Postprint version

LÓPEZ-GAY, Antonio; SPIJKER, Jeroen; (…); MÓDENES, Juan A. et al. (2021) "Sociodemographic determinants of intraurban variations in COVID-19 incidence: the case of Barcelona". *Journal of Epidemiology & Community Health* (Online First: June 22, 2021): 1-7 (ISSN: 0143-005X). http://dx.doi.org/10.1136/jech-2020-216325

Title page: Sociodemographic determinants of intra-urban variations in COVID-19 incidence. The case of Barcelona

1. Antonio López-Gay

tlopez@ced.uab.cat

Department of Geography, Universitat Autònoma de Barcelona, Barcelona, Spain

Centre d'Estudis Demogràfics, Barcelona, Spain

Postal address: Carrer de Ca n'Altayó, Edifici E2. Universitat Autònoma de Barcelona. 08193 Bellaterra, Barcelona/Spain

2. Jeroen Spijker

Centre d'Estudis Demogràfics, Barcelona, Spain

3. Helen V. S. Cole

Barcelona Lab for Urban Environmental Justice and Sustainability, Universitat Autònoma de Barcelona, Barcelona, Spain

4. Antonio G. Marques

Department of Signal Theory and Communications, King Juan Carlos University, Madrid, Spain

5. Margarita Triguero-Mas

Barcelona Lab for Urban Environmental Justice and Sustainability, Universitat Autònoma de Barcelona, Barcelona, Spain

6. Isabelle Anguelovski

Barcelona Lab for Urban Environmental Justice and Sustainability, Universitat Autònoma de Barcelona, Barcelona, Spain

7. Marc Marí-Dell'Olmo

Agència de Salut Pública de Barcelona, Barcelona. Spain

CIBER Epidemiología y Salud Pública (CIBERESP), Madrid. Spain.

Institut d'Investigació Biomèdica (IIB Sant Pau), Barcelona. Spain.

8. Juan A. Módenes

Department of Geography, Universitat Autònoma de Barcelona, Barcelona, Spain

Centre d'Estudis Demogràfics, Barcelona, Spain

9. Dolores Álamo-Junquera

Agència de Salut Pública de Barcelona, Barcelona, Spain

10. Fernando López-Gallego

Heterogeneous Biocatalysis Laboratory, CICbiomaGUNE

11. Carme Borrell

Agència de Salut Pública de Barcelona, Barcelona, Spain

CIBER Epidemiología y Salud Pública (CIBERESP), Madrid, Spain

Word count (excluding title page, abstract, references, figures and tables): 2,997

Sociodemographic determinants of intra-urban variations in COVID-19 incidence. The case of Barcelona.

Abstract

Background

Intra-urban sociodemographic risk factors for COVID-19 have yet to be fully understood. We investigated the relationship between COVID-19 incidence and socio-demographic factors in Barcelona at a fine-grained geography.

Methods

This cross-sectional ecological study is based on 10,550 confirmed cases of COVID-19 registered during the first wave in the municipality of Barcelona (population 1.64 million). We considered 16 variables on the demographic structure, urban density, household conditions, socioeconomic status, mobility, and health characteristics for 76 geographical units of analysis (neighborhoods), using a Lasso analysis to identify the most relevant variables. We then fitted a multivariate Quasi-Poisson model that explained the COVID-19 incidence by neighborhood in relation to these variables.

Results

Neighborhoods with: (i) greater population density, (ii) an aged population structure, (iii) a high presence of nursing homes, (iv) high proportions of individuals that left their residential area during lockdown and/or (v) working in health-related occupations were more likely to register a higher number of cases of COVID-19. Conversely, COVID-19 incidence was negatively associated with (vi) percentage of residents with post-secondary education and (vii) population born in countries with a high Human Development Index.

Conclusion

Like other historical pandemics, the incidence of COVID-19 is associated with neighborhood socio-demographic factors with a greater burden faced by already deprived areas. Because urban social and health injustices already existed in those geographical units with higher COVID-19 incidence in Barcelona, the current pandemic is likely to reinforce both health and social inequalities, and urban environmental injustice all together.

Sociodemographic determinants of intra-urban variations in COVID-19 incidence. The case of Barcelona.

Main text

1. INTRODUCTION

Global flows of people, resources, and capital involved in the production and maintenance of urban life facilitates the spread of infectious disease and the emergence of pandemics.[1] After appearing in China in late 2019, the first cases of COVID-19 were confirmed in Spain and elsewhere in Europe, by late January 2020. Previous research on virus transmission has shown that socioeconomic and cultural factors at the individual, household and neighborhood levels, are essential mechanisms for community spread of the virus.[2 3]

Individual-level risk factors such as gender, age or race/ethnicity are known to influence infectious diseases incidence,[4 5] including COVID-19.[6 7] Although infection rates are similar between genders, men are more likely to have comorbid conditions (such as hypertension, diabetes, obesity and cardiovascular diseases) that are also risk factors associated with worse COVID-19 outcomes.[8 9] Women, however, are often more exposed because of their more frequent dedication to care professions.[10] Older people are also known to be more susceptible to COVID-19 and show higher fatality rates.[11] In contrast, the role that children play in disease transmission is still unclear as they are rarely the index case[12] and are less likely to transmit COVID-19 to adults.[13] On the other hand, school closures are likely to have led to increased childcare by seniors,[14] potentially increasing risk of transmission.

