News from the front: Excess mortality and life expectancy in two major epicentres of the COVID-19 pandemic in Italy

Simone Ghislandi[1], Raya Muttarak[2], Markus Sauerberg[3], Benedetta Scotti[1]

[1] Department of Social and Political Sciences, Bocconi University, Milan
[2] Wittgenstein Centre for Demography and Global Human Capital (Univ. Vienna, IIASA, VID/ÖAW), International Institute for Applied Systems Analysis, Laxenburg
[3] Wittgenstein Centre for Demography and Global Human Capital (Univ. Vienna, IIASA, VID/ÖAW), Vienna Institute of Demography, Austrian Academy of Sciences, Vienna

Abstract

Existing studies commonly rely on national official reports to estimate the impact of COVID-19 on population health and human life. However, relying on national reports is problematic because classification and estimation of COVID-19 mortality are not consistent across countries. Likewise, delay coronavirus test results and shortage of testing kits can result in undercounting of coronavirus deaths. To overcome these problems, this study exploits all cause daily death registrations data provide by the Italian Statistical Office (ISTAT) from 1st January to 4th April 2020. This allows us to: 1) calculate excess mortality in 2020 compared to the years 2015 to 2019; and 2) estimate life expectancy on a seasonal and annual basis. We focus our analysis on Bergamo and Brescia, the two hardest hit provinces in Lombardy, northern Italy. Given the clustering nature of the epidemic, focusing on the areas with high concentration of severe illness and deaths allows us to capture the true impact of COVID-19 on mortality and life expectancy, which are likely to be underestimated in the national level data. We find that on the period 1 Jan to 4 April 2020, seasonal life expectancy in Bergamo reduced by around 8.1 and 6.5 years compared to 2019 for men and women respectively (4.5 and 3.4 years in Brescia). The drop in period life expectancy for 2020 may total up to 3 years in the case of men and 2 years in the case of women. Such a sharp decrease in life expectancy has not been experienced in modern history since the Second World War. This study shows that, in the absence of public health interventions to reduce the spread of the virus, COVID-19 has set life expectancy in Bergamo and Brescia back to the Italian life expectancy of 15 years ago.

Keywords: COVID-19, death registration, excess mortality, life expectancy, Italy, underestimation

Corresponding authors:

Simone Ghislandi (simone.ghislandi@unibocconi.it)
Benedetta Scotti (benedetta.scotti@unibocconi.it)

Introduction

According to the World Health Organization (WHO) records, on 31 December 2019 a pneumonia of unknown cause detected in Wuhan, China was first reported to the WHO Country Office in China. On the 20th of January 2020 the media around the world report the news of the human-to-human transmission of this new coronavirus, while on the 30th of January the WHO declares the outbreak a Public Health Emergency of International. While Italy, Iran and South Korea had already experienced a significant outbreak of the new coronavirus, it was only on 11 March 2020 that the COVID-19 was declared a pandemic by WHO. There are now confirmed cases of COVID-19 in every continent except Antarctica.

What is particularly worrying about the new coronavirus is its high degree of transmissibility and infectivity along with its great potential to develop into severe illness and death (Wang et al. 2020). Indeed, to minimise the impact of COVID-19 pandemic on public health, over 100 countries worldwide had impose strict restrictions on movement of people and social gatherings (Hale et al. 2020) following the drastic measures taken in China in an attempt to 'flattening the curve'. Italy, the first country outside China to experience a large-scale health impact of COVID-19, imposed a (partial) nationwide lockdown since 9 March 2020. Whilst the Italian government was praised by WHO for an implementation of such a measure (never employed in modern democratic nations since WWII), given the evidence that the virus spread under-detected in the northern part of the country already since late January 2020, strict measures might have been imposed too late. As a consequence, the outbreak has put unprecedented strain on the Italian healthcare system, resulting in an exceptionally high number of coronavirus deaths, currently the second highest in the world after the US.

