbourhoods, against initial plans to design measures at much bigger areas. In this sense, other studies[14] reach a similar conclusion about the importance of detailed spatial data. Concretely, when analysing SARS in Hong Kong different spatial methods were tested and it was concluded that (p. 1556) '(...) the selection of such districts for quarantine could have been better informed by these ellipses indicating directional bias and associated physical distance in disease transmission'. This is a crucial issue, as in most of cases quarantine areas are designed using big administrative spatial units without properly identifying real disease's clusters.

Another key point when dealing with such a pandemic is the capacity to react quickly to move public resources to the most affected areas. To illustrate this point, Fig. 4 shows
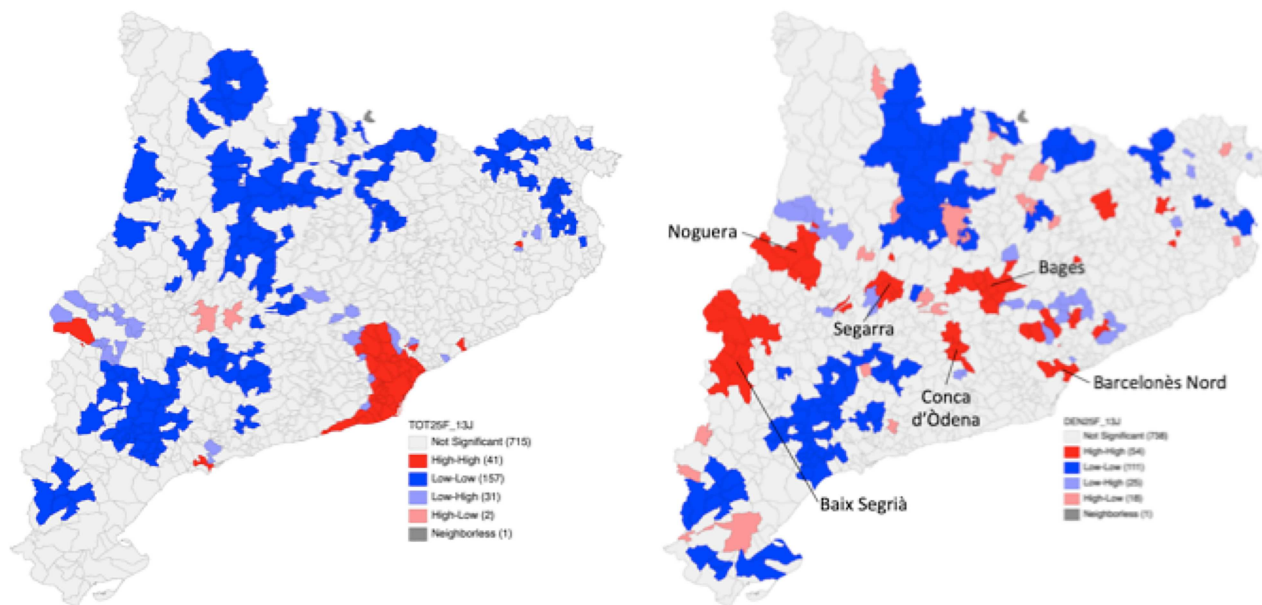
Source: own elaboration.

Fig. 3 LISA of reported positive cases. Note: density of positive cases is calculated as number of cases per 100 000 inhabitants (according to population data of 2019).

weekly data about local spatial autocorrelation of density of positive cases. In this sense, during the first 2 weeks of the pandemic the number of reported positive cases was very low and without a clear spatial pattern, but from the third week the cluster of the Conca d'Òdena appeared and it was remarkable until the seventh week. Similarly, the cluster in Baix Segrià appeared in 17th week and disseminated quickly to neighbour areas. In the meanwhile, smaller clusters showed up in several areas, without any specific public actions implemented there.