Individual socioeconomic factors such as level of education, income, employment status and type of occupation are also thought to impact risk of COVID-19. Although initial COVID-19 outbreaks emerged from international (business) travel and winter holidays,[15] subsequent trends reveal that those working in specific occupations, especially front-line, ”essential” jobs in health, care, retail and hospitality, are more at risk for infection.[16 17] Individuals living in poverty and other marginalized populations are more susceptible to infectious diseases.[5] For instance, in the US context, racialized minorities (especially African Americans) are vulnerable social groups that exhibit higher than average rates of infectious diseases. This has been attributed to systematic and interpersonal racism, and poorer access to health care facilities and other health-promoting resources.[18]

Public health researchers have also long acknowledged the importance of neighborhood-level socio-demographic and physical characteristics—including racial and economic residential segregation, and the spatial distribution of affordable and fresh food, or public transport—for understanding health outcomes.[19 20] Structural contexts and neighborhood environments can therefore create uneven poor living conditions and lasting environmental injustices for lower-income or immigrant residents living in certain areas of a city,[21] resulting in health inequity by neighborhood. In fact, during the 1918 influenza pandemic, researchers already found a significant association between disease transmissibility and neighborhood-level social characteristics such as population density, illiteracy, and unemployment.[4]

Emerging research on COVID-19 shows similar patterns and pathways.[22] For example, people living in denser neighborhoods, with poor and overcrowded housing conditions have an elevated risk of infection as social contact in these living scenarios is more likely.[11 23] Urban connectivity, mobility and the mode of transport also play an important role in the spread of COVID-19.[24] At the neighborhood level, greater use of private motor vehicles and less public transport mobility means less exposure to infection.[25] Likewise, infection rates may be lower where part of the (more mobile, international and national) population was able to leave before movement restrictions or where a higher proportion of people was able to work from home during lockdown. Conversely, rates may be higher where more essential workers live (occupations that are over-represented by women and immigrants from low-income countries) as they are more likely to commute. Overall, higher mortality rates from COVID-19 are associated with poorer neighborhood conditions, including a scarcity of health care facilities.[26] The number of nursing and retirement homes has also been associated with a greater number of infections in the neighborhood.[27]

To date, COVID-19 research on spatial variations has been mainly set at the national or sub-national levels. At this level of analysis, it is very difficult to disentangle the different intervening factors behind risks and exposures to COVID-19 as this approach fails to reveal the diverse patterns within these larger geographies. There is therefore a need to focus on geographically smaller units to allow to better account for confounding factors[28] and enhance the predictive accuracy and interpretability of the resulting statistical model. As of late 2020, neighborhood-level studies of socio-spatial inequality in COVID-19 infection and mortality have primarily focused on the United States and United Kingdom.[29 30] Very little is known about such patterns in mainland Europe,[31] especially so in much denser and mixed-use urban environments. To address these shortfalls, we investigated the relationship between COVID-19 incidence and a comprehensive diversity of intra-urban socio-demographic factors in Barcelona, Spain.

2. METHODS

2.1. Study design and study population

This cross-sectional ecological study used data from the COVID-19 Register of the Barcelona Public Health Agency. During the first wave, Spain registered one of the highest per capita number of cases in Europe, making analysis at the local scale more reliable. Barcelona became one of the initial hotspots in the country, possibly due to its international position in tourism, business, education and research. [32]

Our study included 10,550 laboratory-confirmed cases of COVID-19 in Barcelona between March 9 and May 3, 2020. We selected these dates to focus on the first outbreak of the pandemic. During this period, tests were essentially performed for those hospitalized or from specific at-risk groups, especially health-care workers, as well as residents and workers in long-term care facilities (LTCF). However, confirmed cases registered in LTCF were excluded, as test campaigns were unevenly implemented across time and space and addresses of residents correspond to those of the LTCF which do not necessarily reflect the socioeconomic position of the residents themselves.

Our geographic unit of observation is the neighborhood. We aggregated addresses of positive tested individuals by neighborhood of residence. Although the municipality of Barcelona (1.64 million inhabitants) is officially divided into 73 *barris* (Catalan for neighborhood), for statistical purposes we have followed the adaptation developed by the Spanish National Statistical Office in several studies.[33] This alternative division is based on the official administrative division, but creates more statistically robust units in terms of population size, merging the least populated with neighboring units and splitting the most populated ones, always according to urban and socio-demographic criteria. Our final division consists of 76 units (henceforth referred to as neighborhoods). They contain an average of 21,500 inhabitants and 1.3 km^2 area. These units are very diverse in terms of wealth, housing characteristics, demographic ageing and health, factors known to be associated with the spread of infectious diseases.

