European Commission

JRC TECHNICAL REPORT

Age, gender, and territory of COVID-19 infections and fatalities

Goujon, A

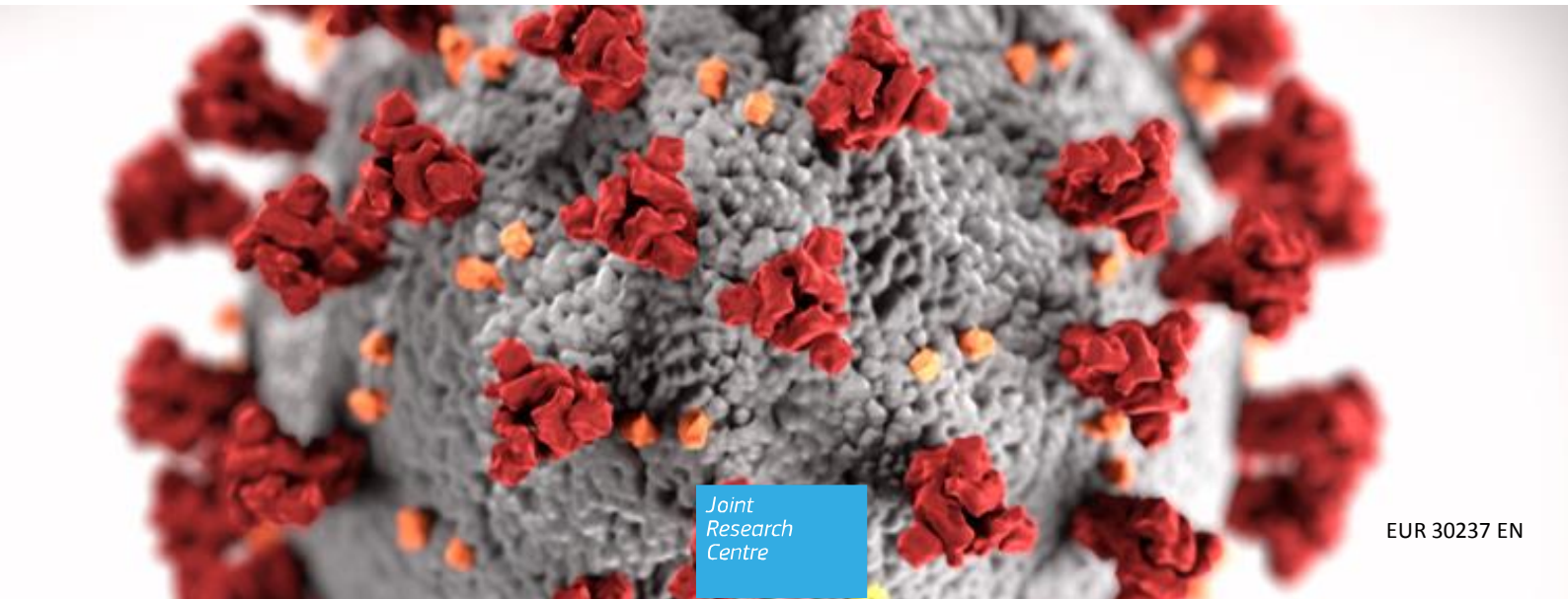
Natale, F

Ghio, D

Conte, A

Dijkstra, L

2020



EUR 30237 EN

Contact information
Name: Anne Goujon
Address: Via E. Fermi 2749, Ispra (VA), Italy
Email: anne.goujon@ec.europa.eu
Tel.: +390332785271

EU Science Hub
https://ec.europa.eu/jrc

JRC120680

EUR 30237 EN

Print	ISBN 978-92-76-19181-0	ISSN 1018-5593	doi:10.2760/642022
PDF	ISBN 978-92-76-19180-3	ISSN 1831-9424	doi:10.2760/838390

Luxembourg: Publications Office of the European Union, 2020

How to cite this report: Goujon, A., Natale, F., Ghio, D., Conte, A., Dijkstra, L., *Age, gender, and territory of COVID-19 infections and fatalities*, EUR 30237 EN, Publications Office of the European Union, Luxembourg, 2020, ISBN 978-92-76-19180-3, doi:10.2760/838390, JRC120680.

Contents

Acknowledgements

Data from The European Surveillance System – TESSy – are provided for the following countries: AT, BE, BG, CY, CZ, DE, DK, EE, EL, FI, FR, HR, HU, IE, IS, IT, LT, LU, LV, MT, NL, NO, PL, PT, SE, SK, UK and released by ECDC.

We thank reviewers at ECDC for providing useful comments on the draft of the paper. However, the views and opinions of the authors expressed herein do not necessarily state or reflect those of ECDC. The accuracy of the authors' statistical analysis and the findings they report are not the responsibility of ECDC. ECDC is not responsible for conclusions or opinions drawn from the data provided. ECDC is not responsible for the correctness of the data and for data management, data merging and data collation after provision of the data. ECDC shall not be held liable for improper or incorrect use of the data.

Authors

Goujon, Anne	European Commission, Joint Research Centre, Unit E6
Natale, Fabrizio	European Commission, Joint Research Centre, Unit E6
Ghio, Daniela	European Commission, Joint Research Centre, Unit E6
Conte, Alexandra	European Commission, Joint Research Centre, Unit E6
Dijkstra, Lewis	European Commission, DG for Regional Policy, Unit B1

Abstract

The COVID-19 virus has spread within the last weeks to a large number of countries worldwide. Increasing evidence suggests that the epidemics does not impact everyone the same. In this note we explore the main demographic differentials in the spread and impact of COVID-19 paying special attention to the combined effect of age and gender, and to the differences at territorial level where population density and size play a main role in the diffusion and effects of the disease in terms of morbidity and mortality. This knowledge complements and refines the epidemiological information about the spread of the virus.

For this analysis, we rely mostly on descriptive exploration of data from The European Surveillance System (TESSy) database developed at the European Centre for Disease Prevention and Control (ECDC) on the number of cases and fatality rates. The analysis at territorial level, which we complement with data from USAfacts about counties in the United States, studies the changes in R0 – the infection reproduction number – across territories with different levels of urbanization.

The findings confirm and define the differential impacts in terms of infections and fatalities. The information is important for stakeholders at EU, national and sub-national levels in charge of designing an exit strategy from COVID-19 and anticipating the rebound for certain segments of the population with differential medical needs, particularly those living in high-density locations.

1 Introduction

Beyond aggregate measures on COVID-19 cases and deaths, several statistics have emerged showing differences that are noticeable in the spread of the COVID-19, between different age groups and sex, and also countries, place of residence and regions.

In this article, we examine some of the main differentials emerging from The European Surveillance System (TESSy) database developed at the European Centre for Disease Prevention and Control (ECDC) and which collects around 30 variables for each COVID-19 case reported by the Member States with full epidemiological detail and the main demographic characteristics of individuals (see Annex 1 for sample description).

This analysis, disaggregated by demographic characteristics of COVID-19 cases and place of residence, provides some evidence to develop an exit strategy and surveillance system to control and monitor the spread of COVID-19 that takes into account the specific characteristics of the virus spread on the entire European territory. For the interpretation of the results, however, it is important to consider the major differences across countries in case and death counting procedures, in the number and type of tests and laboratory capacity, as well as the different phases of the epidemic that different territories go through and that strongly influence the results of this analysis.

This report extends the analysis implemented by the demographic team in JRC-KCMD. The first report compared cases and fatalities of COVID-19 by age and by provinces and regions in some of the most affected countries, with the main objective to evaluate possible effects of the demographic characteristics of population on the epidemic outcomes (Natale et al. 2020a). When considering fatalities in relative terms across age groups, we do not find significant differences between countries, since fatalities are uniformly concentrated in the population over 60 years of age. Also while there are clear geographical patterns in the diffusion of the disease, these patterns are not related to differences in the share of elderly across provinces and regions. The second report evaluated the potential socio-economic vulnerability and epidemic risk by age, particularly focusing on the potential role of social interactions for different groups in various settings in the exit strategy from COVID-19 (Natale et al. 2020b).

Old age has been identified as a risk factor to be affected compared to young and middle-age which has mostly to do with the lower immunity due to frailty and higher prevalence of chronic illnesses in the elderly population (Verity et al. 2020). A further factor of heterogeneity in mortality due to COVID-19 is gender where large sex differences exist with men having higher risk compared to women (Dowd et al. 2020, Caramelo 2020), which might have to do with the sex hormones like testosterone and oestrogen that seems to be key in adapting the body's immune response and the presence of other risk factors, e.g. diabetes, hypertension and cardio-vascular diseases, that are affecting more men than women. In this article, we analyse how the gender differential in the number of cases and fatalities varies by age, showing that for all countries in our sample the difference peaks for the persons aged 60-69 years.

While density dependence – when the density of a population regulates its growth rates - has been shown to have a greater influence on mortality for nonhuman hosts such as animals and insects (Greenhalgh 1992), the fact that the spread of COVID-19 (like that of Middle East Respiratory Syndrome (MERS) and Severe Acute Respiratory Syndrome (SARS) occurs between close contacts via aerosols and droplets (Peeri et al. 2020) renders the study of place of residence (urban/rural) and population density highly relevant (see also Rocklöv and Sjödin 2020). Furthermore, population density together with other complex ecological factors play a large role in influencing the chance of emergence of zoonotic infections. Using the data from TESSy for several European Member States and data for the US, we show as expected that the degree of urbanization at regional/provincial/county levels has a substantial impact on the reproduction number R0. This finding has implications for the containment of the epidemic which is more challenging in urban regions and for the exit strategy which could be differentiated between distinctive population settlements as a necessary requirement to avoid a

rebound of the disease in high-density localities that have a high incidence of infection. This corroborates analysis that were done in the framework of other SARS epidemics (Fang et al. 2009).

2 Gender differentials by age

2.1 Overview

Figures 1.A and 2.A display the age structure of the COVID-19 epidemic by gender for the whole sample. Overall, the cumulative number of diagnosed positive cases is particularly large among the population between 35 and 65 years of age. This group which consists mostly of a population of working-age represents 50 percent of confirmed positive cases, indicating that the infection does not only affect the elderly (75+ age-group is 20 percent of all cases).

The figures also capture gender differences in positive cases : more cases are notified among men aged from 55 to 80 years old compared to women, while higher numbers of positive cases are reported among women aged from 15 to 55 years and above 80. It shows that the male disadvantage is particularly pronounced for the age group 65-70.