As the COVID-19 outbreak spread aggressively in other countries like Spain, France, the UK, the US, country disparities in COVID-19 deaths became distinct. In particular, Italy is shown to have the highest case fatality rate while Germany, Austria and Norway are among the countries the lowest rate (Global Health 5050 2020). Accordingly, the media has constantly trying to dig out the underlying reasons behind the country differentials in COVID-19 deaths. Comparing COVID-19 deaths across countries, however, is challenging for several reasons, ranging from disparities in testing coverage, the criterial for defining a COVID-19 death to practice in post-mortem examination (García-Basteiro et al. 2020). One way to overcome the inconsistency in the differences in measurement of incidence and fatality is to consider excess mortality. Data on weekly or daily excess deaths, for instance, would allow scientists to compare the observed deaths in 2020 with the mortality patterns typically observed in previous years (Leon et al. 2020).

Exploiting daily death registrations data published by the Italian Statistical Office (ISTAT), this paper provides the first evidence-based assessment of the impact of COVID-19 on mortality and life expectancy in the first hardest hit epicentres outside China. Three main features characterise the present analysis. First, COVID-19 deaths are estimated using excess mortality measures rather than relying on the official report of the number of deaths attributed to COVID-19. This way the mortality estimation does not suffer from biases related to sampling in testing and reporting policies. Second, the analysis focusses specifically on Bergamo and Brescia, the two hardest hit provinces in Lombardy, northern Italy. The two provinces were sometimes compared to Wuhan given their position as epicentres of the outbreak, which experienced a rapid escalation in the number of infections resulting in an insufficiency of healthcare resources and intensive care units (Porcheddu et al. 2020).

The highly clustered nature of local transmission results in high concentration of severe illnesses and deaths in one area (Jia et al. n.d.). Therefore, the direct impact of COVID-19 on mortality

and average life expectancy is likely to be felt at a sub-national level rather than in a whole country. Finally, at the moment of writing, the epidemics in Bergamo and Brescia appear now to have been contained given that the epidemic curve has presented a downward trend in the first wave of infections. The results presented here are thus based mainly on hard number observed in these regions rather than on some forms of simulations with underlying assumptions. Focusing on the areas where the outbreak appears to have slowed down allows us to capture a fuller picture of the impact of the virus on mortality and life expectancy. Hence, by using complete and reliable data on the number of deaths for the first largest-scale outbreak outside China, this paper provides the first unbiased account of the mortality impact of COVID-19 and its consequence on life expectancy.

In the analysis, we compare mortality rates in 2015-2019 and 2020 across age and gender categories and provide estimations of how life expectancy changes following the COVID-19 pandemic. The motivation driving our analysis is discussed in the next session, where the main methodological choices regarding the measures and the geographical areas are discussed in depth. In the background section, we provide a short description of the policy measures implemented for tackling the epidemic. Next, the set of results are divided in two parts. First, rough excess mortality rates are presented. These values are the used to construct potential life expectancy based on different assumptions for the rest of the year 2020. In the final discussion part, we underline the most important lessons to be drawn from the evidence and conclude.

Motivations

This paper aims to provide the first evidence-based assessment of the human life consequences of the COVID-19 pandemic based on unbiased weekly death registrations data. In particular, the choice of using overall daily mortality rather than official COVID-19 figures is motivated by two major concerns. First, during the last couple of months it has become clear that there exists no uniform way of classifying COVID-19 death. There is a variation across countries in counting mortality “from” and mortality “with” COVID-19. Italy, for instance, include any persons who tested positive for the virus when they died in the official COVID-19 death toll statistics regardless of the pre-existing diseases that may have caused death (Onder et al. 2020). Likewise, whilst countries like Belgium and France after 2nd April 2020 include coronavirus deaths outside hospitals in its official daily report, countries like the UK, Italy or Spain do not. Some countries such as Belgium, the UK and the US record COVID-19 deaths also from suspected cases and do not require a positive laboratory diagnosis like in Germany and the Netherlands (Roser et al. 2020). In addition, when the epidemic worsens, the counting of fatalities becomes more difficult. People dying at home or in long-term care facilities might indeed not be tested at all simply because resource allocation prioritises emergency operations. Even within the same country, it is difficult to compare the number of cases and deaths in different regions given disparities in regional policies and testing strategies and capacities (Odone et al. 2020). Therefore, weekly death registrations data from all causes are more comparable across countries and regions (Leon et al. 2020). This is important given the need to share generalisable information in order to track the progress of the epidemics, to guide the implementation of prevention and control measures as well as to plan when to relax social distancing measures.