2.2. Intra-urban socio-demographic covariates

A total of 16 neighborhood-level indicators on demographic structure, socioeconomic status, urban and household density, mobility and health characteristics were initially chosen based on earlier established associations with COVID-19 (see Table 1 for sources, expected association with COVID-19 and summary statistics). Specifically, we included information on the proportion of (i) young people (ages 0-15) and (ii) elderly (70 years and older), and (iii) the percentage of the population aged 70+ who was male. Socioeconomic indicators included were (iv) mean income per person, (v) age-standardized ratio of population with at least post-secondary education, (vi) percentage of the population born in foreign countries with a high Human Development Index (HDI) and (vii) low HDI. We also included (viii) population density, (ix) average number of persons per dwelling and (x) people living alone. We obtained mobility data on: (xi) the availability of private transportation), and (xii) mobility during lockdown. We also captured the presence of (xiii) transient populations (measured as the rate of inhabitants automatically deregistered by the municipality, which occurs when foreign residents fail to renew their registration), as cumulative infection may be lower in areas with hypermobile groups (e.g.

international students) that were likely to leave the city due to the pandemic. We also incorporated (xiv) the number of LTCF beds per 1,000 inhabitants and (xv) the percentage of economically active population in the health sector. Lastly, we included (xvi) the life expectancy at birth as a proxy for general health status.

Table 1. Covariates used in the study: Hypothesized association with COVID-19, definitions, sources and summary statistics before transformation (when required#)

Covariate	Abbreviation	Assoc. with	Operationalisation and measurement	Period	Mean	Min.	Max	S.D.
DEMOGRAPHIC STRUCTURE								
1. Percentage 0-15	P0-15	+ [a]	% of the population aged 0-15 (ln)	2019	16.2	8.3	22.8	3.1
2. Percentage 70+	P70+	+ [b]	% of the population aged 70+	2019	12.7	7.6	18.8	2.1
3. Old-age sex ratio	OLDSEXR	men+ [c]	% of the 70+ population who is male	2019	38.8	33.5	45.7	2.3
SOCIOECONOMIC STATUS								
4. Income	INCOME	– [d]	Mean Income (€) per person (ln)	2016	15,199	7,790	29,144	4,633
5. Post-secondary education	UNI	+/– [d]	Age-standardized ratio of post-secondary educated population*	2019	1.0	0.2	1.8	0.4
6. HDI high migrants	MIGRHI	+/– [d]	% of the population born in foreign countries with a high HDI (ln)	2018	6.7	1.3	22.8	4.5
7. HDI low migrants	MIGRLOW	–/+ [d]	% of the population born in foreign countries with a low HDI (ln)	2018	17.5	6.5	46.7	7.9
DENSITY (URBAN AND HOUSEHOLD)								
8. Population density (urban)	URBDENSITY	+ [e]	Inhabitants (in hundreds) per hectare, excl. forest parks	2019	2.9	0.2	5.9	1.4
9. Persons per dwelling	HHSIZE1	+ [e]	Average number of persons per dwelling	2019	2.5	2.1	3.0	0.2
10. One person households	1PHH	– [e]	Age standardized relative ratio of persons living in one person households	2019	1.3	0.9	2.1	0.2
MOBILITY								
11. Vehicles per inhabitant	VEHICLES	– [f]	Number of registered vehicles per 100 persons aged 20+	2019	47.6	18.9	89.1	12.5
12. Mobility during lockdown	MOBILE	+ [g]	Percentage of people who left the area of residence during lockdown. Daily average, March 9-May 3 2020 (ln)	2020	10.7	5.5	19.6	2.8
13. Population deregistration	PADRON	– [h]	Automatic deregistrations from the municipal population register (expiration of the register)/1000 (ln)	2017-18	27.3	7.6	81.9	14.1
HEALTH								
14. Nursing homes (LTCF)	NURSHOME	+ [i]	Number of nursing and retirement homes (measured in beds) per 1000 people (square root)	2020	9.0	0.0	81.5	12.3
15. Health workers	EMPLHEALTH	+ [j]	Percentage of economically active population in health sector	2011	6.9	2.0	11.5	1.9
16. Life expectancy	LE_0	– [k]	Life expectancy at birth	2006-17	84.0	80.3	86.3	1.3

Notes: Expected or observed association between covariate and mortality at the neighborhood level (main unit of analysis): a) The role that children play in disease transmission is still unclear as studies have shown that they are rarely the index case.[12] On the other hand, school closures are likely to have led to increased childcare by elders (who are a high-risk group)[14]; b) [11]; c) Although infection rates are similar, the severity of disease (hence COVID-19 reporting) and death is two times greater for men than for women;[41] d) As an average for the studied period, a negative association with SES indicators is expected: higher income/educated/high HDI neighborhoods lead to lower infection rates, although the association may be positive early on the epidemic;[15] e) Denser neighborhoods and overcrowding is associated with more social contact; [11 15 23] f) More vehicles, less mobility in public transport, less exposure to infection;[24] g) More mobility in the geographical unit of analysis, implies greater exposure to infection; h) Infection rates may be lower in neighborhoods where part of the population has left; i) More nursing and retirement homes, more infections in the neighborhood;[27] j) More health personnel, the greater the infection rate. This is for two reasons: they are more exposed to the virus and have systematically been tested. Note that there are 25 geographical units with repeated values due to census aggregation. k) LE_0 As a proxy for health status (the severity of COVID-19 is positively associated with coexisting illnesses).[42]