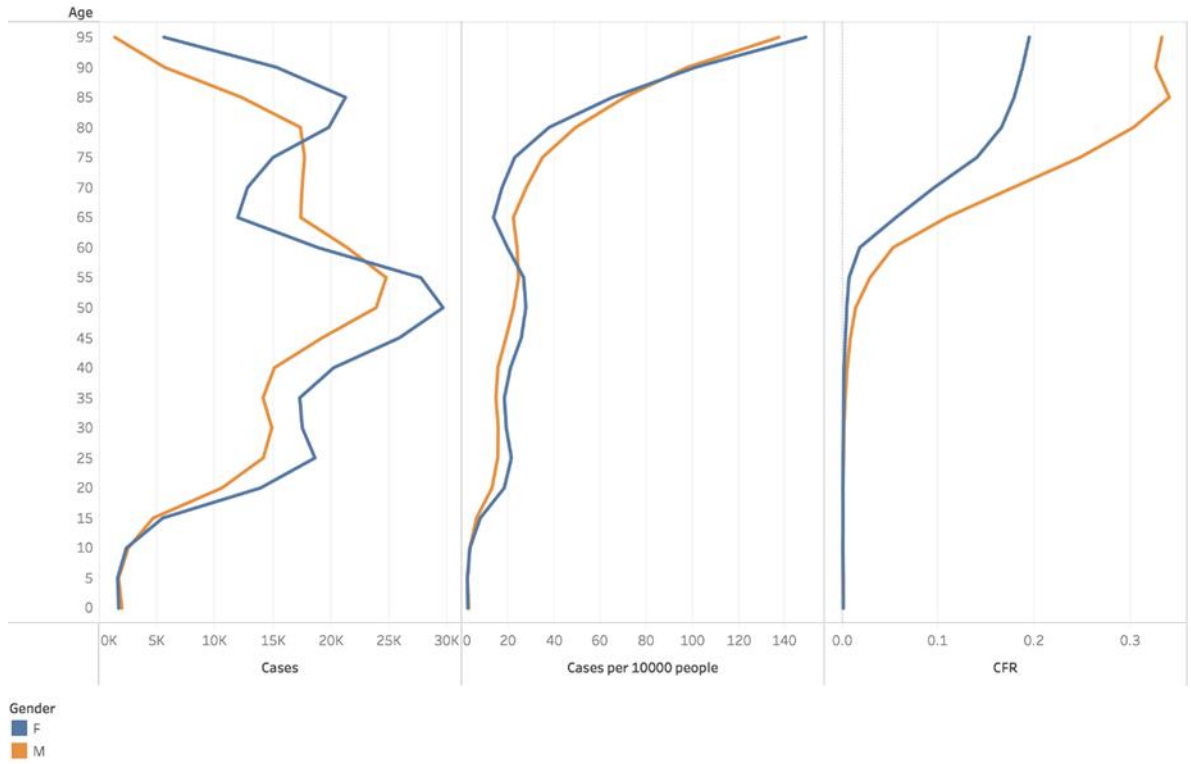
The fact that the diagnosed cases are higher among women below 55 years and among men in the age-group 55-80 years is visible in Figure 3 for a selected number of Member States. However, some national specificities can be observed. In Italy and Belgium, men aged between 60 and 80 are respectively 1.5 and 1.4 more likely to be reported as positive as women (Figure 3). The higher number of cases for men could be linked to the fact that testing, especially in the early phases, of the pandemic was primarily performed on critical cases. Since elderly men face more serious consequences than women, they could, as a result, be more likely to be tested. For some other countries like Germany and Portugal, the male to female cases ratio is close to 1 around the 60-75 age groups.

When related to the total population by age groups, the proportion of diagnosed with COVID-19 patients is higher among women under the age of 50; in the sample, there are indeed 10 infected women against 8 infected men for every 10,000 persons living in the countries below the age of 20 years. The proportion is about the same for men and women around the age of 50-55 years; yet, from the age of 58 onward, the male prevalence among notified positive cases becomes evident (Figure 1.B and 2.B). The comparison across countries reported in our sample (Figure 4) confirms that positive cases among the male population above the age of 60 years are overrepresented in relative terms compared to female ones in many countries. The diagonal of Figure 4 indicates the equal distribution between gender; values below the diagonal indicate the prevalence of cases among men. At a higher number of cases per 10,000 population, the male disadvantage seems also larger compared to countries with lower relative numbers, except for Ireland.

Since the number of positive cases in the dataset corresponds largely to the number of symptomatic individuals who were present at health facilities, this gives a broad indication[1] that COVID-19 related morbidity and mortality risks are of particular concern for men. As Figure 1.C shows, the case fatality rate (CFR -- the proportion of deaths from COVID-19 compared to the total number of people diagnosed with the disease) confirms the male disadvantage: the CFR is higher among men for all age groups. Overall, the male mortality disadvantage peaks at the age of 85 when the CFR for men is 9.7% higher than that for women (24.1% and 14.4% respectively). Figure 5 shows that COVID-19 related fatalities are higher among men aged 60+ in all countries, with the exception of Cyprus and Slovakia. Gender differences are particularly substantial in the Netherlands, where the CFR for men aged 60+ is 8 percentage points higher than for women (25% and 17% respectively), similar to Italy, where the male CFR is 22 percent against 15% for women.

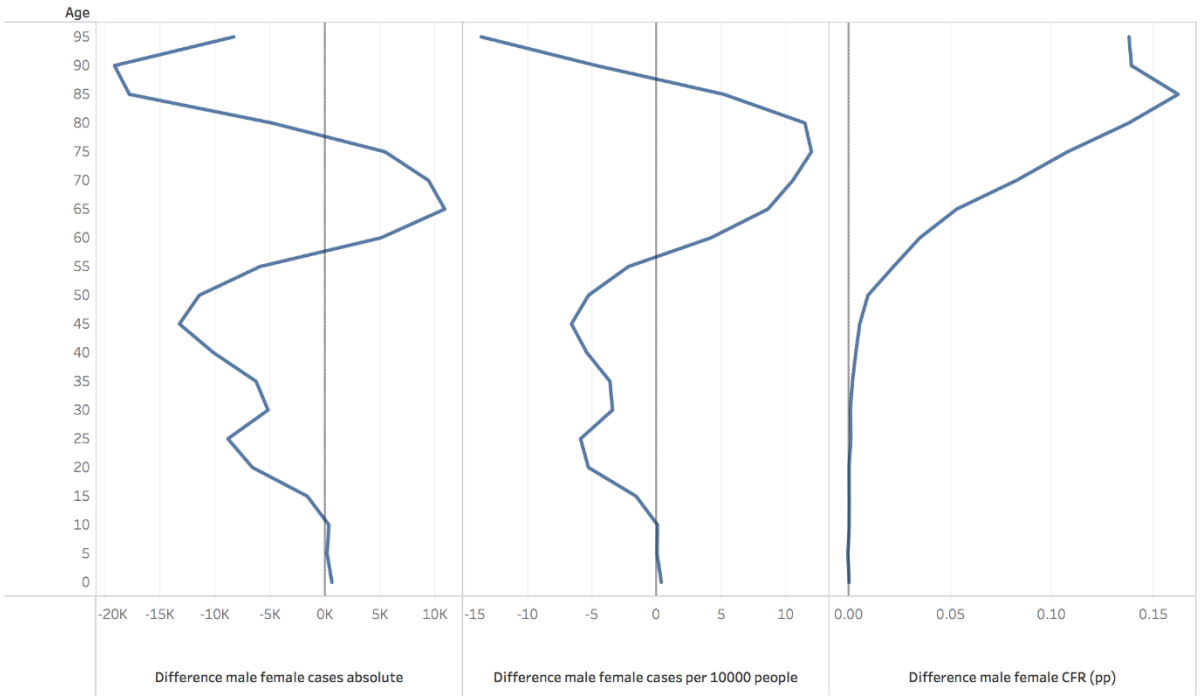
[1] Accurate analyses of COVID-19 effects should be conducted at the end of the pandemic, when the number of infected persons (symptomatic and asymptomatic individuals) and deaths will be counted. As well, many countries have been testing some groups more widely irrespectively of symptoms, e.g. health care workers, patients in long term care facilities, etc.

Figure 1 (A,B,C). Distribution of cases in absolute numbers (A), cases per 10,000 people (B) and case fatality rates (C) by age and gender



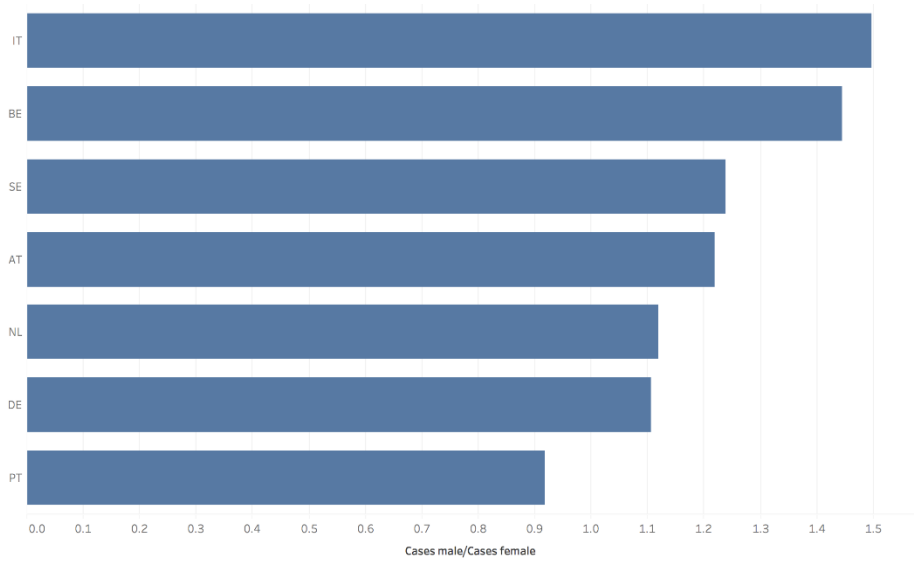
Source: ECDC TESSy (May 11, 2020).

