COVID-19 Cases and Case Fatality Rate by age

Fabrizio NATALE, Daniela GHIO, Dario TARCHI, Anne GOUJON, Alessandra CONTE

Abstract

The report provides a comparison of cases and fatalities of COVID-19 by age and by provinces and regions in some of the most affected countries. The main objective of the comparison is to evaluate possible effects of the demographic characteristics of population on the epidemic outcomes. The analysis is preliminary since it is based on constantly evolving data while the COVID-19 pandemic is still unfolding. Age distributions are presented by large age classes and using estimated values for harmonised 5-year age groups. In addition to Case Fatality Rates, we calculate three main indicators: Synthetic CFR, Relative Illness Ratio and Relative Mortality Ratio. Finally, we examine the relation between the geographical patterns of cases and fatalities and the demographic profiles of the population. The distribution of both cases and fatalities across ages shows several anomalies. However our findings seem to exclude the fact that an older population alone may justify the high number of fatalities recorded in particular in the case of Italy.

Key messages

- Given the fact that fatalities caused by COVID-19 are mostly concentrated in older age groups who suffer from underlying medical conditions and with considerable differences between males and females, for the management of sanitary response, it is important to avail of harmonised data on cases and fatalities by age and gender.
- Such data needs to be adjusted considering the underlying differences in age structure.
- Despite the demographic adjustments, the distribution of both cases and fatalities across ages shows several anomalies. A striking anomaly is in the case of Italy which has a high Case Fatality Rate for ages above 60 years in respect to other countries.
- The literature points to the fact that differences in age distributions of the Case Fatality Rate may originate from the procedure adopted in each country for testing and reporting positive cases and fatalities attributable to COVID-19 and to demographic factors.
- In our analysis, when considering fatalities in relative terms across age groups rather than the CFR, we don't find significant differences between countries, since 94% of fatalities are uniformly concentrated in the population over 60 years of age.
- A territorial analysis of the distribution of cases and fatalities shows that there are clear geographical patterns in the diffusion of the disease, but these patterns are not related to differences in the share of elderly across provinces and regions.
- These last two findings seem to exclude the fact that an older population alone may justify the high number of fatalities recorded in particular in the case of Italy.

Introduction

While the COVID-19 pandemic is still unfolding, it is of paramount importance to collect and analyse epidemiological data by age and compare distribution across countries.

Since the disease has the worst consequences for older age groups, the sanitary response has to be tailored considering the demographic structure of the population both at national and regional level. Demographic science has an important role to play in explaining differences in fatality rates across countries [1].

The purpose of this note is to provide a first analysis of data on cases and fatalities by age collected from online sources and reports by Ministries of Health in the most affected countries.

To our knowledge such data is still not systematically collected and harmonised at international level and the analyses conducted so far in early papers in the epidemiology literature are not including cross country comparisons.

Since the data is not harmonised these comparisons need to be considered with caution. Differences in the distribution by age of cases and fatalities may be strongly affected by the sampling strategies and ways of reporting fatalities attributable to COVID-19, rather than by differences in morbidity and mortality.

The lack of harmonisation ultimately affects the comparability of Case Fatality Rates (CFR), through different types of biases as described in the Box below.

In the note we present figures of cases, fatalities and of the CFR by age updated as of 30/03/2020. These figures are presented by large age classes and using estimated values for harmonised 5-year age groups.

In addition to the CFR, we calculate three main indicators: Synthetic CFR (SCFR), Relative Illness Ratio (RIR) and Relative Mortality Ratio (RMR). These indicators are used to adjust for differences in age structure and underlying baseline general mortality. The first indicator provides a single measure of mortality to COVID-19 for the entire country while the other two indicators give a profile by age respectively for cases and mortality.

Finally in the last chapter we include an analysis of data on cases and fatalities at provincial and regional level and we examine their relation with the age structure of population in the same territories through simple correlation.

Limitations in using the case fatality rate (CFR) to monitor and compare the spread of the epidemic across countries and over time

The CFR is most-often referred to as the case-fatality rate or ratio. The CFR is a measure of the severity of the condition as it corresponds to the proportion of the population with a certain condition who die from that condition during the reference period. Only ascertained cases are included in the denominator.

During an outbreak the CFR can be over- or underestimated [1]. As the COVID-19 pandemic is still unfolding, the individual outcome (recovery or death) is known only for a partial proportion of infected patients. In contrast to SARS, the majority of COVID-19 infections are asymptomatic or mild and do not require hospitalization [2]. Differential case ascertainment (in turn determined by symptomatology, care-seeking behaviour, clinical access, and testing criteria) thus becomes an additional and important source of bias. The number of positive cases depends on the screening policy at national and sub-national level, and the adoption of procedures (i.e. who should be tested and which requirements are adopted) are decided at international level by the WHO. Yet, screening policy has also changed over time; furthermore, its application can differ at local level, specifically within countries that have a regionalised health system.

Finally, CFR is a crude indicator, which does not account for the changes in the demography of positive cases and deaths during the different stages of an epidemic, which are commonly characterised by a concentration of infections at the early stage among the most vulnerable age groups.

Cases and Case Fatality Rates by broad age groupings

Ministries of Health in the affected countries are producing regular updates of the evolution of the pandemic within their territories. Reports in most countries include figures on the evolution of confirmed cases and fatalities at the national level, by regions and by age. Table 1 and 2 in the Annex provide a summary of the raw data extracted from online sources as of 30/03/2020.

In order to compare figures across countries, Figure 1 provides a higher level of aggregation of cases, fatalities and CFR by three main age classes of 0 to 20, 20 to 60 and over 60. Whenever available, updates are included in the figure with symbols of different shading.