The type of transformation that was applied to make the variable values more normally distributed is mentioned in brackets. ln refers to natural logarithm. See also sub-section 2.4.1.

*Here we first applied the age-specific percentage of post-secondary educated persons in the municipality to the residents of every neighborhood by age group. We got, then, an estimated number of post-secondary educated population for every unit of analysis, which we compared with the observed one. If the observed was greater than expected, the indicator is greater than 1, if lower, less than 1.

Data sources: 1, 2, 3, 5, 6, 7, 9, 10, 12, 13) Population Register (National Statistics and Barcelona City Council); 4) Atlas de distribución de renta de los hogares (National Statistics); 8) Urban indicators (Barcelona City Council); 11) Economic indicators (Barcelona City Council); 12) Estudio de movilidad durante el estado de alarma por COVID-19 (National Statistics); 14) Location list and characteristics of nursing homes (Consorci de Serveis Socials de Barcelona); 15) 2011 Population Census (National Statistics); 16) Health indicators (Agència de Salut Pública de Barcelona).

2.3. Statistical analyses

2.3.1. Data transformation

The distribution of each neighborhood-level socio-demographic indicator and covariate was first assessed for normality using visual inspection of QQ plots and the Smirnov-Kolmogorov test for normality. Accordingly, we log-transformed: (i) young population, (ii) income, (iii) foreigners from high HDI countries, (iv) foreigners from low HDI countries, (v) mobility during lockdown, and (vi) transient populations. We also used a square root transformation for the nursing homes variable.

2.3.2. Multiple variables model

To fit the total number of cases observed in each unit of analysis, we relied on a generalized linear model (Quasi-Poisson regression) that takes into account the total population as an offset as well as the socio-demographic variables. Given the relatively large number of covariates included in the study and the potential multicollinearity among them, we ran a Lasso analysis to automatically identify the most relevant variables.[34] In the context of generalized linear regression modeling and prediction, Lasso performs both variable selection and regularization to enhance prediction accuracy and interpretability of the statistical model. The hyperparameter of the Lasso-regularized maximum likelihood estimator was set using cross validation and, once Lasso identified the most informative variables, we fitted the final Quasi-Poisson model that explained the COVID-19 incidence for each unit of analysis considered. Finally, variable elasticities were calculated. This enables estimating the increase of cumulative incidence (and predict the total number of positive cases) for a 1% change in a particular covariate and thereby compare the effect of the different covariates.

3. RESULTS

The intra-urban geography of the COVID-19 cumulative incidence in Barcelona during the period of study reveals a strong proximity among the units with the highest and lowest values (Figure 1). Northern neighborhoods (mainly located within the districts of Nou Barris and Horta-Guinardó) have the highest incidence values, with some of them exceeding 1,000 cases per 100,000 inhabitants during the eight weeks of observation. On the other hand, the incidence in the geographical units located in the southeast of the city (i.e. historic center) is less than a third of that in the worst-affected neighborhoods.

From the initial 16 variables considered, the Lasso method selected as meaningful to explain the observed COVID-19 levels the following seven (see also supplemental material): (i) elderly, (ii) high education, (iii) foreigners from high-HDI countries, (iv) population density (urban), (v) mobility during lockdown, (vi) LTCF, and (vii) health workers. These variables are mapped in Figure 2.

Results of our Quasi-Poisson model confirm that the associations between the final selection of variables and the intra-urban COVID-19 incidence in Barcelona are all in the expected direction (Table 2). Neighborhoods that are densely populated, with a higher number of older adults, with more numerous LTCF and with higher proportions of individuals that left their area of residence during lockdown were statistically more likely to have a higher number of cases of COVID-19 during the first outbreak of the pandemic. The work in health-related occupations variable was significant at the 0.063 level. Conversely, the association with COVID-19 cases is negative with the other two socioeconomic factors: post-secondary educated residents and population born in high-HDI countries, with the second one being less relevant (note that while the cross-validation analysis of the lasso-regularized 16-variable regression deems the high-HDI variable meaningful, the p-value associated with the 7-variable regression casts doubts about its statistical significance). Considering the effect of the factors on the number of COVID-19 infections in a neighborhood of Barcelona with average characteristics, a 1% increase in older people or mobility during lockdown would lead to almost 30 extra cases, while a neighborhood with a 1% higher ratio of post-secondary educated inhabitants leads to 26 fewer cases during the observed period according to our model. We finally ran a Global Moran's I test to assess the potential spatial autocorrelation of the model's residuals, but results were not significant (see supplemental material).