A second advantage of using overall mortality measure is that it allows us to capture the negative spillovers associated with the drop in supply of hospital-based services due to COVID-19. The reduced access to hospitals, emergency rooms (often closed during the outbreak), acute hospital beds and clinical activities could contribute to peaks in mortality for other medical conditions that are not directly related to the epidemic. Obviously, these deaths should be considered as death as an *indirect consequence* of COVID-19. In other words, a study that consider excess all-cause

mortality during the period of the epidemic like ours provides a complete assessment of the impact of COVID-19 on population health and human life.

Most of the existing literature has relied on mortality measures, especially case-fatality rates (CFR) (CDC COVID-19 Response Team 2020; Giangreco 2020; Khafaie and Rahim 2020; Onder et al. 2020), which are not informative for international and historical comparisons. Since case-fatality rates are calculated as the number of deaths divided by the number of confirmed cases, the absence of accurate estimation of the infection rates in a reference population makes the denominator in the CFR relies heavily on testing strategies and capacities (Odone et al. 2020). Although much of the attention from the media and WHO has been devoted to increasing the number of coronavirus testing to as many people as possible (WHO 2020), little has been said about a sampling strategy. The only possible way to obtain reliable information about infection rates is through selecting and testing a representative sample of the population (Galeotti and Surico 2020), an approach only very few countries have actually implemented. Only recently in April 2020, Iceland and Austria were the first countries to carry out a random-population screening test. This is understandable, since testing has been used as a measure of containment through isolating infected patients (together with tracing). However, testing only suspected cases implies a selection bias in a sampling approach which fails to provide unbiased estimates of infection rates in a population, no matter how many tests a country runs. The figures on infection rates based on unrepresentative sample testing thus are likely to be misestimation of the true prevalence.

Another important feature of the present study is its specific geographical focus on Bergamo and Brescia, the two hardest hit, relatively small provinces in Lombardy, northern Italy. Due to the clustering nature of the virus outbreak, an analysis at the national level would underestimate the scale of the impact of COVID-19 on mortality and life expectancy. With a total population of approximately 2.3 million in the two provinces, a large-scale excess mortality can indeed affect life expectancy of a population in these small areas. In addition, the containment measures through lockdown were introduced too late (Signorelli et al. 2020) given that the virus had already been circulated undetected in the region as early as in January 2020 (L'Eco di Bergamo 2020; Reuters 2020). This allows the epidemic to develop in its "purest" form, showing the potential impact of the virus in absence of any intervention measures to reduce its spread. Hence, when indicators like life expectancy are calculated, the reference population should naturally be the one experiencing the consequences of the epidemic. On the contrary, if the national data are considered and the outbreaks have been contained, the negative impact of the epidemic on potential years of life loss might end up being heavily underestimated.

Finally, our study extends beyond current studies on the health impact of COVID-19 which typically focus on mortality by focusing on life expectancy related measures. Whilst studying COVID-19 mortality rates are certainly important, they need to be collapsed in an index which is universal enough to provide a reliable measure of human life lost. In this regard, life expectancy is a consistent, objective measure of human development since it represents overall wellbeing of populations in a simple and non-bias form (Ghislandi et al. 2019; Sen 1998). Furthermore, reliable measures of life expectancy are available for some countries since the nineteenth century. This makes it possible to use life expectancy for historical comparisons of the human costs associated to major events

COVID-19 outbreak in Bergamo and Brescia

In the early morning of the 20th February 2020, the first severe local case of COVID-19 was diagnosed in Europe, in a small hospital in Codogno, a municipality south east of Milan.

Authorities reacted by tracing the connections of patient one, but it soon became clear that identification of a patient zero would be difficult, if not impossible. Indeed, later studies have shown that COVID-19 was present in the Lombardy region since the first half of January 2020 (Cereda et al. 2020; L'Eco di Bergamo 2020; Reuters 2020). The area around Codogno was immediately declared a "red zone" and a complete lockdown was enforced in 11 municipalities, while universities and schools in Lombardy were immediately closed. The region was about to enter what is probably the most serious public health and economic crisis after-WWII period. However, in the ten days following the presence of patient one, the nature of the epidemic was downplayed by most media and authorities, including some health experts. Testing strategies changed in due course, with only patients showing a combination of symptoms being tested. Lockdowns were gradually and slowly introduced only when it became clear that important outbreaks were on the way. On 8 March 2020 the whole Lombardy was declared "orange zones", just one day before the rest of Italy. Non-essential activities were however stopped only on 23rd March 2020.