Figure 2 (A,B,C). Distribution of the difference between male and female cases in absolute numbers (A), cases per 10000 peopl e (B) and case fatality rates (C) by age



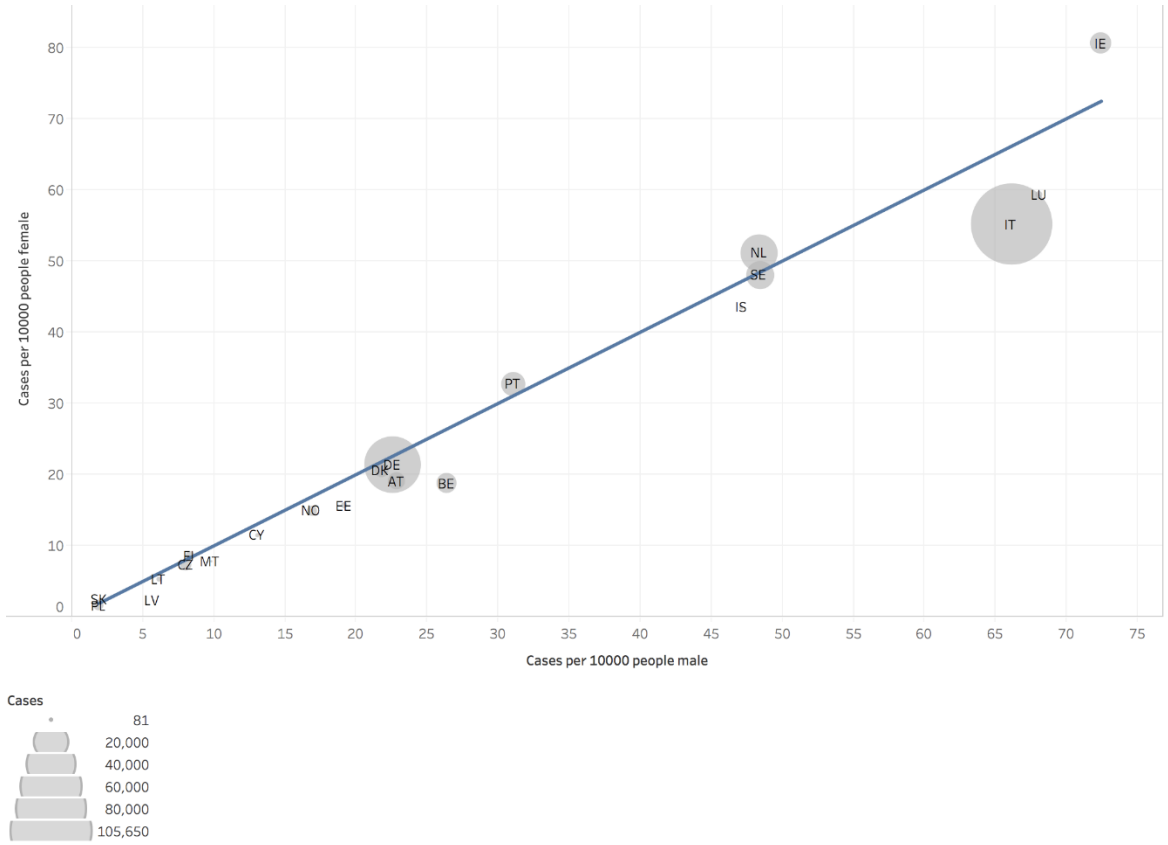
Source: ECDC TESSy (May 11, 2020).

Figure 3. Distribution of the male to female ratio of cases by age and country for ages between 60 and 80



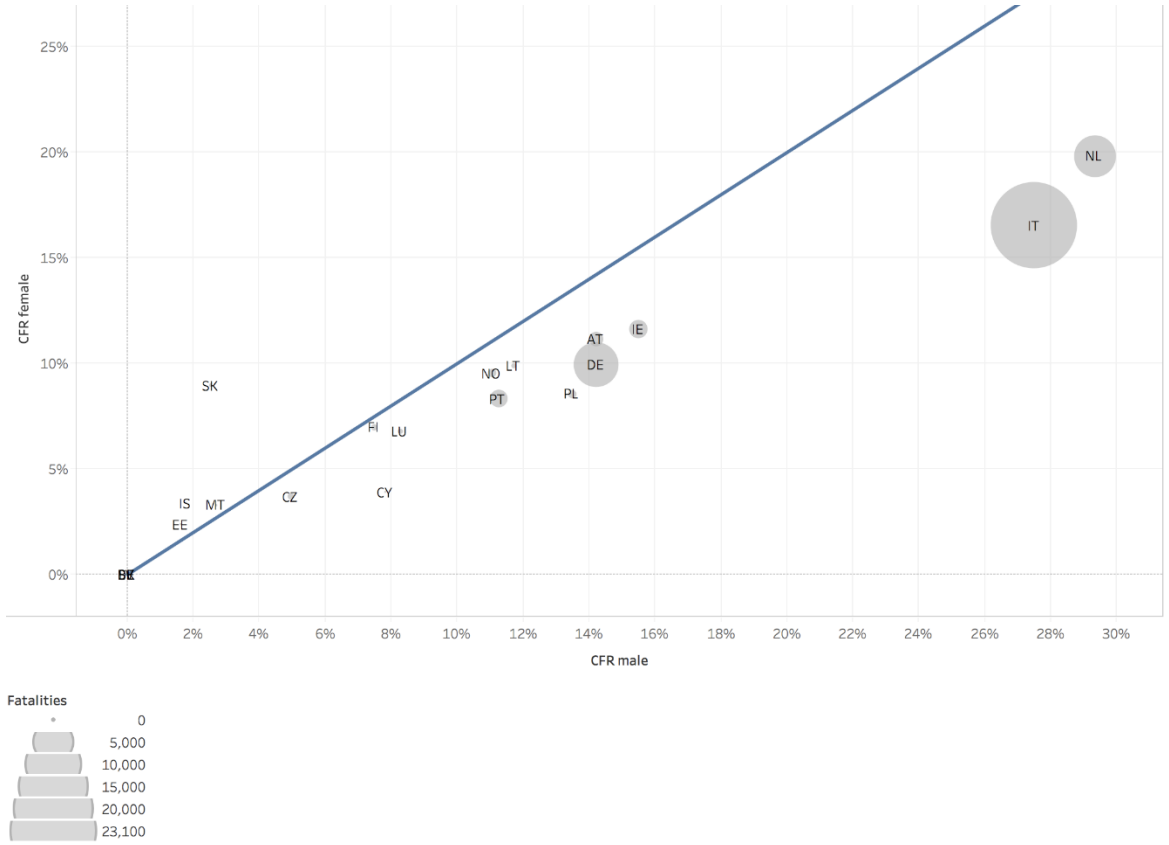
Source: ECDC TESSy (May 11, 2020).

Figure 4. Comparison between cases per 10000 people above age 60 for men and women by country. Values below the diagonal indicate that the number of cases is higher for men than for women in relative terms



Source: ECDC TESSy (May 11, 2020).

Figure 5. Comparison between the case fatality rates for men and women above age 60 by country. Values below the diagonal indicate that the CFR for men is higher than for women.



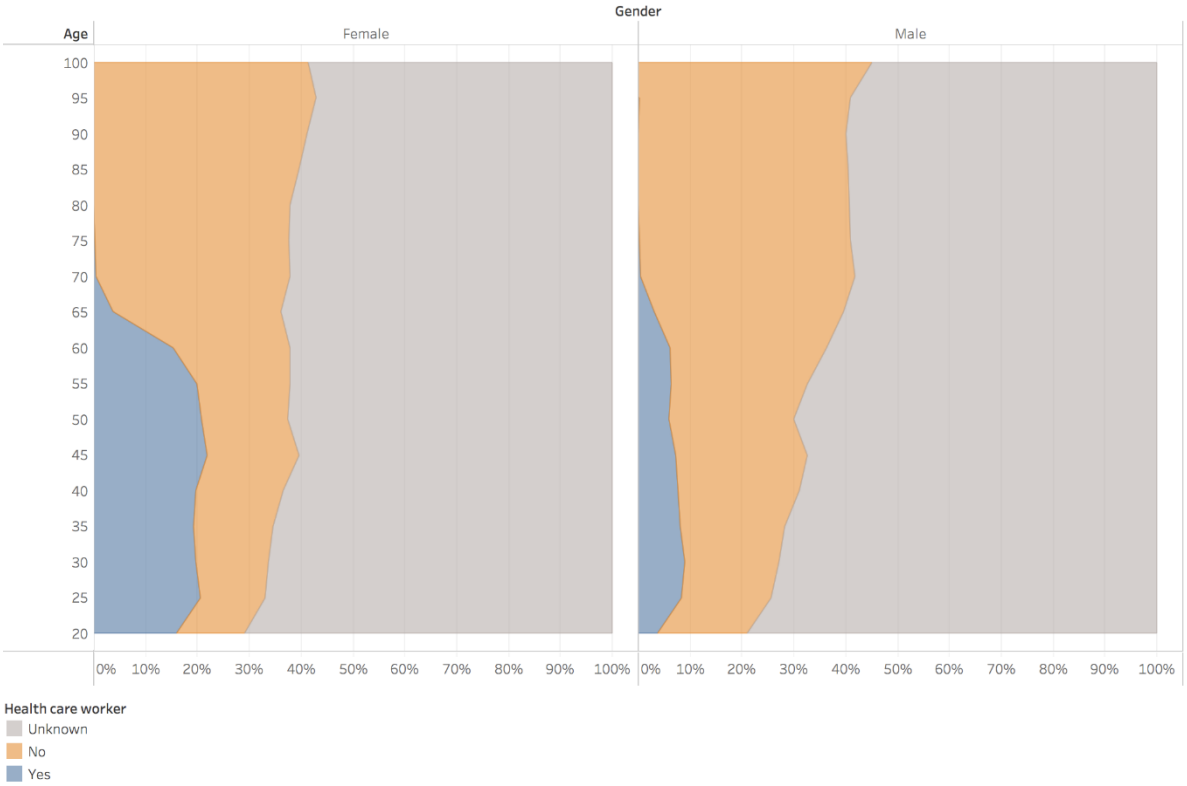
Source: ECDC TESSy (May 11, 2020).