Table 2. Results of the generalized linear (Quasi-Poisson regression) analysis of social and demographic factors on COVID-19 infection rates in Barcelona from March 9th to May 3rd 2020

Estimated Coefficients:

	Estimate	Standard Error	t-statistic	p-value	Elasticity#			
					Change in cumulative incidence		Equivalence in cases	
					%	95% CI	Number	95% CI
(Intercept)	-5.9826	0.3791	-15.7830	0.0000				
Percentage 70+	0.0228	0.0088	2.6000	0.0114	+0.29%	[0.07%, 0.51%]	28.0	[7.0, 49.0]
Post-secondary education	-0.2694	0.1011	-2.6660	0.0096	-0.27%	[-0.47%, -0.07%]	-26.0	[-45.0, -6.8]
HDI high migrants	-0.0901	0.0693	-1.3000	0.1979	-0.09%	[-0.22%, 0.05%]	-8.6	[-22.0, 4.4]
Population density (urban)	0.0295	0.0140	2.0990	0.0396	+0.09%	[0.01%, 0.16%]	8.2	[1.0, 15.9]
Mobility during lockdown	0.2751	0.1182	2.3280	0.0229	+0.27%	[0.04%, 0.50%]	26.4	[4.0, 48.7]
Nursing homes (LTCF)	0.0279	0.0126	2.2160	0.0301	+0.04%	[0.00%, 0.08%]	4.0	[0.0, 7.6]
Health workers	0.0260	0.0137	1.8930	0.0626	+0.18%	[-0.01%, 0.37%]	17.3	[-1.0, 35.2]

Notes: 76 observations (neighborhoods). #Elasticity is a form of standardized measure that estimates the relative change in the infection rate as a consequence of a relative change in a determinant. Here, we considered the effect of a 1%

increase in the factor on the infection rate, keeping all other factors the same. The equivalence of cases is based on the number of cases for Barcelona as estimated by the model (9,646) using the mean city-level values of each covariate as shown in Table 1.

4. DISCUSSION. INTERPRETATION AND IMPLICATIONS.

4.1. Discussion

Our results confirm that incidence of COVID-19 is related to several intra-urban socio-demographic factors. In Barcelona, higher rates of infection were found in geographical units that were more densely populated, had more residents aged 70 or over, observed high levels of mobility during lockdown, contained more nursing home facilities, and had the highest levels of people working in health-related occupations. Conversely, neighborhoods with relatively more residents with high levels of education and with an immigration background from high-HDI countries registered fewer COVID-19 infections.

Our results are mostly in line with other indicators of spatial health inequalities for Barcelona which indicate that residents in neighborhoods located in the North of the city—generally lower-income neighborhoods, with lower education, denser areas, and higher immigration from lower-HDI countries (as an indicator of ethnicity)—also have lower life expectancy and suffer more from chronic diseases.[35] The same exposures that put residents at risk of general poor health and comorbidities also have implications for risk of COVID-19 infections.[8 9]

The environmental justice literature further demonstrates several causal pathways which may account for health differences by neighborhood socioeconomic status by showing that, for example, neighborhoods with high percentages of low-income and non-university educated residents historically have more environmental hazards,[36] putting residents at greater exposure to risks leading to greater related health impacts. Because urban social and health injustices already existed in those neighborhoods with higher COVID-19 incidence in Barcelona, including poor housing conditions, and at greater risk for economic disadvantage among others, the current pandemic is likely to reinforce health and social inequalities and urban environmental injustice. People living in these neighborhoods have less of a social safety net during times of both health and socioeconomic stress. They are thus more likely to face an unjust burden in overcoming the pandemic and its economic consequences.

During Spring 2020, the lockdown in Spain limited mobility strictly to those working in essential services, including low-wage jobs that require commuting by public transit to other parts of the city, which predicts higher COVID-19 incidence in geographical units with higher numbers of commuters. In their case, additional health inequalities are likely to manifest because essential workers are often underpaid and under-protected, in positions that require close interactions with the public. Additionally, they may already suffer from underlying health conditions due to their lower socio-economic status, as recent research suggests.[37] As non-essential workers are losing their jobs or facing less pay, these hardships affect lower educated (and logically income) communities more, and jeopardizes their ability to overcome the pandemic in the long-term.[38] In contrast, more privileged residents have greater ability to financially and physically recover. The negative association we found between infection and neighborhoods with high percentages of individuals with post-secondary degree and/or born in high HDI countries can be understood from a dual perspective: first, the presence of this type of residents is closely associated with neighborhoods dominated by middle and upper socioeconomic households, which, in addition, were more likely to work remotely. Second, this group is increasingly formed by young mobile and transient populations,[39] who had the chance to return to their home countries at the initial stage of the pandemic.