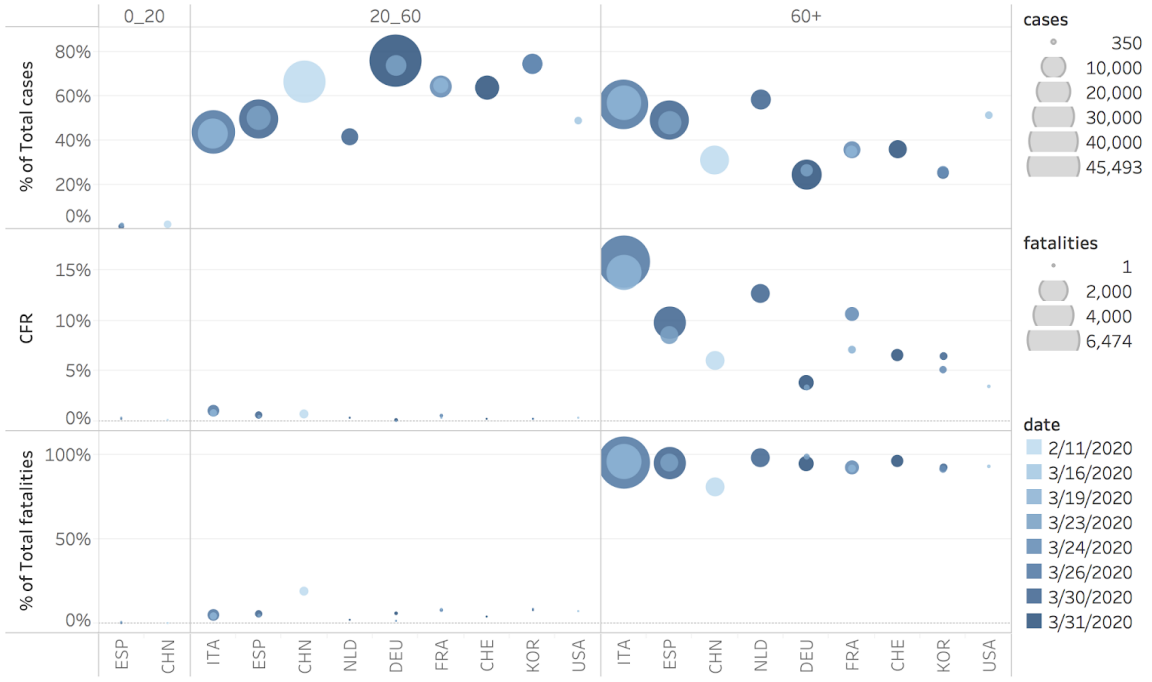

Figure 1 COVID-19 cases, fatalities and CFR by broad age groups. Cases and fatalities are shown as a percentage of the total in each country. The shading of the symbols corresponds to different updates of the data and the size is representing the number of cases (upper charts) and the number of fatalities (two lower charts).

Countries with the highest share of cases in the age group 20-60 are Germany (76%), South Korea (75%), China (67%) and France (64%). In Spain and the USA, cases are equally distributed between the age group 20-60 and over 60. Netherlands and Italy concentrate most of the reported cases in the group over 60, respectively 58% and 56%.

Since the number of cases is the denominator for the calculation of the CFR, the anomaly about the age distribution of cases in Italy[1] and the Netherlands also affects the value of the CFR. The CFR for Italy in the class over 60 years of age is 16%, for Netherlands 13%, while it is 7% in the other countries summed up together.

The bottom part of the figure shows the relative distribution of fatalities across the age groups in each country. When considering fatalities in relative terms the anomalies of Italy and Netherlands disappear. All countries have a similar distribution of fatalities and when summing the values for all countries, around 94% of fatalities are concentrated in the age group of 60 years and over. The main difference remains for China which has a share of 81% in the case of the class over 60 years of age and 19% in the group 20-60.

By assuming a high risk of deaths in the case of older people, a country with an older population should have not only more fatalities in absolute terms but also a higher share in the older group. The fact that this is not observed in the data leads to the suspicion that differences in age structure alone at national level are not sufficient to explain the high number of fatalities in some countries.

When comparing changes over time there are no particular differences emerging in the distributions by ages for all indicators considered, therefore in the subsequent analyses we refer to the latest data available.

[1] The anomaly of the CFR for Italy is reported also in [4]

Comparison of Case Fatality Rates across countries

The higher CFR in the older age classes in Italy compared to other countries with available data is confirmed when considering a continuous distribution with 5-year age groups (Figure 2).[2]

The CFR for all countries is starting to increase after age 50, however, the increase is steeper in the case of Italy and Netherlands.

The other countries have more consistent distribution of the CFR which is not exceeding 20% in the age class of 80-85[3] in contrast to 25% in the case of Italy and 22% for Netherlands.

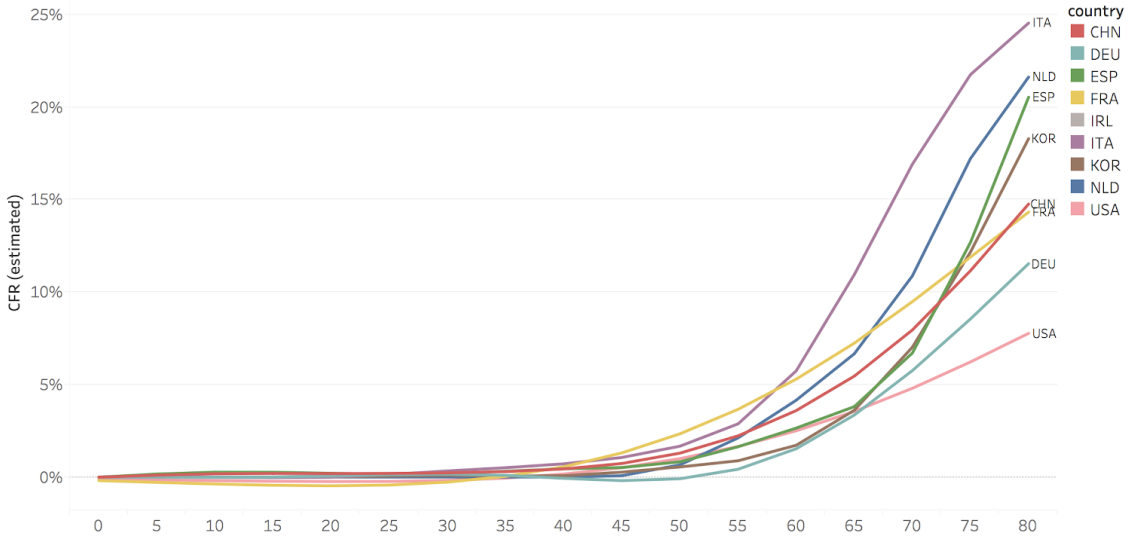

Figure 2 COVID-19 cases and CFR by harmonised 5-years age groups.