2.2 How to explain gender differentials

Our results suggest that there is an age and gender divide in COVID-19 cases and fatality rates which does not necessarily reflect the population structure. This is not unprecedented as seen from other epidemics. During the Ebola epidemic, women were more likely to be infected by the virus, given their role as caregivers within families and their overrepresentation as health-care workers (Davies and Bennett, 2016). Chinese official sources reported the higher risk that female health workers incurred in Hubei province, where more than 90% of health-care workers were women (Boniol et al., 2019). In some European member states, there has been increasing evidence that outbreaks in elderly homes have spilled over via the health care personnel, which is a sector that commonly reports a majority of female employees[2]. It is clear that such spill over could occur outside elderly homes, exposing patients assisted at home to become additional clusters of infections. The higher number of cases among women below the age of 55 in European Member States could also stem in part from their larger participation in the health sector as well as from their higher care-giving responsibilities with the family compared to men. Figure 6 shows that the proportion of health care professionals among diagnosed cases is substantially larger for women than for men. However, the data have to be interpreted carefully since healthcare workers could be over-represented as they are much more likely to be tested.

[2] See for instance: https://www.ecdc.europa.eu/sites/default/files/documents/covid-19-rapid-risk-assessment-coronavirus-disease-2019-eighth-update-8-april-2020.pdf (May 24, 2020)

Figure 6. Percentage of healthcare workers among cases by age and gender. Overall healthcare workers represent 6.6% (25,963) of the total of cases



Source: ECDC TESSy (May 11, 2020).

The male disadvantage is evident in COVID-19 fatality. Although current disaggregated datasets are incomplete and should be cautiously interpreted, the higher fatality rate for men may derive from gender-based immunological differences [Chen et al., 2020], or be associated with comorbidities, including hypertension, cardiovascular diseases and drinking alcohol more commonly observed among men [Editorial The Lancet, 2020].

Previous outbreaks have shown the relevance of analysis of the role of women in society into the response efforts and definition of policy interventions. These results also confirm that the current pandemic is linked to the specific demographic characteristics of the population resulting in outcomes that affect the population in an uneven way. This evidence therefore has significant implications for the definition of an exit strategy. The effectiveness of measures to gradually lift the lockdown and reopen the economy will also depend on the integration of the demographic profile into its strategy.

3 Territorial differences

With the rapid spread of the COVID-19 pandemic, attention has focused primarily on the hardest-hit urban regions in terms of number of cases and deaths, including Wuhan, Milan, Madrid, London and New York City. Although the incidence of the virus is highest globally in the dense and populated urban centres, the epidemic has also spread outside of big cities, with possible important implications for the evolution of the reproduction number and mortality between urban and rural regions.

While it is commonly assumed that high population density increases the vulnerability of some regions to epidemics due to the high frequency of social and economic interactions, this is not clearly supported by the health geographical evidence on the evolution of earlier pandemics (Garrett, 2007; Chowell et al., 2008; Li et al., 2018). According to Chandra et al. (2013) on the influenza pandemic of 1918–19 in India, there is also a threshold level of population density of 175 people per square mile above which social distancing and quarantine policies are more effective.

However, population density is certainly an essential element in understanding virus transmission, even if it is one among many of the key factors determining vulnerability of some regions to the virus. From small villages in the Lombardy and Veneto regions in Northern Italy to large global metropolises, COVID-19 has hit hard in very different places, showing that the link between population density and the virus transmission is not easy (see Figures A2.1 and A2.2 in the Annex). Moreover, urban and rural regions differ in many dimensions, including population volume, age, education, wealth, social capital and type of work. All these factors can jointly influence the vulnerability and resilience of different places to the virus.

Figure 7 provides a descriptive overview of the duration of the epidemic in each province (days between first and last reported case) grouped three typologies of degree of urbanisation (See Appendix 1 for details on the sample composition and the time span for each country[3]). For this analysis, we follow the Eurostat classification of urban-rural regional typology[4] to indicate the category of a region - urban, intermediate and rural - based on the share of local population living there. The sample consists of 187 urban regions, 374 intermediate regions and 286 rural regions at different stages of the epidemic (see Annex 2). Since the first COVID-19 case was detected, the pandemic has evolved more rapidly and with a higher incidence in most urban regions, where a large share of the European population lives[5]. The spread of the virus decreases along with the degree of population density, from intermediate to rural regions. Overall, the lower population density outside urban regions seems to have potentially contributed to a lower incidence of the virus[6]. Whereas intermediate and rural regions in Europe are not immune to the risks associated with the spread of the virus, they might be particularly vulnerable to the virus due to the population demographics (higher share of elderly population and of chronic diseases than in urban regions), income level and limited access to medical resources due to a shortage of health workers, hospitals and intensive care beds.

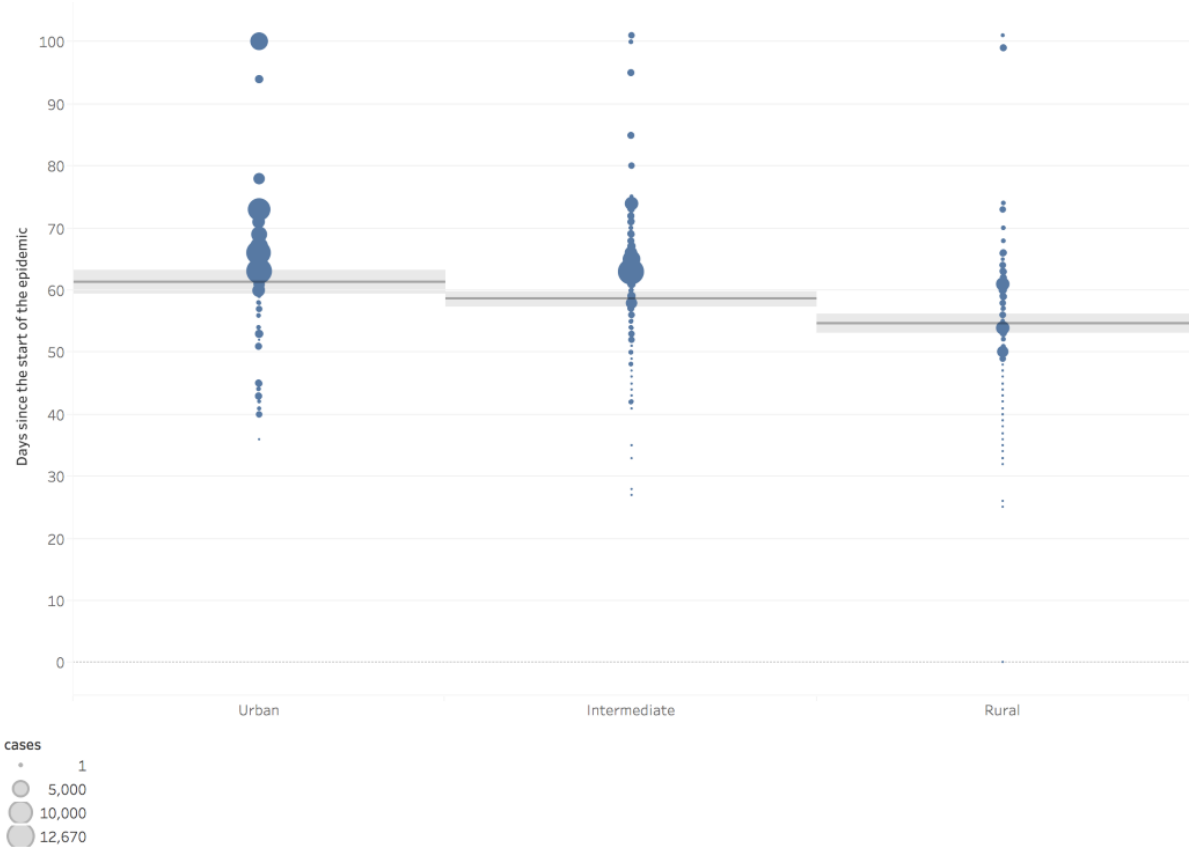
[3] It is important to note that the availability of data on the virus spread is not homogeneous in all regions analysed and that these variations may depend on the number of tests carried out in each place.

[4] Eurostat's definition for urban-intermediate-rural regions are applied to NUTS level 3 regions. The urban-rural typology (https://ec.europa.eu/eurostat/web/rural-development/methodology) is a classification based on the following three categories: a) predominantly urban regions, NUTS level 3 regions where more than 80% of the population lives in urban agglomerations; b) intermediate regions, NUTS level 3 regions where more than 50% and up to 80% of the population live in urban agglomerations; c) predominantly rural regions, NUTS level 3 regions where at least 50% of the population live in rural grid cells.

[5] According to Eurostat, in 2019, 40% of the European population lived in urban regions, 39% lived in intermediate regions, and 21% lived in rural regions.

[6] Since the beginning of the epidemic, outbreaks and positive cases have been more limited in rural and intermediate regions.

Figure 7. Days between the first and the most recent reported case by urban-rural regional typology of EU provinces. Each circle is representing a province, the size is proportional to the number of cases and the horizontal line is the average with 95% confidence by the urban-rural regional typology



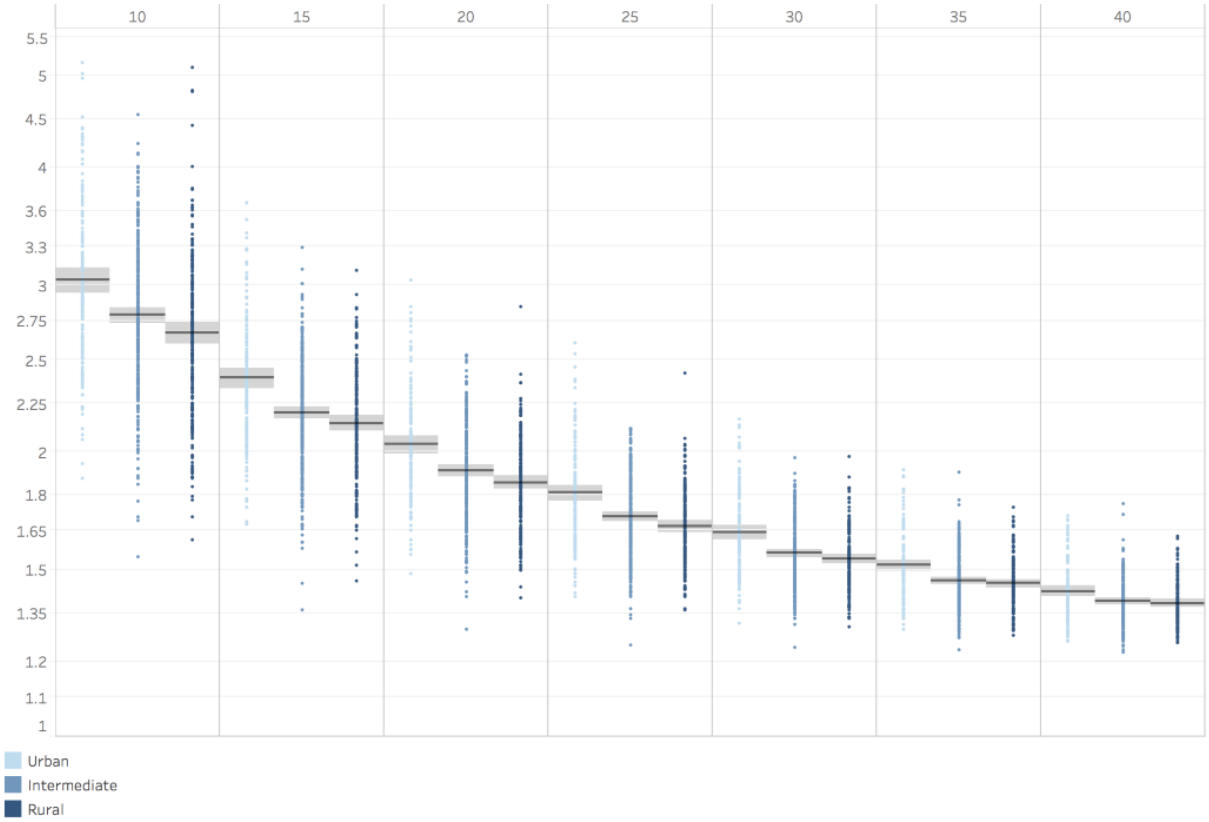
Source: ECDC TESSy (May 11, 2020).