Discussion

Main findings of this study

This paper has summarized the evolution of the first 21 weeks of the COVID-19 disease in Catalonia (Spain) through local spatial autocorrelation measures. We have shown that the dissemination of the disease has a spatial pattern that deserves to be analysed. Our results show a very important heterogeneity in terms of spatial growth of the pandemic that, in turn, requires specific policies to be adopted at very narrow spatial levels.
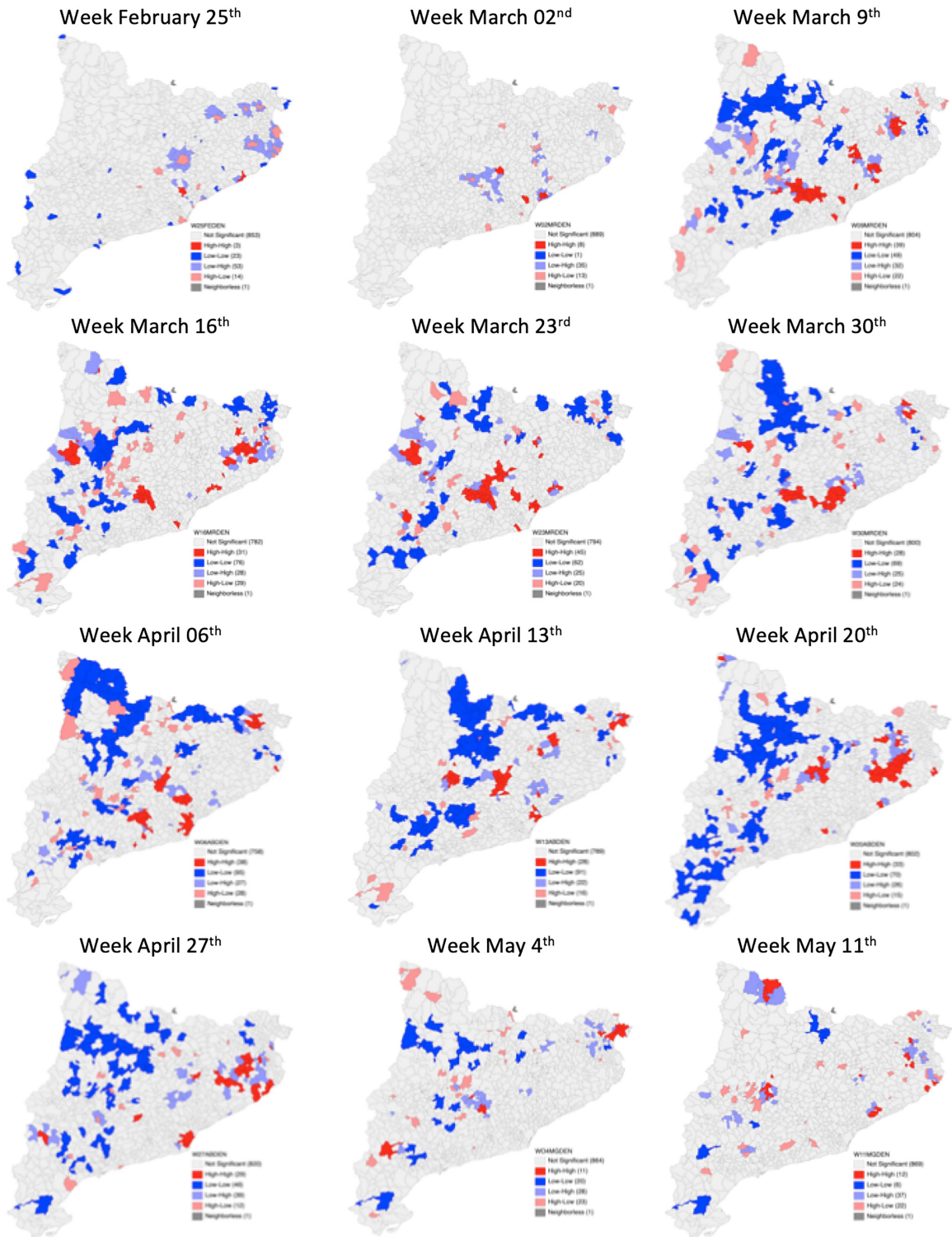
What is already known on this topic

Previous empirical research dealing with spatial dimension of coronaviruses include a wide diversity of methodologies such as correlation analysis for China[10] (COVID-19); global

(Moran's I) spatial autocorrelation for China[11] (COVID-19), for Beijing[12] (SARS) and for Hong Kong[13] (SARS); LISA for Beijing[12] (SARS); joint-count spatial statistic for Beijing[9] (SARS); nearest neighbour analysis for Hong Kong[13] (SARS); nearest neighbour hierarchical clustering for Beijing[7] (SARS); bayesian maximum entropy (BME) models for China[8] (SARS); space–time clustering for the USA[2] (COVID-19); or spatial panel data model for China[1] (COVID-19), but the empirical evidence for Europe is more scarce.