[2] In order to allow comparison across countries and address the issue of difference in reporting the age distribution in the figure has been fitted on the original data using a Spline nonparametric regression technique.

[3] The CFR is not estimated for ages above 80-85 since the data are not available in most countries.

Demographic adjustment of the Case Fatality Rate

As above described, there are many biases that can affect the CFR during the course of an emerging epidemic, such as the deficit in harmonised screening procedures and the lack of individual longitudinal data on the total number of cases.
However, bearing in mind this limitation, the CFR needs to be corrected for the changing demographics of infected cases and fatalities over time. In fact, as demonstrated by the overlapping age pyramids in Figure 3, the age-structure of positive cases in Italy has been evolving over time.

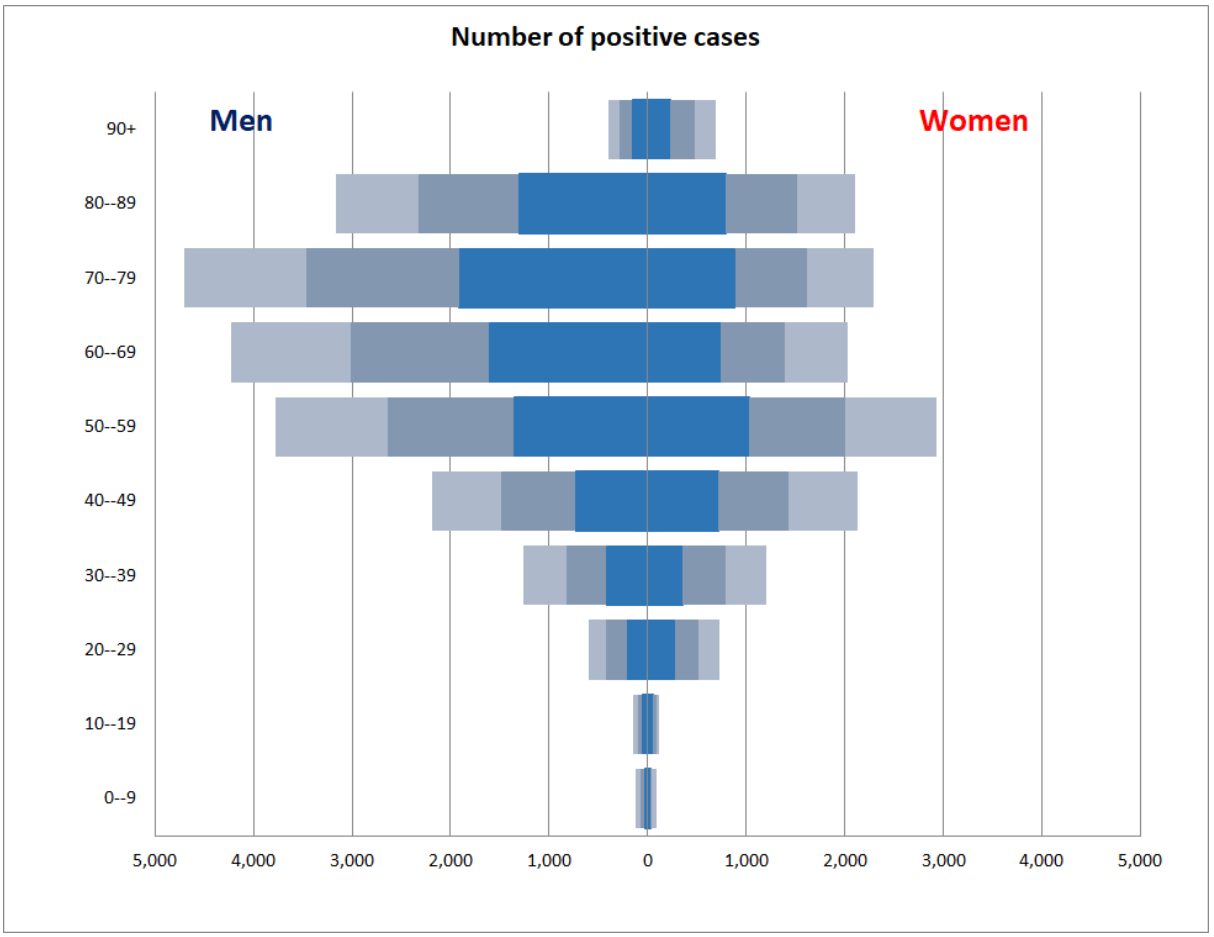

Figure 3 Age structure of positive cases, Italy. Legend: dark blue bars correspond to cases recorded as of March 12; light grey bars correspond to cases recorded as of March 16; light blue bars correspond to cases recorded as of March 20.

Synthetic CFR (SCFR)

As mentioned before CFR is a crude rate, representing the proportion of persons with a particular condition (cases) who die from that condition. In order to estimate the impact of the age structure, we calculate age-specific CFR ($CFR_{i,,i+n}$) for each age group between age i and age i+n (*i, i+n)* using Equation (1)

$$CFR^{i}_{t} = D^{i}_{t,t+1} / C^{i}_{t,t+1} \quad \text{(Eq. 1)}$$

where $D_{i,t+1}$ corresponds to the number of deaths at age i in the time between time t and t+1, and $C_{i,t+1}$ corresponds to the number of positive cases for the same age-group and during the same period.