Figure 8 shows the territorial disparity in the contagiousness of the virus, as measured by the Basic Reproduction Number R0[7], corresponding to the average number of secondary infections from one case at the beginning of the epidemic. We track the number of cases through the daily estimated R0 to determine the different levels of infection by regions (the calculation method is explained in Annex 3).

The differences in contagiousness based on reported cases follow the territorial demarcation lines. The virus mainly affects the population where they are the more numerous, i.e. in urban regions. For those who do not live in urban regions, the impact is smaller. In the first 10 days of the epidemic, the estimated R0 (with 95% CI) was equal to 3 in urban regions, i.e. each new positive case of COVID-19 has produced three new cases in urban regions. The average for intermediate regions at the beginning of the epidemic was about 2.8 new cases per infected person, followed by rural regions with an average of 2.6. During the course of the infection, R0 values have decreased in the three different geographical units; on the 40th day since the outbreak began, in the three regions R0 values continue to be well above 1, which is the level above which the infection will continue to spread[8].

[7] We consider R0 rather than R since the analysis is carried out at the early stages of the epidemic.
[8] According to the WHO, the pandemic influenza virus (H1N1) 2009, or swine flu, had a R0 between 1.2 and 1.6, which made it easier to control its spread than viruses with greater transmissibility.

Figure 8. R0 by the urban-rural regional typology of EU provinces[9]. The calculation is repeated considering increasing temporal windows (10-40 days) since the start of the epidemic. Each circle is representing the calculation of R0 for a province and the horizontal line is the average of R0 with 95% confidence intervals by the urban-rural regional typology



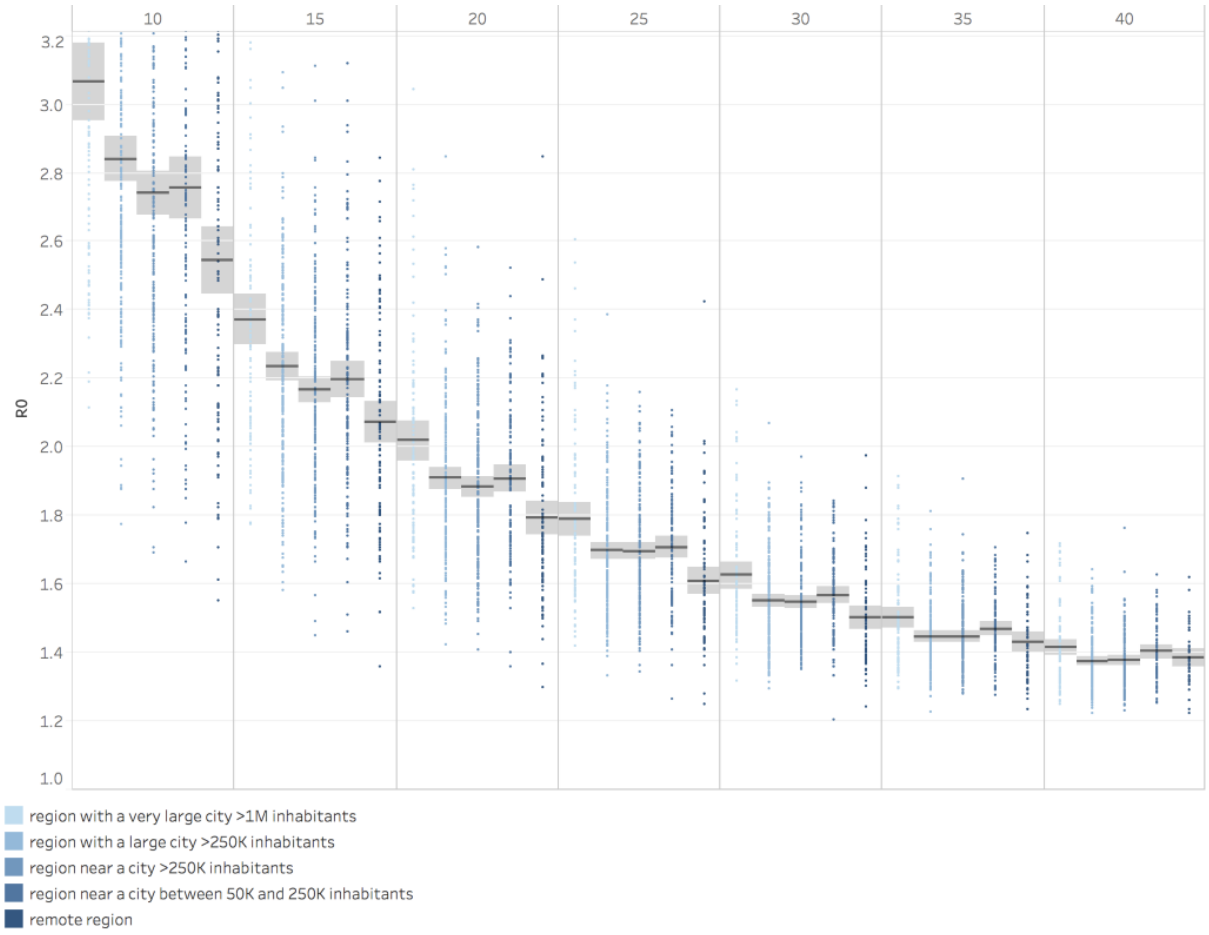
Source: ECDC TESSy (May 11, 2020).

The following series of graphs illustrates the results of the several sensitivity analyses that have been conducted using different territorial classifications, based on absolute measures of settlement population size. In particular, Figure 8 displays the results according to 5 different territorial classes for Europe, while Figure 9 shows 9 classes for the counties in the United States based on data from USAfacts.org (see Annex 1), which allows us to see more clearly the course of the infection. In all cases, virus transmission depends largely on the population size. These additional territorial classifications confirm the results of an epidemic infection that - at each stage - mainly affects the largest and most densely populated regions and a progressive reduction in the number of infections as population size decreases. Moreover, with time, the regional disparities become smaller. Overall, this analysis confirms a clear relationship between settlement population size and viral infections across populations, although a limitation of this analysis could be that testing intensity and access to testing would be higher in more densely populated areas, but is unlikely to overturn the gradient that we observed. Population size is a factor determining the level of infection at each stage of the epidemic, and regions/counties with large cities with more than 1 million inhabitants have the highest R0, while less populated regions have the lowest values. In the case of the United States, the gradient of R0 for detailed population size classes is robust for all time periods, with decreasing R0 values during the course of the epidemic. There are few exceptions in Europe to the general trend observed with some

[9] Traditionally the OECD has classified TL3 regions as predominantly urban, intermediate, or predominantly rural regions (respectively 3_PU, 3_IN,3_PR). This typology is mainly based on population density in each local unit, combined with the existence of urban centres where at least one-quarter of the regional population reside.

remote regions having higher R0 than more populated regions which would require further investigation.