What this study adds

On a similar analysis about SARS pandemic in Beijing's districts,[12] it is shown that local clusters evolve quickly as disease spreads. Although spatial scale of both analyses differs, they present similar results in terms of this spatial variability regarding high–high clustering at local level, a circumstance suggesting the existence of complex transmission mechanisms. In general terms, results from Fig. 4 indicate that it is key to monitor the spatial evolution of the pandemic to isolate significantly affected areas and to concentrate public resources in these areas. Currently, most of attention of policy makers is devoted to borders' control measures aiming to stop international spread of the pandemic,[15] especially in terms of international air traffic.[10] Our results suggest that analogous measures taken at lower spatial levels are also needed in order to prevent similar spread inside national borders through, for instance, public transport systems.

Week February 25th   Week March 02nd   Week March 9th

Week March 16th   Week March 23rd   Week March 30th

Week April 06th   Week April 13th   Week April 20th

Week April 27th   Week May 4th   Week May 11th

Source: own elaboration.

Fig. 4 LISA of reported positive cases (Weekly data, density). Note: density of positive cases is calculated as number of cases per 100 000 inhabitants (according to population data of 2019).
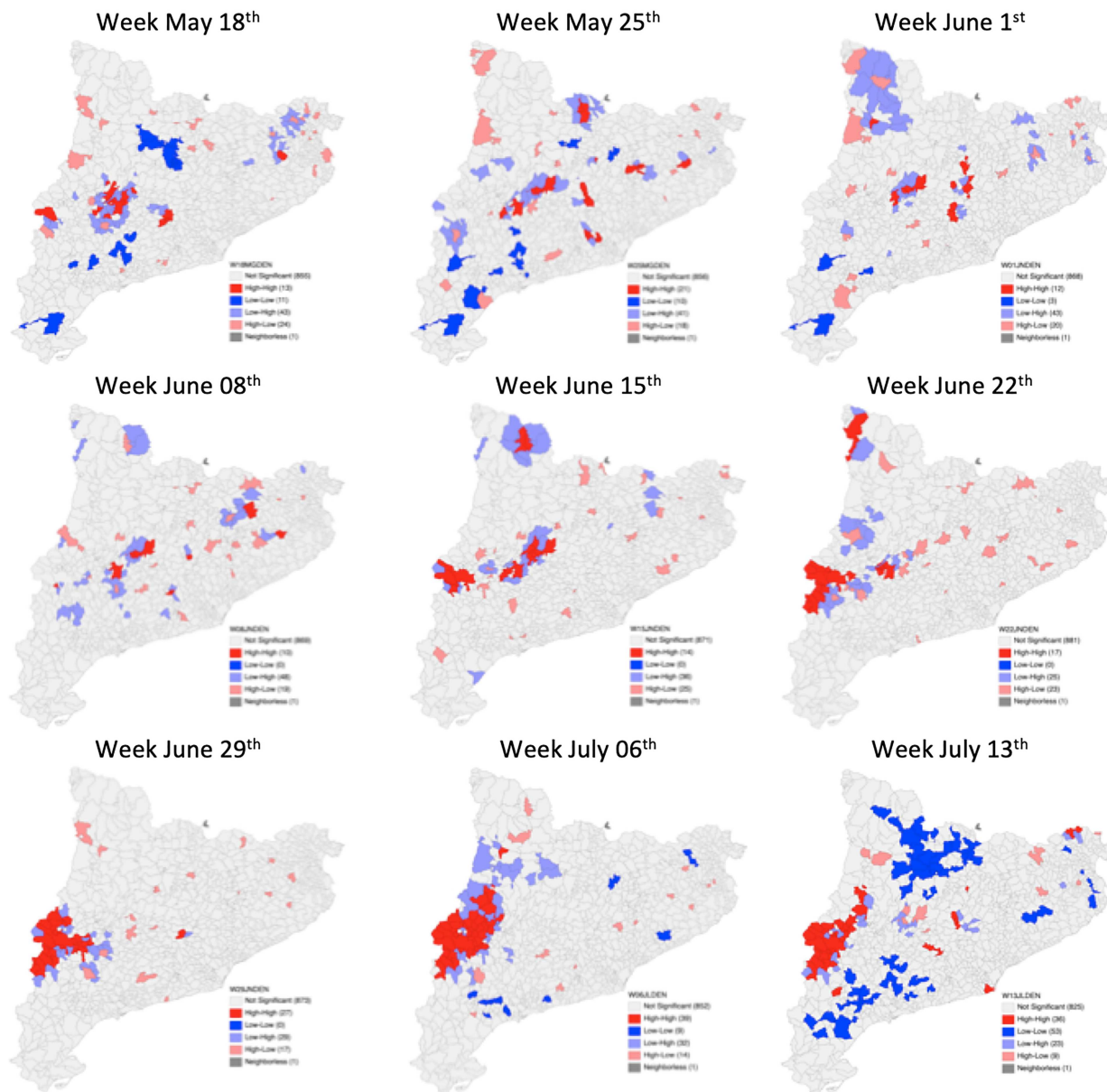
Source: own elaboration.