We aggregate CFRs into a synthetic CFR (SCFR) that is calculated as shown in Equation (2):

$$SCFR_{t} = \sum_{n}^{i} CFR_{t} \quad \text{(Eq. 2)}$$

with n being the width of the age intervals.

The SCFR can be interpreted as the average number of deaths per confirmed case of COVID-19.

Country	Gender	Date	Cases	Deaths	CFR	SCFR	Mean age at death
USA (New York)	Total	Mar-26	21848	267	1.22	2.12	82.38
Belgium	Total	Mar-25	4909	178	3.63	3.7	80.94
Spain	Total	Mar-31	55484	3337	6.01	5.97	83.35
Spain	Men	Mar-31	27788	2103	7.57	6.92	83.42
Spain	Women	Mar-31	27682	1237	4.47	4.92	83.41
Spain	Total	Mar-25	32816	1326	4.04	4.09	83.81
Spain	Men	Mar-25	16444	828	5.04	4.72	83.93
Spain	Women	Mar-25	16363	498	3.04	3.38	83.76
Spain	Total	Mar-22	18959	805	4.25	4.53	84.29
Spain	Men	Mar-22	9736	518	5.32	5.07	84.31
Spain	Women	Mar-22	9182	287	3.13	3.85	84.5
Italy	Total	Mar-30	94312	10026	9.23	8.60	82.11
Italy	Men	Mar-30	52206	6930	13.27	10.61	81.81
Italy	Women	Mar-30	41549	3083	7.42	6.24	83.10
Italy	Total	Mar-23	57989	5019	8.66	7.09	82.44
Italy	Men	Mar-23	33399	3544	10.61	8.44	82.12
Italy	Women	Mar-23	24045	1449	6.03	5.23	83.74
Italy	Total	Mar-20	35731	3047	8.53	7.01	82.6
Italy	Men	Mar-20	20688	2139	10.34	8.16	82.21
Italy	Women	Mar-20	14374	890	6.19	5.49	83.89
Italy	Total	Mar-16	24879	1697	6.82	5.58	83.06
Italy	Men	Mar-16	14621	1697	6.82	5.58	83.06
Italy	Women	Mar-16	9867	493	5	4.35	84.28
Italy	Total	Mar-12	13317	785	5.89	4.77	83.79
Italy	Men	Mar-12	7815	572	7.32	5.65	83.59
Italy	Women	Mar-12	5113	212	4.15	3.6	84.66
China	Total	Feb-22	44672	1023	2.29	3.77	81.9

Table1 CFR and SCFR by country and gender

Overall Italy shows higher fatality rates values than China and Spain. The difference is reduced when the SCFR is considered instead of the CFR. In addition, looking at the four different time points available for Italy, the SCFR at the beginning of the epidemic, on 12 March, is not much different (4.77%) than that of China (3.77%), and very close to the SCFR of Spain (4.55%). For Belgium, which has a significantly lower number of total cases, the SCFR value is lower and close to that of China.

During the latest period of observation since March 16, COVID-19 severity in Italy has evolved. On March 19, the adjusted SCFR increased from 5.6% to 7%, while the mean age at death decreased from 83.0 to 82.6 years. The adjusted SCFR for men had reached 8.2% with an estimated mean age at death at 82.2 years, against an adjusted SCFR among women that stood at 5.5% with an estimated mean age at death at 84 years. On March 23, the adjusted SCFR was 0.08% higher than on March 19, against 0.13% increase detected when CFR is used, and against 1.5% increase from March 12 to March 19. Looking at female SCFR, there is a first sign of a decline, from 5.49 to 5.23%.

Using SCFR, *estimated* trajectories of epidemic stages can be traced (Figure 4).