The nature and the adequacy of the response to the COVID-19 epidemic has been discussed elsewhere and is still ongoing (Winfield 2020). In this context this brief excursus is useful in order to understand the timing and the geographic distribution of the events. Specifically, the timing is crucial here. Bergamo has been the first and hardest hit province in Lombardy. By 14 march 2020, all beds in intensive care units (ICUs) of the main hospital of the city were already filled. As it will become clear later, the wave of hospitalizations and deaths started around the very first days of March. Indeed, it has already been noticed that Bergamo is a case where the "flattening of the epidemic curve" policies were introduced too late, and hence not being very effective. A similar situation, with a few days delay, can be said for Brescia.

Data

We rely on a compendium of administrative data provided by the Italian National Institute of Statistics (Istat). We use daily death counts for all causes at the municipality level, disaggregated by sex and age class (0, 1-4, 5-9, 10-14, …, 95-99, 100+). They cover the period between 1st January and 30th April for the years 2015-2019, while for 2020 the observation period extends up to 4th April. Importantly, Istat does not collect and release data on daily death counts for all municipalities[1]. As of today, these data are available only for municipalities which meet the following three criteria: (i) have been included in the recently established National Register of Resident Population (ANPR), (ii) have recorded at least 10 deaths between 1st January and 4th April, and (iii) have recorded an increase of at least 20% in the number of deaths registered between 1st March and 4th April 2020 compared to the 2015-2019 average over the same period. In both our areas of interest, i.e. the provinces of Bergamo and Brescia, about 72% of municipalities are included in the ANPR, totaling to 175 and 149 respectively. Out of these, Istat has released daily death data relative to 2020 for 122 municipalities in the province of Bergamo and 101 municipalities in the province of Brescia.

In addition to this dataset, we rely on four further sources of information. First, we make use of annual death counts at the municipality level, disaggregated by sex and by age class (0, 1-4, 5-9, 10-14, …, 95-99, 100+). These data extend up to 2017. Second, we obtain data on resident population at the municipality level, disaggregated by sex and single-year age classes (0,1,…,100+), on January 1 of years 2015-2019. We reclassify age classes so as to match the

[1] These data have been collected and released by Istat *ad hoc* to better document the evolution of local mortality upon the Covid19 epidemics outbreak. As such, Istat has been focusing so far on a subgroup of municipalities where excess mortality in 2020 is most likely attributable to the epidemics.

ones used by Istat for daily and annual death counts (0, 1-4, 5-9, 10-14, …, 95-99, 100+), and aggregate data accordingly. Third, we use data on monthly live births, disaggregated by sex, at the municipality level. These data are available up to December 2019. Finally, we collect data on monthly death counts, disaggregated by sex, at the municipality level. These data are available up to December 2019.

Methods

We estimate the resident population by age and sex in the provinces of Bergamo and Brescia on 1st January 2020 by the following steps. First, we aggregate resident population data by year, age class and sex at the province level. Then, we estimate the female and male resident population in each age class at the provincial level on 1st January 2020 by means of linear extrapolation based on years 2015 to 2019. Likewise, we use available data on monthly live births between January 2015 and December 2019 to estimate monthly births over 2020 in the provinces of Bergamo and Brescia.

We measure excess mortality at any time *t* of 2020 as the difference between the observed and the expected number of deaths in 2020 in *t*. We defined the expected number of deaths in *t* as the average number of deaths observed in *t* over the period 2015-2019.