Fig. 4b continued

Limitations of this study

As this study uses municipality data its findings are of special interest for small municipalities but, unfortunately, less useful for big ones that contain several neighbourhoods. In this sense, availability of COVID-19 data at intra-urban level may allow to provide similar insights for bigger cities.

Conclusion

It is evident that the evolution of COVID-19 is not easy to predict and control and that at the very beginning public health measures are implemented without a clear knowledge about the most appropriate strategies but, at the same time, the efficiency of policy responses increases with the availability and use of data about the spatial dimension of COVID-19. This is why it is so important to map reported cases of the disease. Public administrations need to carry out an active surveillance of the evolution of COVID-19 in order to take better and quicker decisions that allow them to use public resources in a more efficient way, and spatial analysis tools are relevant instruments to understand the spatial diffusion process of the pandemic. Finally, our results suggest that it is

needed to invest public resources in data monitoring systems in order to better collect and process public health data.

Supplementary data

Supplementary data are available at the *Journal of Public Health* online.

Funding

Author contributions

J.M.A.C. collected, analysed and interpreted the data. J.M.A.C. wrote, revised and submitted the manuscript.

References

1 Guliyev H. Determining the spatial effects of COVID-19 using the spatial panel data model. *Spatial Stat* 2020;**38**(1–10).

2 Desjardins MR, Hohl A, Delmelle EM. Rapid surveillance of COVID-19 in the United States using a prospective space-time scan statistic: detecting and evaluating emerging clusters. *Appl Geogr* 2020;**118**(1–7).

3 Dalton C, Corbett S, Katelaris A. Pre-emptive low cost social distancing and enhanced hygiene implemented before local COVID-19 transmission could decrease the number and severity of cases. *SSRN* **3549276**:2020.

4 Xu B, Gutierrez B, Mekaru S *et al.* Epidemiological data from the COVID-19 outbreak, real-time case information. *Sci Data* 2020;**7**(106):1–6.

5 Zhang N, Huang H, Su B *et al.* A human behaviour integrated hierarchical model of airborne disease transmission in a large city. *Build Environ* 2018;**127**:211–20.

6 Cooley P, Brown S, Cajka J *et al.* The role of Subway travel in an influenza epidemic: a new York City simulation. *J Urban Health* 2011;**88**(5):982–95.

7 Wang J, McMichael AJ, Meng B *et al.* Spatial dynamics of an epidemic of severe acute respiratory syndrome in an urban area. *Bull World Health Organ* 2006;**84**(12):965–8.

8 Cao C, Chen W, Zheng S *et al.* Analysis of spatiotemporal characteristics of pandemic SARS spread in mainland China. *Biomed Res Int* 2016;1–12.

9 Meng B, Wang J, Liu J *et al.* Understanding the spatial diffusion process of severe acute respiratory syndrome in Beijing. *Public Health* 2005;**119**:1080–7.

10 Lau H, Khosrawipour V, Kocbach P *et al.* The association between international and domestic air traffic and the coronavirus (COVID-19) outbreak. *J Microbiol Immunol Infect* 2020; forthcoming.

11 Kang D, Choi H, Kim JH *et al.* Spatial epidemic dynamics of the COVID-19 outbreak in China. *Int J Infect Dis* 2020;**94**: 96–102.

12 Cao ZD, Zheng DJ, Zheng XL *et al.* Spatio-temporal evolution of Beijing 2003 SARS epidemic. *Sci China Earth Sci* 2010;**53**: 1017–28.

13 Lai PC, Wong CM, Hedley AJ *et al.* Understanding the spatial clustering of severe acute respiratory syndrome (SARS) in Hong Kong. *Environ Health Perspect* 2004;**112**:1550–6.

14 Moran P. The interpretation of statistical maps. *J R Stat Soc* 1948;**59**:185–93.

15 Wells CR, Sah P, Moghadas SM *et al.* Impact of international travel and border control measures on the global spread of the novel 2019 coronavirus outbreak. *PNAS* 2020;**31**, **117**:7504–9.