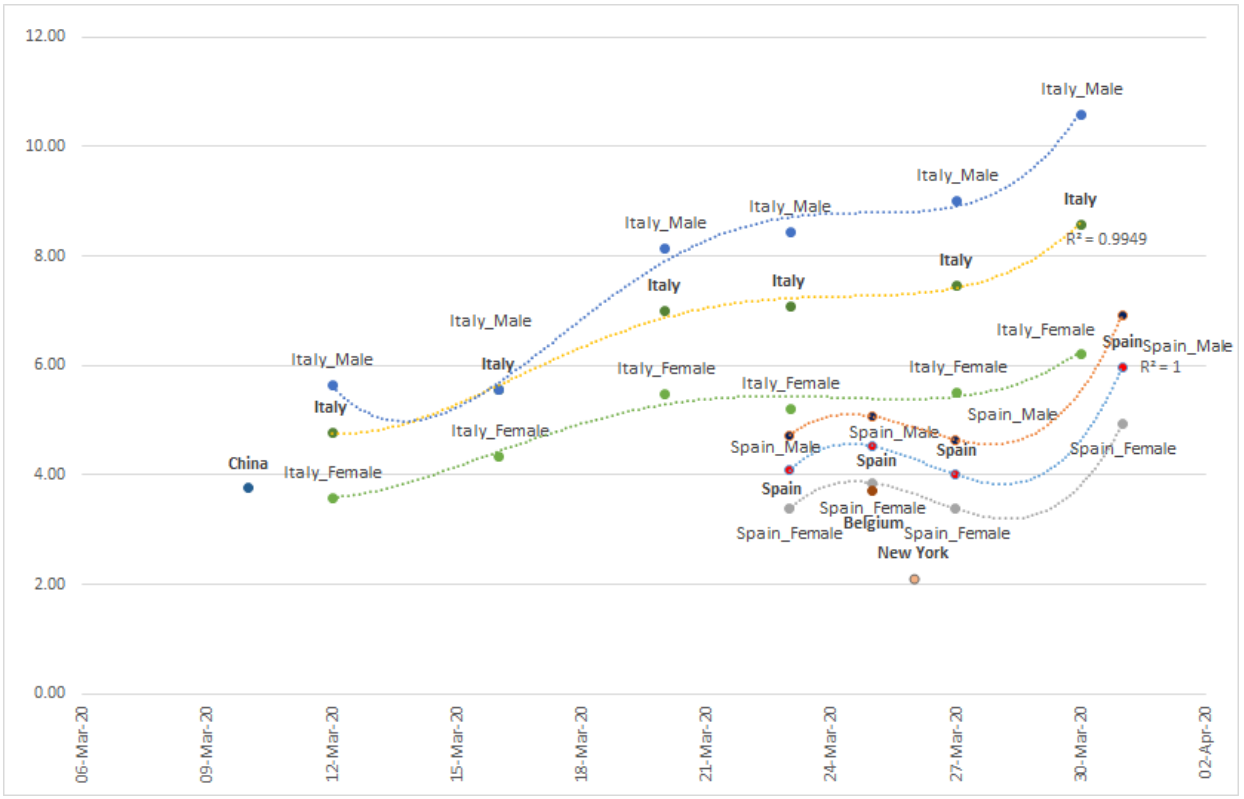

Figure 4 Overview of epidemic stages

The application of the method can be extended to the regional/sub-national level, as soon as appropriate data will be available. We are already working to harmonise data at regional levels for Italy. In addition, it is planned to systematically re-calculate the indicator at the country level on a regular basis (ideally daily but weekly at the minimum) over time as new updated data become available.

Relative Illness Ratio and Relative Mortality Ratio

To further compare differences in the age distribution of cases and fatalities across countries we rely on the following two indicators, which have been already used to explore age profiles for influenza epidemics [3] and recently for an early epidemiological analysis of COVID-19 in China [5]. Both measures provide a way to standardise absolute figures of morbidity and mortality by age and showing differences in risk across population groups.

Relative Illness Ratio

The Relative Illness Ratio (RIR) is a measure of morbidity calculated according to the following formula (Equation 3) as the share of Covid-19 cases in each age group in respect to the total number of morbidity cases divided by the share of population in the same age group.

$$RIRi = \frac{(Ccovid_i / \sum_i Ccovid_i)}{(P_i / \sum_i P_i)} \quad \text{(Eq. 3)}$$

where *Ccovid* is the estimated number of cases of COVID-19 in age group *i* and *P* is the population in the same age group.

A value above 1 indicates that there is a higher proportion of cases in the age group relatively to other age groups, after taking into account differences in the size of the population.

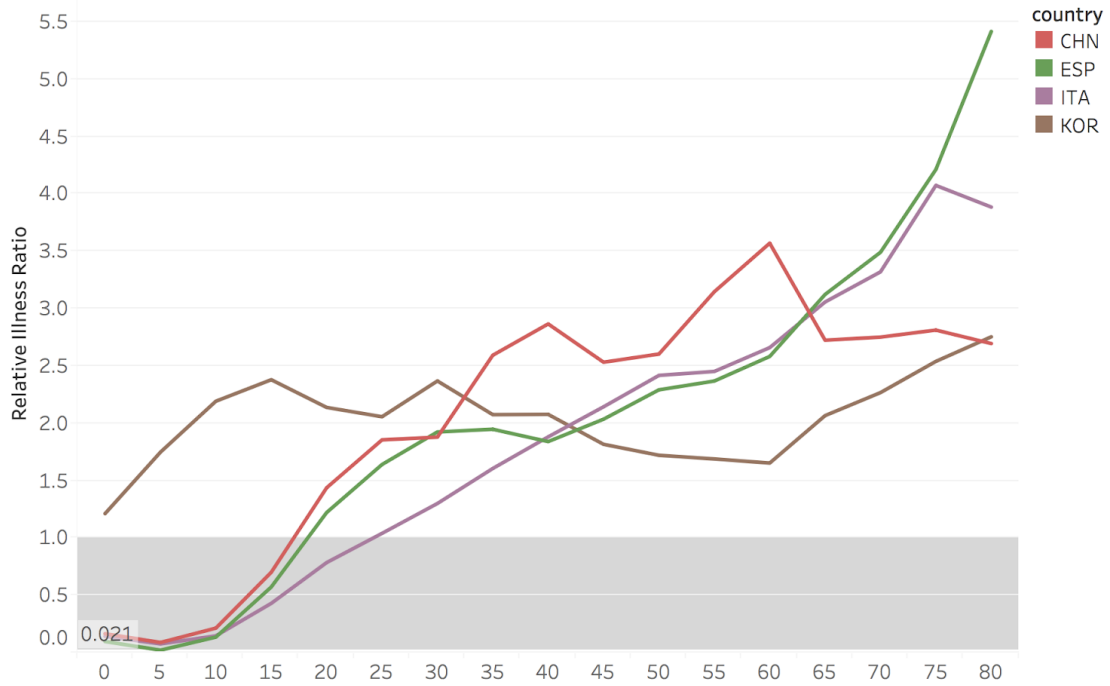

Figure 5 Relative Illness Ratio by 5-year age groups[4].

[4] The indicators are only presented for China, South Korea, Spain and Italy where the number of observations are sufficient to estimate a standardised age distribution.

The RIR in China appears skewed towards the age of 40. This is even more evident in the case of South Korea. The higher proportions of cases in younger age groups for these two countries could be explained by a more extensive testing and hence a better measurement of morbidity at these ages, compared to Italy.

Relative Mortality Ratio

The Relative Mortality Ratio (RMR) is calculated as the share of the fatalities to COVID-19 for each age group in respect of the total number of fatalities, divided by the baseline mortality rate for all causes of death (Equation 4).

$$RMRi = \frac{(Fcovid_i / \sum_i Fcovid_i)}{(M_i / \sum_i M_i)} \quad \text{(Eq. 4)}$$

where *Fcovid* is the number of fatalities linked to COVID-19 in age group *i* and *M* is the total fatalities[5] for all causes in the same group.