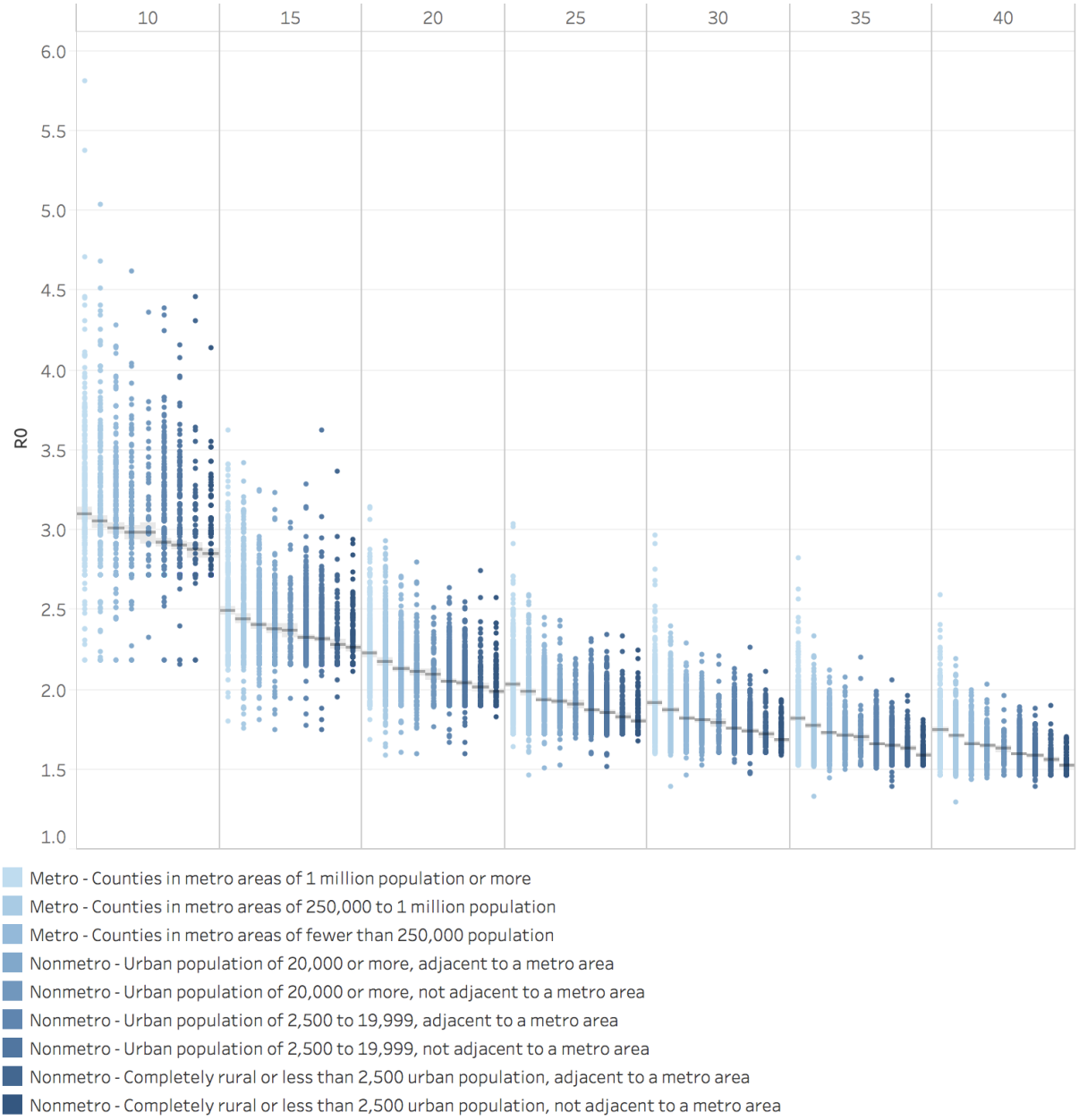
Figure 9. R0 by metro region by size and non-metro region by distance to city by size for EU provinces[10]. The calculation is repeated considering increasing temporal windows (10-40 days) since the start of the epidemic. Each circle is representing the calculation of R0 for a province and the horizontal line is the average of R0 with 95% confidence intervals



Source: ECDC TESSy (May 11, 2020).

[10] This method classifies metro and non-metro regions. Metro regions are classified according to the population size of the functional urban area (FUA) into metro region with a FUA of at least 1 million inhabitants and a metro region with a FUA between 250,000 and 1,000,000 inhabitants. Non-metro regions are classified according to the distance to a FUA and the population size of that FUA into: region near a FUA with more than 250,000 inhabitants, region near a FUA with between 50,000 and 250,000 inhabitants and a remote region.
https://www.oecd-ilibrary.org/urban-rural-and-regional-development/classifying-small-tl3-regions-based-on-metropolitan-population-low-density-and-remoteness_b902cc00-en

Figure 10. R0 by the degree of urbanisation of US counties[11]. The calculation is repeated considering increasing temporal windows (10-40 days) since the start of the epidemic. Each circle is representing the calculation of R0 for a county and the horizontal line is the average of R0 with 95% confidence intervals by the degree of urbanisation



Source: USAFacts https://usafacts.org/visualizations/coronavirus-covid-19-spread-map/ (April 27, 2020).

[11] The 2013 Rural-Urban Continuum Codes form a classification scheme that distinguishes metropolitan counties by the population size of their metro area, and nonmetropolitan counties by degree of urbanization and adjacency to a metro area. https://www.ers.usda.gov/data-products/rural-urban-continuum-codes.aspx

4 Conclusion

Results from the analysis suggest that there is an age and gender divide in COVID-19 infection and fatality rates which does not necessarily reflect the population structure. There are more reported cases of diagnosed women by COVID-19 compared to men in the population below the age of 60 years. One possible reason is that women represent a high percentage of workers in the health care sector and care facilities and are therefore at the forefront of the fight against COVID-19. Moreover, by working in health and care facilities they have a potentially higher chance of being tested.

Between 60 and 80 years old, more men were diagnosed positively compared to women. The male disadvantage is particularly pronounced for the age group of the 65-70 years old, especially in Italy. A tentative explanation for the higher number of cases for men in Italy compared to other countries is that only hospitalized patients were tested (especially at the beginning of the epidemic) and elderly men who are facing more serious consequences are more likely to be hospitalized. Although this is also the case for other countries (like France or the Netherlands), this pattern is more visible in Italy.

While in absolute terms there are overall more women than men diagnosed with COVID-19 above 80 years of age, in relative terms - to the size of the total population - we observe that the male population above the age of 60 years is overrepresented compared to the female ones in the large majority of European countries. The COVID-19 related mortality risks are of particular concern for men as the case fatality rate is higher among men for all age groups.

The analysis of differences in the infection reproduction number along the continuum of provinces/regions/counties in Europe and in the United States shows a clear disadvantage for urban regions, especially at the start of the epidemic. The difference between regions of residence becomes smaller as the epidemic develops and the reproduction number decreases.

Our analysis faces some limitations because we consider only diagnosed cases and fatalities as reported by TESSY data, at the national and sub-national levels There are also significant differences in the number of tests and capacity of testing between countries and regions and over time. As a result, the completeness of case-based data varies by country (compared to the overall number of reported cases)(see Annex Figure A1.1).

The analysis shows the importance of accurate and timely collection of data that differentiate and categorize COVID-19 affected populations beyond the reporting of the sheer total number of cases and fatalities.

Overall, this paper shows the relevance of a demographic differentiated analysis of the COVID-19 infection in the European member states population. The observed significant heterogeneity by age, sex and place of residence of the population will be a key information to guide the preparedness and response In healthcare facilities, the definition of policy interventions, particularly with the design of the exit strategy which should integrate demographic and geographic profiling. This could possibly lead to gradual virus deconfinement plans depending on the structure (e.g. prolonged confinement for residents and personnel in retirement houses), or on the place of residence (e.g. special plans for large cities). The observed gender differences point at the need to pay special attention to health care workers – regardless whether they are male or female.

References

Boniol M, et al., 2019. Gender equity in the health workforce: analysis of 104 countries: Working Paper 1. Geneva: World Health Organization.

Caramelo F, Ferreira N, Oliveiros B. 2020. Estimation of risk factors for COVID-19 mortality - preliminary results. MedRxiv preprint. doi: https://doi.org/10.1101/2020.02.24.20027268

Chandra S, Kassens-Noor E, Kuljanin G and Vertalka J. 2013. A geographic analysis of population density thresholds in the influenza pandemic of 1918–19. International journal of health geographics, 12(9).

Chen N et al. 2020. Epidemiological and clinical characteristics of 99 cases of 2019 novel coronavirus pneumonia in Wuhan, China: a descriptive study. Lancet, 395 (10223):507-513.

Chowell G, Luís MAB, Niall J, Wladimir JA, Cécile V. 2008. The 1918–1919 influenza pandemic in England and Wales: spatial patterns in transmissibility and mortality impact. Proceedings of the Royal Society B, 275(1634):501–509.

Davies SE, Bennett B. 2016. A gendered human rights analysis of Ebola and Zika: locating gender in global health emergencies. International Affairs, 92(5): 1041–1060.

Dowd JB, Andriano L, Brazel DM, Rotondi V, Block P, Ding X, Liu Y, Mills MC. 2020. Demographic science aids in understanding the spread and fatality rates of COVID-19. Proceedings of the National Academy of Sciences. https://doi.org/10.1073/pnas.2004911117

Editorial The Lancet, 2020, The gendered dimensions of COVID-19, April 11, 2020, doi: https://doi.org/10.1016/S0140-6736(20)30823-0

Fang LQ, De Vlas SJ, Feng D, Liang S, Xu Y-F, Zhou J-P, Richardus JH, Cao W-C. 2009. Tropical Medicine and International Health, 14 suppl. 1: 14-20. https://doi.org/10.1111/j.1365-3156.2008.02189.x

Garrett TA. 2007. Economic Effects of the 1918 Influenza Pandemic: Implications for a Modern-day Pandemic. In Working paper CA0721. 2010th edition.: Federal Reserve Bank of St. Louis Web site.

Greenhalgh D. 1992. Some results for an SEIR epidemic model with density dependence in the death rate. IMA Journal of Mathematics Applied in Medicine & Biology (1992) 9, 67-106.

Li R, Richmond P, and Roehner BM. 2018. Effect of population density on epidemics. Physica A: Statistical Mechanics and its Applications, 510: 713-724.

Natale F, Ghio D, Tarchi D, Goujon A, Conte A. 2020a. COVID-19 cases and case fatality rate by age. Ispra: Joint Research Centre.

Natale F, Ghio D, Goujon A, Conte A. 2020b. Socio-economic vulnerability and epidemic risk by age. Implications for the exit strategy from COVID-19. Ispra: Joint Research Centre.

Obadia T, Haneef R, and Boëlle P-Y. 2012) The R0 package: A toolbox to estimate reproduction numbers for epidemic outbreaks. BMC Medical Informatics and Decision Making, 12(1): 147. https://doi.org/10.1186/1472-6947-12-147

Peeri, NC. et al. 2020. The SARS, MERS and novel coronavirus (COVID-19) epidemics, the newest and biggest global health threats: what lessons have we learned? International Journal of Epidemiology, 2020, 1–10, doi: 10.1093/ije/dyaa033.

Verity, R et al. 2020. Estimates of the severity of coronavirus disease 2019: a model-based analysis. The lancet, doi: 10.1016/S1473-3099(20)30243-7.

Wallinga J and Lipsitch M. 2007. How generation intervals shape the relationship between growth rates and reproductive numbers. Proceedings of the Royal Society B: Biological Sciences, 274(1609), 599–604. https://doi.org/10.1098/rspb.2006.3754

Zhanwei D, Xu X, Wu Y, Wang L, Cowling BJ, and Ancel Meyers L. 2020. Serial Interval of COVID-19 among Publicly Reported Confirmed Cases. Emerging Infectious Disease Journal, 26(6). https://doi.org/10.3201/eid2606.200357.

List of figures

Annexes

Annex 1. Data and methods

The European Surveillance System (TESSy) database developed at the European Centre for Disease Prevention and Control (ECDC) includes detailed information on 574 000 positive cases as of May 11. There are noticeable differences in data provision between countries, reflecting differences in the epidemic phase and testing procedures[12] adopted during the periods of data collection. Table A1.1 and Figure A1.1 and A1.2 provide an overview of the sample composition.