We calculate seasonal life expectancy between 1st January and 4th April for the years 2015-2020 (separated for men and women). For this purpose, we calculate seasonal age-specific mortality rates for each year. The numerators (i.e. the death counts) are recorded daily. Thus, aggregation over the period 1st January and 4th April at the provincial level is straightforward. The corresponding exposures (i.e., the denominators) were estimated as follows. Starting from the estimated resident population on 1st January in 2020, we counted age-specific person-days up to 4th April. Theoretically, these counts are a function of four demographic events, namely births, ageing, migration, and dying. The daily inflow of births is estimated by using monthly birth data and assuming that these births are uniformly distributed over each month. The effect of ageing, i.e. individuals might transition in- and out of a given age interval, is modeled by giving each individual the probability of 1/365 for turning one year older during the observation period. The outflows due to deaths are straightforward as age-specific death counts are known on a daily basis. We neglect migration because the population under study is quarantined. In sum, the exposed population at day *t* in age group *x* is given by the population alive at day *t-1* in age group *x* plus those who age into age group *x* at day *t* minus those who either die in age group *x* or age out of the age group *x* in day *t.* We express the obtained daily exposure values in terms of person-years by multiplying them by 1/365 (1/366 for leap years). Then, we derive age-specific mortality rates for the period 1st January to 4th April. Finally, life tables are built following standard procedures outlined by the Human Mortality Database protocol. We estimate confidence intervals for life expectancies by bootstrapping using Monte Carlo simulation methods, assuming death counts follow a binomial distribution (Andreev and Shkolnikov 2010; Chiang 1984). Further, the method proposed by Arriaga (1984) allows for decomposing changes in life expectancy into age-specific contributions. We applied this approach in order to assess, which age-groups have primary contributed to the change in seasonal life expectancy between 2019 and 2020.

When estimating life expectancy at the end of the year, harvesting (i.e. the reduction in mortality rates following peak mortality associated to shock events) should be taken into consideration. Hence, we estimate life expectancy for the entire 2020 considering two alternative scenarios, which we define as "conservative" and a "business as usual". In the conservative scenario (full harvesting), we assume the mortality rate at the end of the year will be the same as in 2017. In this approach, the main effect of COVID-19 would be to anticipate mortality to the first quarter of

the year 2020. This scenario is probably not a very likely one, but it represents a clearly defined upper bound for life expectancy in 2020. In other scenario, harvesting is set to zero and we assume that after the end of the epidemic, mortality rates will go back to the average levels recorded in the previous years. There are good reasons to assume that mortality will not reduce substantially after the end of the first wave of COVID-19 infections. For example, because of most of the resources and health services have been directed to COVID-19 patients, people with other medical conditions such as cancer were under-treated. Likewise, although most vulnerable people have been doing self-isolation and locking themselves inside, it is highly possible that they would get the virus later after restriction measures are lifted. As a matter of fact, mortality after the epidemic wave might actually ***increase***, especially among the working age population, as a consequence of the lockdowns and the likely economic recessions as has been observed, for example, in Greece (Laliotis et al. 2016). However, for this paper, we are not making assumptions on the indirect impact of COVID-19 pandemic on mortality after the epidemic ends. Therefore, if anything, years of life lost due to COVID-19 for the year 2020 may in fact be underestimated.

Based on the evolution of daily mortality observed until 4th April, we posit that the epidemics in our area of interest, i.e. the provinces of Bergamo and Brescia, should be over by 30th April. To estimate life expectancy for the entire 2020, we need to project the evolution of deaths after 4th April. We use cubic spline extrapolation to project the number of daily deaths in each age class from 4th April until 30th April. After 30th April, we project mortality until the end of the year based on two scenarios. In the "conservative" scenario, we assume that after the end of the COVID-19 epidemics, deaths in each age class will occur such that the total number of deaths in 2020 equals the total number of deaths registered in 2017[2]. If the cumulative number of deaths by the end of the epidemics in any age class exceeds the total number of deaths registered in 2017, we assume zero deaths will occur after the end of the epidemics in that age class. In the "business as usual" scenario, we assume that mortality will return to the average levels recorded in 2015-2017. These are the most recent years for which data on annual death counts are available. Thus, it is possible to derive the number of deaths occurring after 30th April as the difference between the annual death counts and deaths registered up to 30th April for the years 2015-2017. As we do not know the daily distribution of deaths after 30th, we make two assumptions. For age groups <50, we assume that deaths occur uniformly over the period 1st May – 31st December 2020. For age groups ≥50, we assume that deaths between May and December 2020 occur following the 2015-2019 average distribution of monthly deaths. We then assume that deaths distribute uniformly within each month between May and December 2020.

We then proceed with the estimation of population exposure for each day between 1st January and 31st December 2020 under both scenarios following the same procedure detailed above. Finally, we aggregate death counts and population exposure values over the entire year to derive age-specific mortality rates and life expectancies under both scenarios.