In the case of RMR, a value above 1 indicates that there is a higher fatality linked to the COVID-19 in the age group compared to other age groups, after taking into account the age specific general mortality rates.

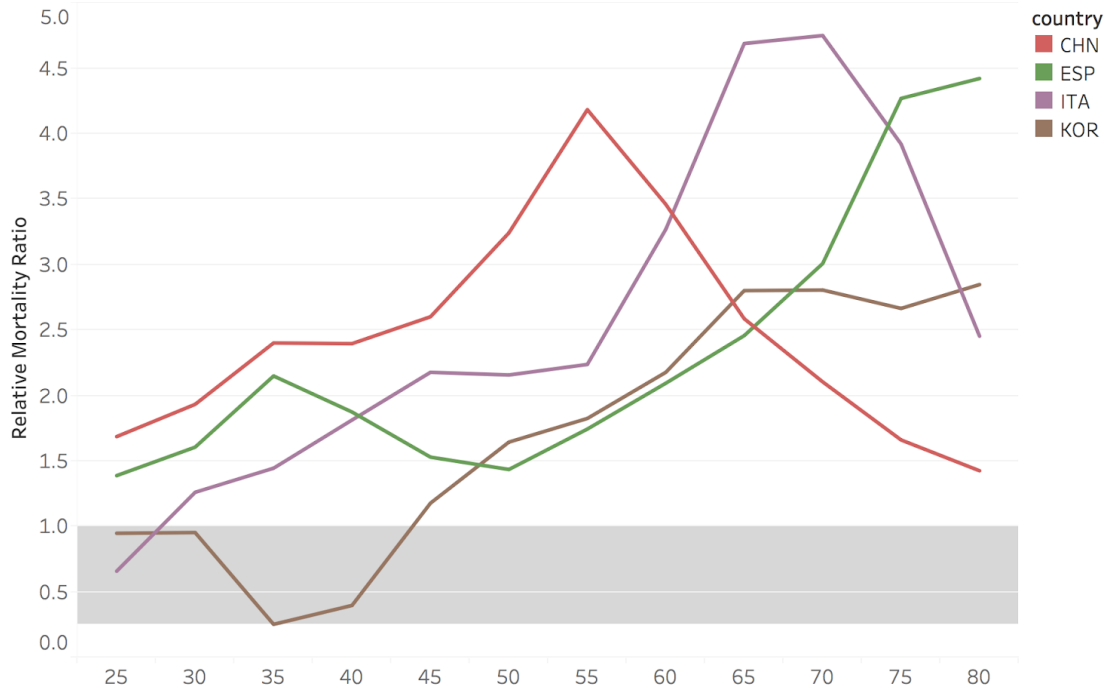

Figure 6 Relative Mortality Ratio by 5-years age groups.

The RMR indicator is reaching the highest value in the case of Italy at age 65-70. The distribution in Italy is similar to the one of South Korea; however the level of risk in the case of Italy is higher for all ages between 55 and 80. For China, the RMR distribution is skewed towards younger age groups with a peak at age 50.

[5] Data for fatalities was obtained from Eurostat demo_magec (2018) in the case of EU countries and from UNDESA (annual average for the period 2015 - 2020) for the other countries.

Cases and fatalities in relation to the share of elderly by regions and provinces

The distribution of cases in respect to the elderly population at the level of European regions and provinces[6] show some clear patterns with a high concentration in a few areas, particularly around the northern provinces in Italy and the southern provinces in Austria and Germany (Figure 7) and around the main cities of Madrid, Paris and Milan. Only in Spain is the spread more evenly distributed between all the regions, with La Rioja and Madrid being more affected.

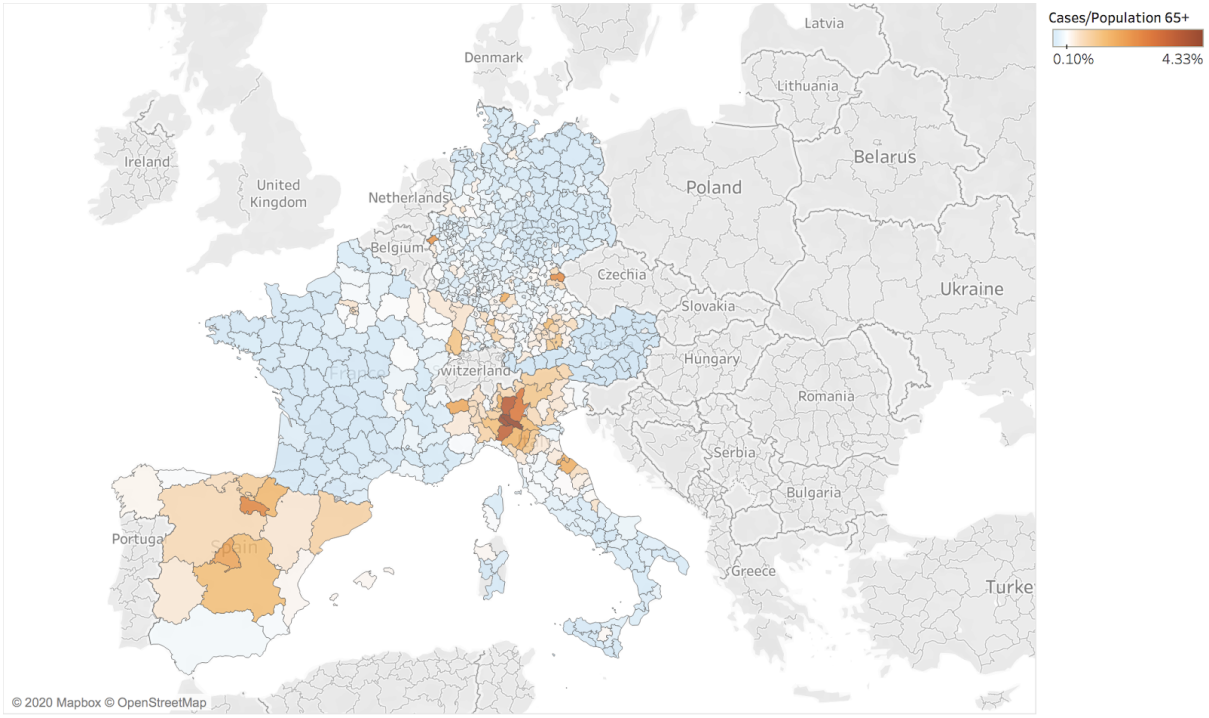

Figure 7 Territorial distribution of the share of cases in respect to the elderly population.