We used the number of cases provided in TESSy without filtering. The published numbers in TESSy might suffer from discrepancies due to the fact they are sourced from different institutions, with different methodologies and across different time zones, and need time to be validated properly. While the numbers have to be used with caution and users should be aware of the limitations, they still provide some valuable information.

Given the low coverage we excluded from the analyses data from Bulgaria, France, Croatia, Hungary and UK.

Data for France at provincial level were not available in ECDC-TESSy and were downloaded from Géodes (Géo données en Santé Publique) (April 29, 2020). The reporting system in Géodes is not exhaustive and the number of reporting establishments varies over time. Some patients, present in the hospital database at a given time, are withdrawn from the database by healthcare establishments when the patient's biological result is negative compared to Covid-19.

The data for the US counties was downloaded from USAFacts (https://usafacts.org) USAFacts.org on 28 April. USAFacts relies on Centers for Disease Control and Prevention (CDC), state- and local-level public health agencies. County-level data is confirmed by referencing state and local agencies directly. The time period covered is 22 January to 20 April 2020 and covers 771 442 confirmed cases.

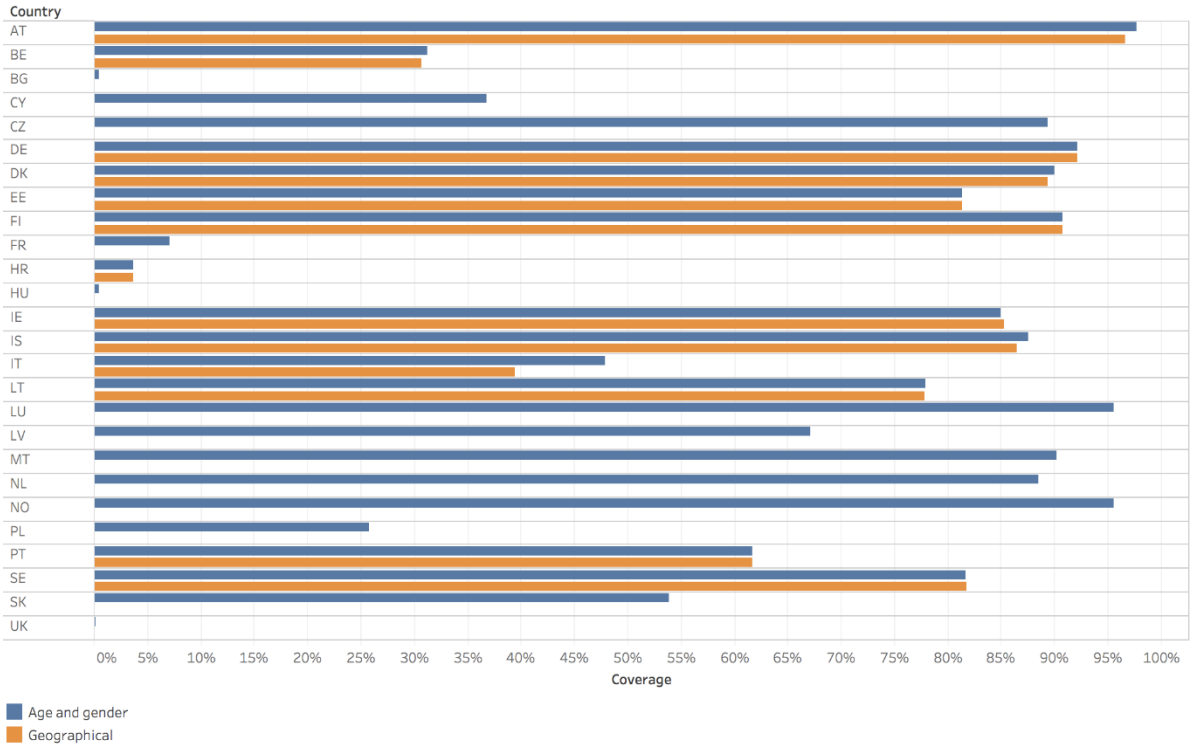
[12] While the World Health Organisation (WHO) provides recommendations on testing all symptomatic individuals with fever, cough and/or difficulty breathing, screening procedures for COVID-19 that are adopted by national health authorities can vary between and within countries.

Table A1.1. ECDC TESSy sample composition

	Cases			% of Total cases			% of Total fatalities		
	Female	Male	Unknown	Female	Male	Unknown	Female	Male	Unknown
AT	8,068	7,668		2.6%	2.9%		1.7%	1.4%	
BE	7,650	6,756	32	2.5%	2.6%	0.7%			
BG	3	3		0.0%	0.0%				
CY	439	439		0.1%	0.2%		0.0%	0.0%	
CZ	4,100	3,729		1.3%	1.4%		0.3%	0.2%	
DE	88,331	80,497	396	28.7%	30.7%	8.7%	17.4%	15.2%	3.5%
DK	5,991	4,365	1	1.9%	1.7%	0.0%			
EE	953	771		0.3%	0.3%		0.1%	0.0%	
EL			2,141			47.1%			86.8%
FI	3,037	2,961		1.0%	1.1%		0.3%	0.2%	
FR	4,551	4,202	1,319	1.5%	1.6%	29.0%	0.6%	0.5%	8.8%
HR	28	46		0.0%	0.0%		0.0%		
IE	12,808	9,508	68	4.2%	3.6%	1.5%	3.1%	2.2%	
IS	822	797		0.3%	0.3%		0.0%	0.0%	
IT	101,591	91,848	522	33.1%	35.0%	11.5%	56.9%	64.1%	0.9%
LT	807	582		0.3%	0.2%		0.2%	0.1%	
LU	1,828	1,896		0.6%	0.7%		0.2%	0.1%	
LV	324	373	1	0.1%	0.1%	0.0%			
MT	199	301	1	0.1%	0.1%	0.0%	0.0%	0.0%	
NL	27,158	15,840	59	8.8%	6.0%	1.3%	15.1%	12.9%	
NO	4,070	4,035		1.3%	1.5%		0.6%	0.5%	
PL	3,268	2,681		1.1%	1.0%		0.5%	0.5%	
PT	15,626	10,925		5.1%	4.2%		2.9%	1.9%	
SE	14,961	11,196	3	4.9%	4.3%	0.1%			
SK	705	692	1	0.2%	0.3%	0.0%	0.1%	0.0%	
UK	47	74	1	0.0%	0.0%	0.0%			

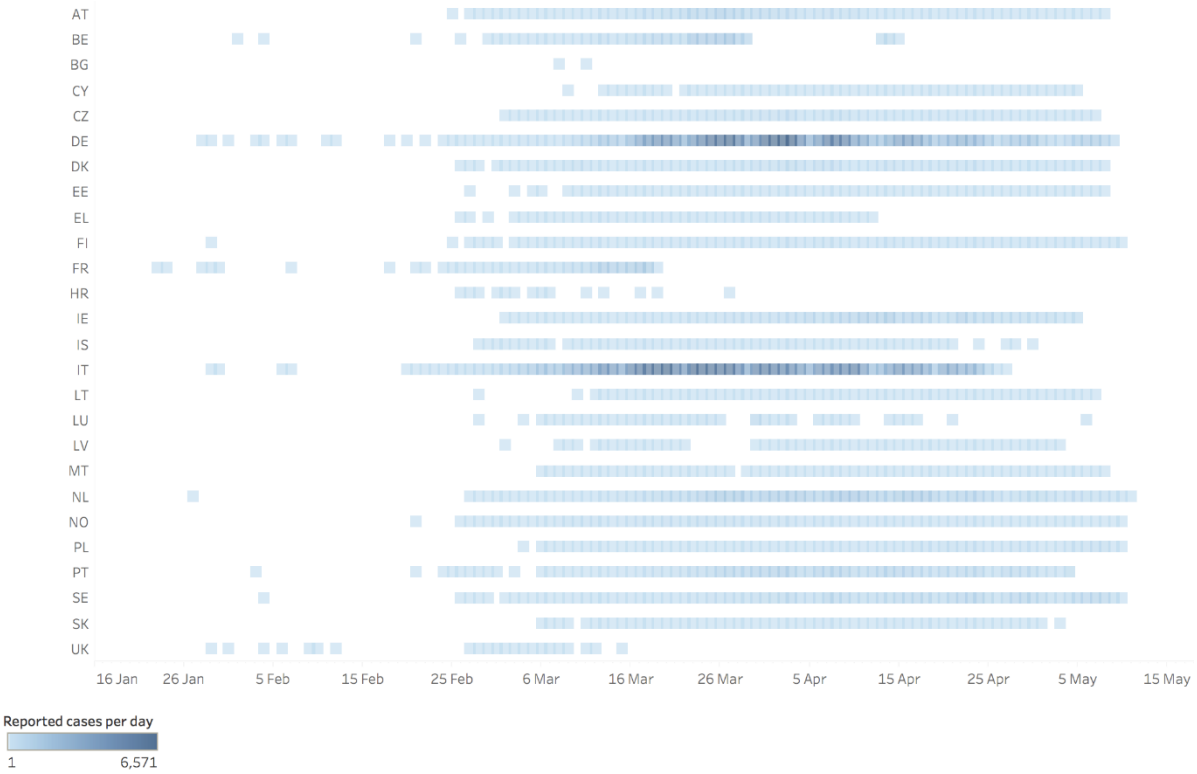
Source: ECDC TESSy (May 11, 2020).

Figure A1.1. Overview of coverage of the data used in the analysis. The bars represent the percentage of cases with detailed information on age, gender and province in respect of the aggregated figures.



Source: ECDC TESSy (May 11, 2020).