Last, results also indicate an expected structural age-related vulnerability, with neighborhoods with a higher percentage of residents over 70 and/or with more nursing homes, predicting higher COVID-19 incidence. Those are thus intersectional social vulnerabilities, particularly important for a context like Spain, which has a high aging population and a high number of residents in nursing homes, many of whom suffer from other comorbid conditions.

4.2. Strengths and limitations

Barcelona is an excellent example to disentangle the spread of the infection within dense and highly mixed use European urban areas. Socioeconomic and urban conditions are significantly different to other urban contexts where most of the research has been conducted. Another strength of our study is that the high number of COVID-19 cases in Barcelona enabled us to test various area-level indicators. In addition, the vast availability of aggregated socio-demographic data at a fine-grained scale allowed us to include many contextual factors that in other studies are often analyzed separately. Nevertheless, using geographically aggregated data also has its limitations, as association found in ecological studies may not necessarily reflect those observed at the individual level. An interesting future line of analysis would be to create buffer zones based on case addresses in order to overcome the limitations of administrative boundaries. Another limitation was that our estimates cover only the municipality of Barcelona and do not include data from the metropolitan area. Last, our measurement of incidence was biased toward more severe patients with COVID-19 as testing procedures were restricted to hospital admissions at this stage of the pandemic. The seroprevalence study conducted between April 27th to May 11th estimated that 7% of the residents in Barcelona's province had developed IgG antibodies against SARS-CoV-2. [40] Assuming this prevalence for the city, the total number of cases that we analyzed represented between 10% and 15% of the people who became infected during our period of study. Therefore, our model is likely to be biased in estimating intra-urban variations of the entire infected population, but not for predicting the most severe cases. Our results may also differ from subsequent waves when massive and rapid COVID-19 testing became available that also detect asymptomatic cases. As the latter is more common among younger people, the predictive value of the Percentage 70+ variable in intra-urban variation of COVID-19 will likely be lower in subsequent waves.

4.3. Final thoughts

Despite initial media and political narratives framing the pandemic as a social equalizer, our analysis shows how vulnerable groups by occupation, age, and ethnicity, who reside in Barcelona neighborhoods with poor pre-existing social and environmental conditions, have statistically higher incidences of COVID-19. With the pandemic, their exposure to overlapping health risks have been compounded by new ones. The COVID-19 pandemic is therefore likely to reinforce existing health and social inequalities, and exacerbate urban environmental injustice in the city. These trends call for public policies and planning interventions to address neighborhood environmental and social factors, strengthen social welfare and health care systems, and improve open green and public spaces to serve as resources and refuges for socially vulnerable groups.

References:

1 Mas-Coma S, Jones MK, Marty AM. COVID-19 and globalization. *One Health* 2020. doi: 10.1016/j.onehlt.2020.100132.

2 Mossong J, Hens N, Jit M, et al. Social contacts and mixing patterns relevant to the spread of infectious diseases. *PLoS Med* 2008;5(3):e74. doi: 10.1371/journal.pmed.0050074

3 Jarvis CI, Van Zandvoort K, Gimma A, et al. Quantifying the impact of physical distance measures on the transmission of COVID-19 in the UK. *BMC medicine* 2020. doi: 10.1186/s12916-020-01597-8

4 Grantz KH, Rane MS, Salje H, et al. Disparities in influenza mortality and transmission related to sociodemographic factors within Chicago in the pandemic of 1918. *Proceedings of the National Academy of Sciences* 2016;113(48):13839-44. doi: 10.1073/pnas.1612838113

5 Quinn SC, Kumar S. Health inequalities and infectious disease epidemics: a challenge for global health security. *Biosecurity and bioterrorism: biodefense strategy, practice, and science* 2014;12(5):263-73. doi: 10.1089/bsp.2014.0032

6 Bambra C, Riordan R, Ford J, et al. The COVID-19 pandemic and health inequalities. *J Epidemiol Community Health* 2020;74:964-968. doi: 10.1136/jech-2020-214401

7 Niedzwiedz CL, O'Donnell CA, Jani BD, et al. Ethnic and socioeconomic differences in SARS-CoV-2 infection: prospective cohort study using UK Biobank. *BMC Med* 2020;18(1):160. doi:10.1186/s12916-020-01640-8

8 Bonow RO, Fonarow GC, O'Gara PT, et al. Association of coronavirus disease 2019 (COVID-19) with myocardial injury and mortality. *JAMA cardiology* 2020. doi: 10.1001/jamacardio.2020.1105

9 Shi Y, Wang Y, Shao C, et al. COVID-19 infection: the perspectives on immune responses. *Cell Death Differ 2020;* 27, 1451–1454. doi: 10.1038/s41418-020-0530-3

10 Ryan NE, El Ayadi AM. A call for a gender-responsive, intersectional approach to address COVID-19. *Global public health* 2020;15(9):1404-12. doi: 10.1080/17441692.2020.1791214

11 Dowd JB, Andriano L, Brazel DM, et al. Demographic science aids in understanding the spread and fatality rates of COVID-19. *Proceedings of the National Academy of Sciences* 2020;117(18):9696-8. doi: 10.1073/pnas.2004911117