In addition to Lombardy and Madrid, high levels of the CFR are recorded in the case of the French provinces of Vosges, Haut-Rhin, Moselle and Territoire de Belfort, close to the German border (Figure 8).

[6] The coverage and geographical resolution of the data varies across countries. See Table 2 in the annex for details.

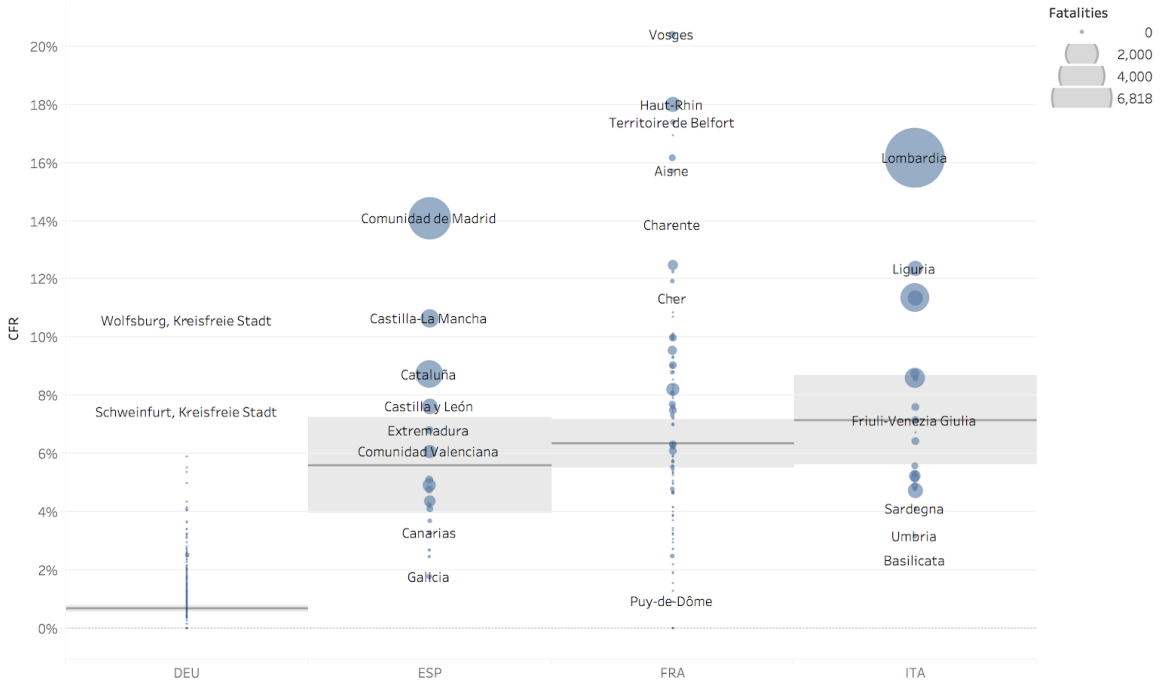

Figure 8 Case Fatality Rate by provinces and regions. Size of the symbol is proportional to the absolute number of fatalities. Horizontal lines indicate the average and the shading 95% confidence intervals.

Looking at the relationship between the share of elderly in the population of the provinces and regions in the few countries with data (Austria, France, Germany, Italy and Spain), no clear patterns emerge (Figure 9). The correlation between the two variables is very low and negative (-0.06). This seems to contradict the idea that the number of fatalities would be largely driven by the share of the elderly population.

Other explanations might be at play that will need to be further researched. While the spread of epidemics is facilitated in densely populated urban areas, it might also be linked to the ties that unite communities and extended families, which are stronger in some regions even if they are less dense rural (or semi-rural) communities, than in others [6].