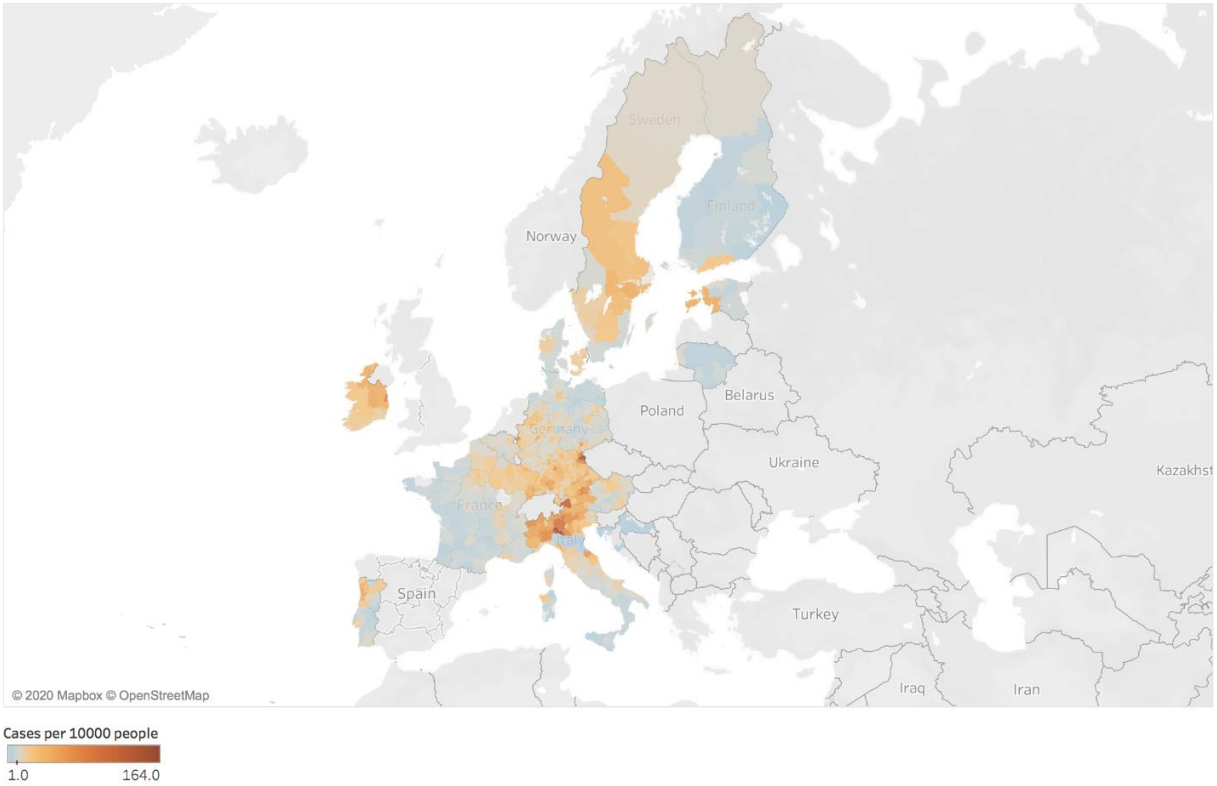
Figure A1.2. Temporal coverage of cases reported in ECDC-TESSy (May 11, 2020).



Source: ECDC TESSy (May 11, 2020).

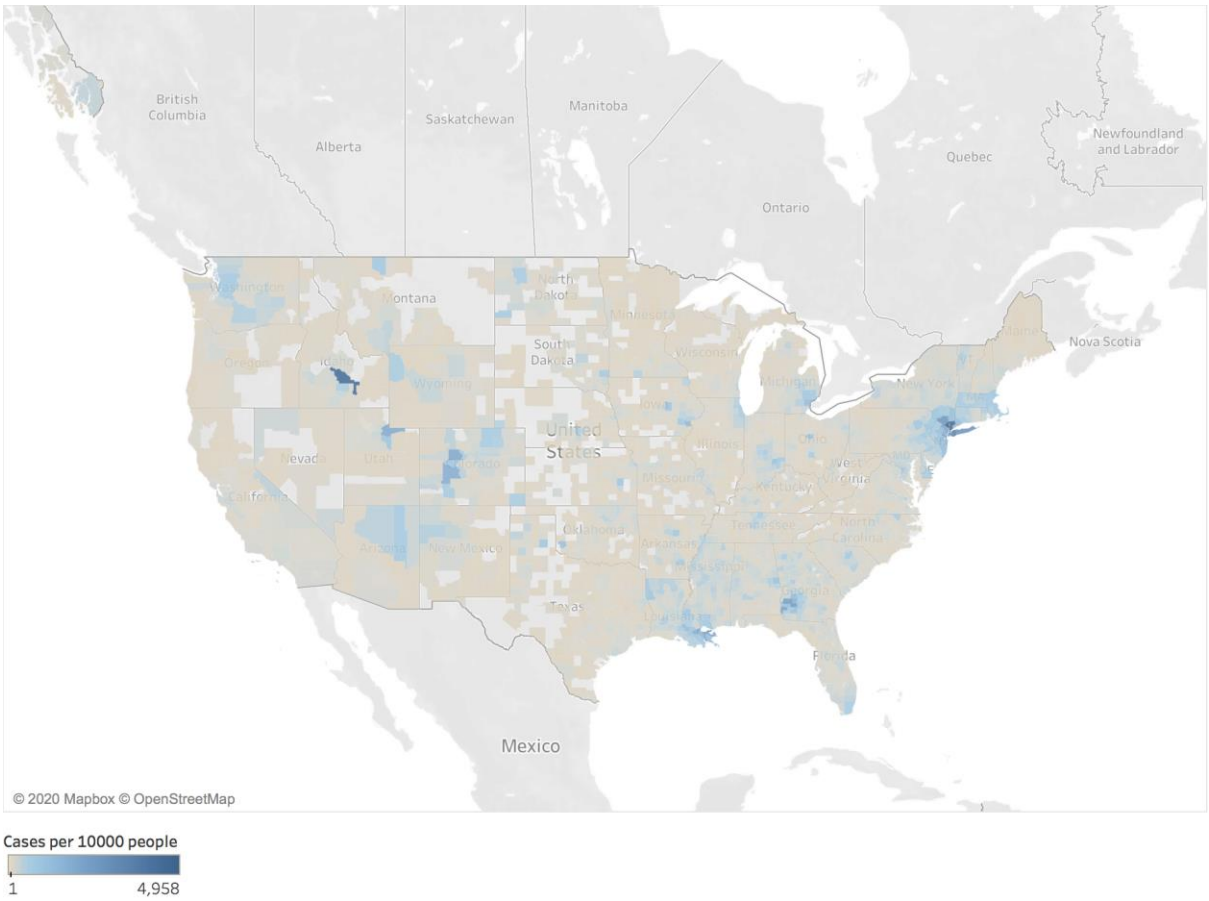
Annex 2 Geographical coverage of territorial data

Figure A2.1. Geographical coverage of the data used in the territorial analysis for the EU. The colours represent the number of cases per 10000 people. The colour scale is centred at the median value of 10.9. The data only includes cases for which information on the place of infection or residence was available at the provincial level (NUTS3)



Source: ECDC-TESSy (May 11, 2020). Data for France at provincial level were not available in ECDC-TESSy and were downloaded from https://geodes.santepubliquefrance.fr/ (April 29, 2020).

Figure A2.2. Geographical coverage of the data on cases in the territorial analysis for the US. The colours represent the number of cases per 10,000 people. The colour scale is centred at the median value of 72



Source: https://usafacts.org/visualizations/coronavirus-covid-19-spread-map/.

Annex 3 R0 calculation method

A key indicator to monitor the spread of a disease and to design control measures and vaccination campaigns is the reproductive number R0. R0 by itself does not tell how fast the disease is spreading but rather how difficult it is to contain it. The value of R0 is affected by the infectiousness of the pathogen, the contacts' intensity and the infectivity of individuals.

Taking as constant infectiousness and infectivity, we can assume that changes in R0 across territories are depending on the effect of population density on the intensity of contacts. The rationale is that higher population density which is reflected in the definition of the urban typology may imply a higher R0 by increasing the probability of contacts between individuals.

One of the most common approaches to calculate the R0 especially in the initial phases of the epidemic is to examine the observed exponential epidemic growth rate of cases (Wallinga & Lipsitch, 2007).

The mathematical relation between R0 and the exponential growth rate is described by the following equation.

$$R = 1/M(-r)$$

Where M is the moment generating function of the (discretized) generation time distribution and r is the exponential growth rate in our case estimated using a Poisson distribution.

The generation time distribution refers to the time lag between infection in a primary case and a secondary case. Since this is often not available, as a substitute we use the serial interval distribution which is defined as the period between the appearance of symptoms in the infector and in the infectee. For COVID-19 we rely on values reported in (Zhanwei Du et al., 2020) which have been calculated examining 468 confirmed cases of 2019 novel coronavirus disease reported in China as of February 8, 2020 (mean interval of 3.96 days, SD 4.75 days).

For the calculation we use the package R0 (Obadia et al., 2012). After constructing the specific epidemic curves for each territorial unit (NUTS3 for the EU and counties for the US) we iterate the calculation of R0 for each curve. Finally, we represent the results by different groupings of rural-urban typologies.

Another important aspect to consider in the calculation, is that R0 changes depending on the phase of the epidemic, starting from the highest value during the exponential growth phase and gradually approaching the target of below 1 which is signalling that the epidemic is dying out, either autonomously or under the effect of containment measures. We want to test if the differences in R0 between territories are persisting independently from the phase of the epidemic. For this we repeat the calculation of R0 considering increasing temporal windows of 5 to 40 days and including in each calculation all the territories which have the epidemic lasting at least for the duration of each temporal window.

GETTING IN TOUCH WITH THE EU

In person

All over the European Union there are hundreds of Europe Direct information centres. You can find the address of the centre nearest you at: https://europa.eu/european-union/contact_en

On the phone or by email

Europe Direct is a service that answers your questions about the European Union. You can contact this service:

- by freephone: 00 800 6 7 8 9 10 11 (certain operators may charge for these calls),

- at the following standard number: +32 22999696, or

- by electronic mail via: https://europa.eu/european-union/contact_en

FINDING INFORMATION ABOUT THE EU

Online

Information about the European Union in all the official languages of the EU is available on the Europa website at: https://europa.eu/european-union/index_en

EU publications

You can download or order free and priced EU publications from EU Bookshop at: https://publications.europa.eu/en/publications. Multiple copies of free publications may be obtained by contacting Europe Direct or your local information centre (see https://europa.eu/european-union/contact_en).

KJ-NA-30237-EN-N

The European Commission's science and knowledge service

Joint Research Centre

JRC Mission

As the science and knowledge service of the European Commission, the Joint Research Centre's mission is to support EU policies with independent evidence throughout the whole policy cycle.

EU Science Hub

ec.europa.eu/jrc

@EU_ScienceHub

EU Science Hub - Joint Research Centre

EU Science, Research and Innovation

EU Science Hub

Publications Office of the European Union

doi:10.2760/838390

ISBN 978-92-76-19180-3