12 Ludvigsson JF. Children are unlikely to be the main drivers of the COVID-19 pandemic–a systematic review. *Acta Paediatrica* 2020; 109:1525-1530. doi: 10.1111/apa.15371

13 Park YJ, Choe YJ, Park O, et al. Contact tracing during coronavirus disease outbreak, South Korea, 2020. *Emerging infectious diseases* 2020;26(10):2465-8. doi: 10.3201/eid2610.201315

14 Davies NG, Kucharski AJ, Eggo RM, et al. Effects of non-pharmaceutical interventions on COVID-19 cases, deaths, and demand for hospital services in the UK: a modelling study. *The Lancet Public Health* 2020; 5(7):375-385. doi: 10.1016/S2468-2667(20)30133-X

15 Mogi R, Spijker J. The influence of social and economic ties to the spread of COVID-19 in Europe. *Journal of Population Research* 2021 (online first). doi:10.1007/s12546-021-09257-1.

16 Gómez-Ochoa SA, Franco OH, et al. COVID-19 in health-care workers: a living systematic review and meta-analysis of prevalence, risk factors, clinical characteristics, and outcomes. *American Journal of Epidemiology,* 2020. kwaa191 doi: 10.1093/aje/kwaa191

17 Lancet. The plight of essential workers during the COVID-19 pandemic. *Lancet (London, England)* 2020;395(10237):1587. doi: 10.1016/S0140-6736(20)31200-9

18 Aral SO, Adimora AA, Fenton KA. Understanding and responding to disparities in HIV and other sexually transmitted infections in African Americans. *The Lancet* 2008;372(9635):337-40. doi: 10.1016/S0140-6736(08)61118-6.

19 Bodor JN, Rice JC, Farley TA, et al. The association between obesity and urban food environments. *Journal of Urban Health* 2010;87(5):771-81. doi: 10.1007/s11524-010-9460-6

20 Diez Roux AV, Mair C. Neighborhoods and health. *Ann N Y Acad Sci* 2010;1186:125-145. doi:10.1111/j.1749-6632.2009.05333.x

21 Agyeman J, Schlosberg D, Craven L, et al. Trends and directions in environmental justice: from inequity to everyday life, community, and just sustainabilities. *Annual Review of Environment and Resources* 2016;41:321-340. doi: 10.1146/annurev-environ-110615-090052

22 Tai DB, Shah A, Doubeni CA, et al. The disproportionate impact of COVID-19 on racial and ethnic minorities in the United States. *Clinical Infectious Diseases.* 2020 Jun 20. doi: 10.1093/cid/ciaa815

23 Esteve A, Permanyer I, Boertien D, et al. National age and co-residence patterns shape covid-19 vulnerability. *Proceedings of the National Academy of Sciences* 2020;117(28): 16118-16120. doi: 10.1073/pnas.2008764117

24 Sa F. Socioeconomic Determinants of COVID-19 Infections and Mortality: Evidence from England and Wales. *Institute of Labor Economics (IZA)* 2020; 159. http://ftp.iza.org/pp159.pdf

25 Harris JE. The Subways Seeded the Massive Coronavirus Epidemic in New York City. *National Bureau of Economic Research, Inc* 2020. http://www.nber.org/papers/w27021.pdf

26 Ji Y, Ma Z, Peppelenbosch MP, et al. Potential association between COVID-19 mortality and health-care resource availability. *The Lancet Global Health* 2020;8(4):e480. doi: 10.1016/S2214-109X(20)30068-1

27 Comas-Herrera A, Zalakaín J, Litwin C, et al. Mortality associated with COVID-19 outbreaks in care homes: early international evidence. *International Long-Term Care Policy Network* 2020.

28 Arpino B, Bordone V, Pasqualini M. No clear association emerges between intergenerational relationships and COVID-19 fatality rates from macro-level analyses. *Proceedings of the National Academy of Sciences* 2020;117(32):19116-21. doi: 10.1073/pnas.2008581117

29 Maroko A, Nash D, Pavilonis B. COVID-19 and Inequity: a Comparative Spatial Analysis of New York City and Chicago Hot Spots*. J Urban Health* 2020; 97 461–470. doi: 10.1007/s11524-020-00468-0

30 van Holm EJ, Wyczalkowski CK, Dantzler PA. Neighborhood conditions and the initial outbreak of COVID-19: the case of Louisiana. *Journal of Public Health* 2020. doi: 10.1093/pubmed/fdaa147

31 Brandén M, Aradhya S, Kolk M,et al. Residential context and COVID-19 mortality among adults aged 70 years and older in Stockholm: a population-based, observational study using individual-level data. *The Lancet healthy longevity* 2020;1(2):e80-8. doi: 10.1016/S2666-7568(20)30016-7

32 Bakıcı T, Almirall E, Wareham J. A smart city initiative: the case of Barcelona. *Journal of the knowledge economy* 2013;4(2):135-48. doi: 10.1007/s13132-012-0084-9