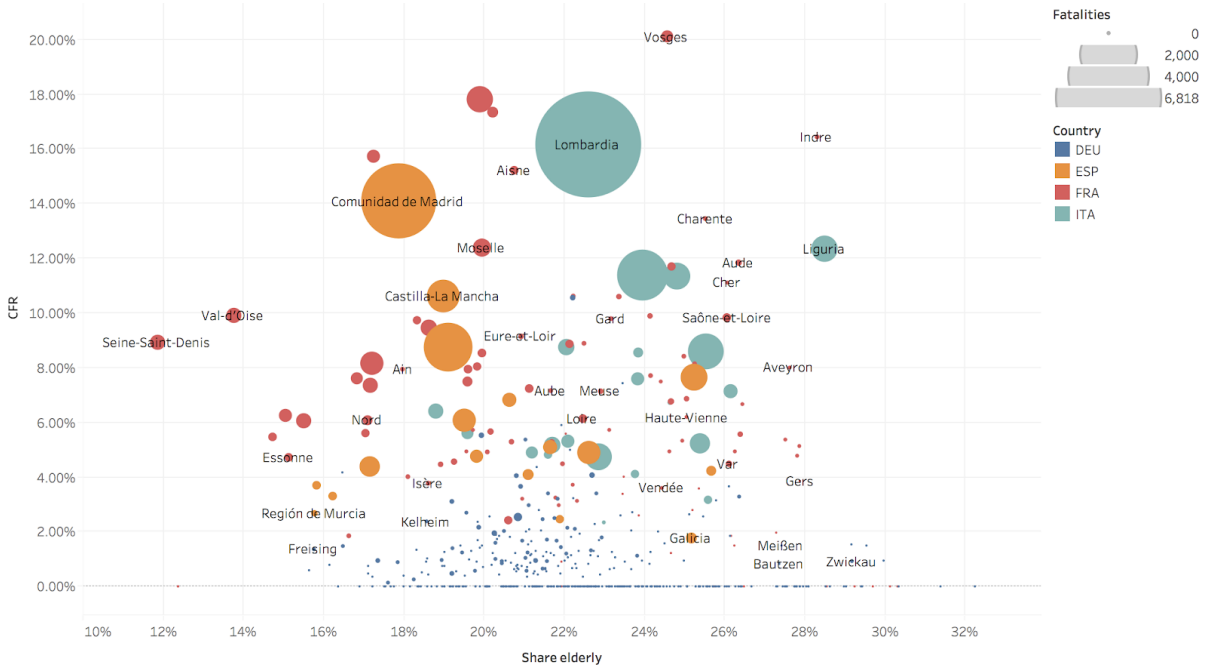

Figure 9 Case Fatality Rate by share of elderly. Size of the symbol is proportional to the absolute number of fatalities.

References

1. Lipsitch M, Donnelly CA, Fraser C, Blake IM, Cori A, Dorigatti I, et al. Potential Biases in Estimating Absolute and Relative Case-Fatality Risks during Outbreaks. Galvani AP, editor. PLoS Negl Trop Dis. 2015;9: e0003846. doi:10.1371/journal.pntd.0003846
2. SARS vs Wuhan coronavirus: Here's how the outbreaks compare - CNN. [cited 27 Mar 2020]. Available: https://edition.cnn.com/2020/01/29/china/sars-wuhan-virus-explainer-intl-hnk-scli/index.html
3. Lemaitre M, Carrat F. Comparative age distribution of influenza morbidity and mortality during seasonal influenza epidemics and the 2009 H1N1 pandemic. BMC Infect Dis. 2010;10: 162. doi:10.1186/1471-2334-10-162
4. Onder G, Rezza G, Brusaferro S. Case-Fatality Rate and Characteristics of Patients Dying in Relation to COVID-19 in Italy. JAMA. 2020 [cited 1 Apr 2020]. doi:10.1001/jama.2020.4683
5. Sun K, Chen J, Viboud C. Early epidemiological analysis of the coronavirus disease 2019 outbreak based on crowdsourced data: a population-level observational study. Lancet Digit Health. 2020;2: e201–e208. doi:10.1016/S2589-7500(20)30026-1
6. Youssef M, Scoglio C. An individual-based approach to SIR epidemics in contact networks. Journal of Theoretical Biology 2011, 283 (1): 136-144.

Annex

Table 1 Sources of data for fatalities and cases at national level

Country	Link	Update
CAN	https://www.canada.ca/en/public-health/services/diseases/2019-novel-coronavirus-infection/health-professionals/epidemiological-summary-covid-19-cases.html	31/03/2020
ITA	https://www.epicentro.iss.it/coronavirus/bollettino/Bollettino-sorveglianza-integrata-COVID-19_26-marzo%202020.pdf	26/03/2020
CHN	https://static-content.springer.com/esm/art%3A10.1038%2Fs41591-020-0822-7/MediaObjects/41591_2020_822_MOESM1_ESM.pdf	19/03/2020
USA	https://www.cdc.gov/mmwr/volumes/69/wr/mm6912e2.htm#T1_down	16/03/2020
ESP	https://www.imedhospitales.com/coronavirus/assets/Actualizacion_59_COVID-19.pdf	30/03/2020
FRA	https://www.santepubliquefrance.fr/maladies-et-traumatismes/maladies-et-infections-respiratoires/infection-a-coronavirus/documents/bulletin-national/covid-19-point-epidemiologique-du-15-mars-2023	24/03/2020
DEU	https://npgeo-corona-npgeo-de.hub.arcgis.com/datasets/dd4580c810204019a7b8eb3e0b329dd6_1	31/03/2020
BEL	https://epidemio.wiv-isp.be/ID/Documents/Covid19/Meest%20recente%20update.pdf	30/03/2020
KOR	https://www.cdc.go.kr/board/board.es?mid=a30402000000&bid=0035	30/03/2020
NLD	https://www.rivm.nl/nieuws/actuele-informatie-over-coronavirus/data	30/03/2020
IRL	https://www.gov.ie/en/publication/4a0175-an-analysis-of-the-2475-cases-of-covid-19-in-ireland-as-of-saturday-/#age-range-affected	28/03/2020
AUT	https://info.gesundheitsministerium.at/	30/03/2020
CHE	https://www.bag.admin.ch/bag/fr/home/krankheiten/ausbrueche-epidemien-pandemien/aktuelle-ausbrueche-epidemien/novel-cov/situation-schweiz-und-international.html#1164290554	31/03/2020

Table 2 Sources of data for fatalities and cases at regional and provincial level

Country	**Geographical resolution and variables available**	**Link**	**Update**
AUT	Provinces (NUTS3): only cases	https://info.gesundheitsministerium.at/	30/03/2020
ITA	Provinces (NUTS3): only cases. Regions (NUTS2): cases and fatalities	https://github.com/pcm-dpc/COVID-19/tree/master/dati-province	30/03/2020
ESP	Autonomous communities and cities (NUTS2): cases and fatalities	https://covid19.isciii.es/	29/03/2020
FRA	Provinces (NUTS3): cases and fatalities	https://geodes.santepubliquefrance.fr/	30/03/2020
DEU	Provinces (NUTS3): cases and fatalities	https://npgeo-corona-npgeo-de.hub.arcgis.com/datasets/917fc37a709542548cc3be077a786c17_0/data	30/03/2020