33 INE. Análisis de la movilidad de la población durante el estado de alarma por COVID-19 a partir de la posición de los teléfonos móviles. Instituto Nacional de Estadística 2020. https://www.ine.es/covid/exp_movilidad_covid_proyecto.pdf

34 Tibshirani R. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)* 1996;58(1):267-88. doi: 10.1111/j.2517-6161.1996.tb02080.x

35 Agència de Salut Pública de Barcelona. La salut a Barcelona 2018. Barcelona: Agència de Salut Pública de Bracelona 2019. https://www.aspb.cat/wp-content/uploads/2019/12/ASPB_salutbarcelona2018.pdf

36 Evans GW, Kantrowitz E. Socioeconomic status and health: the potential role of environmental risk exposure. Annual review of public health 2002;23(1):303-31. doi: 10.1146/annurev.publhealth.23.112001.112349

37 van Dorn A, Cooney RE, Sabin ML. COVID-19 exacerbating inequalities in the US. *Lancet* 2020;395(10232):1243. doi: 10.1016/S0140-6736(20)30893-X

38 Cole HV, Anguelovski I, Baró F, et al. The COVID-19 pandemic: power and privilege, gentrification, and urban environmental justice in the global north. *Cities & Health* 2020:1-5. doi: 10.1080/23748834.2020.1785176

39 Cocola-Gant A, Lopez-Gay A. Transnational gentrification, tourism and the formation of 'foreign only'enclaves in Barcelona. Urban Studies. 2020:0042098020916111. doi: doi.org/10.1177/0042098020916111

40 Pollán M, Pérez-Gómez B, Pastor-Barriuso R, et al. Prevalence of SARS-CoV-2 in Spain (ENE-COVID): a nationwide, population-based seroepidemiological study. *The Lancet* 2020;396(10250):535-44. doi: 10.1016/S0140-6736(20)31483-5

41 Bischof E, Wolfe J, Klein SL. Clinical trials for COVID-19 should include sex as a variable. *The Journal of Clinical Investigation* 2020;130(7). doi: 10.1172/JCI139306

42 Guan WJ, Ni ZY, Hu Y, et al. Clinical characteristics of coronavirus disease 2019 in China. *New England journal of medicine* 2020;382(18):1708-20. doi: 10.1056/NEJMoa2002032

Summary box:

What is already known on this subject?

- Previous research on virus transmission has shown that individual, household, and neighborhood-level socioeconomic and cultural factors are associated with viral transmission.

- Most of COVID-19 research on spatial variations has been mainly set at the national or sub-national regional level. Because of the internal heterogeneity of these units, it is very difficult to disentangle the different intervening demographic and socioeconomic factors behind risks and exposures to COVID-19.

- The limited research on the COVID-19 pandemic at the neighborhood level (mainly in US and UK) identifies the effect of sociodemographic determinants, like socioeconomic status or ethnicity.

What this study adds?

- We analyze the spread of COVID-19 in Barcelona, a very dense and highly segregated city in Southern Europe, where the first outbreak led to very high levels.

- We test a wide range of sociodemographic and urban characteristics, including mobility during lockdown, 16 variables in total, in order to predict intra-urban variations in COVID-19 infections at the neighborhood level in Barcelona.

- The COVID-19 pandemic is likely to reinforce existing health and social inequalities, and exacerbate urban environmental injustice. These trends call for public policies and planning interventions that must address historic poor neighborhood environmental and social factors, strengthen social welfare systems, and improve open green and public spaces in cities.

Contributors: AGM, ALG, and FLG conceived the study. ALG, FLG, MMD, and JAM collected data, calculated indicators and built the final dataset. AGM, ALG, MMD, JS, and MTM conducted the statistical analyses. IA, CB, HC, AGM, ALG, JS and MTM, wrote the paper. ALG was principal investigator of the study. All authors contributed to the interpretation of data, and read, edited and approved the final manuscript.

Competing interests: None declared.

Data availability: Our data are accessible to researchers upon reasonable request for data sharing to the corresponding author. Our dataset has been built based on publicly available data in the referred repositories.

Ethics approval: No ethical approval was sought for this study as it used aggregated, anonymous, and publicly available data, collected at the neighborhood level.

Funding: This project has been funded by the following programs: H2020 European Research Council (GREEN LULUs SG / GA678034 & HEALIN CoG / GA864616); Ministerio de Ciencia e Innovación (CSO2016-79142-R; GLOBFAM / RTI2018-096730-B-I00I3; 'Maria de Maeztu' Program / CEX2019-000940-M; 'Ramón y Cajal' Program / RYC-2013-14851; 'Juan de la Cierva' Program / FJCI-2017-33842 & IJC2018-035322-I); Agència de Gestió d'Ajuts Universitaris i de Recerca (DEMFAMS / 2017 SGR 1454).; Talent Research Program, Universitat Autònoma de Barcelona.

Patient consent for publication: Not required.