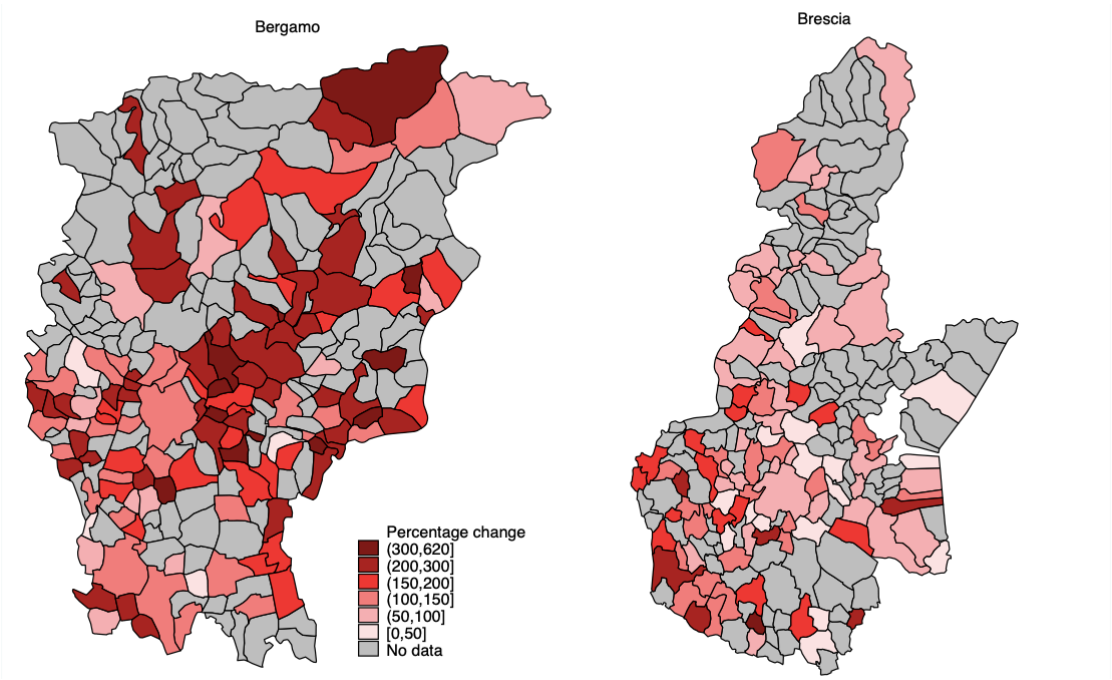
Results

Bergamo and Brescia represent the highest number of municipalities with an outlier number of deaths. Their distribution across the provinces is displayed in Figure 1. Clearly, Bergamo province shows a higher number of "red" areas. In addition, Figure A1 (in Appendix) shows the same municipalities within Lombardy, comparing the same graph with Veneto, the region that, together with Lombardy, was the epicenter of the very first clusters of COVID-19 in the last week

[2] 2017 is the last year for which data on yearly death counts disaggregated by age and sex at the municipality level is available.

of February. It is evident that the COVID-19 outbreak hit Lombardy much harder than Veneto and most of the losses are observable in the areas of Bergamo and Brescia. Indeed, in the whole country no other provinces display such a monumental upward trend in mortality.

Figure 1: Excess mortality in the provinces of Bergamo and Brescia



The maps plot the % change in total number of deaths recorded between Jan 1-Apr 4, 2020 with respect to the 2015-2019 average.

Table 1 shows the age and sex distribution of the total excess mortality by province. Compared to the average number of people that died in the same period in the five years before (2015-2019), the excess number of deaths between 1 January 2020 and 4 April 2020 for Bergamo and Brescia sums to 6549. 64% of mortality of this excess mortality is witnessed in Bergamo. COVID-19 affected slightly more men than women (56% of excess deaths) whilst at the national level 65% of COVID-19 deaths involve male mortality (Global Health 5050 2020). However, this almost gender balance in the number of deaths does not reflect an equality in the relative mortality risk between sexes. In general, men are up to 3.5 times more likely to die from COVID-19 than women. In fact, the risk of dying for men is higher than for women across all ages except for the under 50 in Bergamo, where the risk ratio is slightly below one.

Age clearly represents the most important predictor of absolute excess mortality related to COVID-19. For example, from Table 1, one can deduce that in our sample a man over 70 years is 10 times and 37 times more likely to die than a person under 70 and under 60 years of age respectively. The number of people dying for COVID-19 peaks in the 80-89 age interval. Note that the proportional increase in mortality due to COVID-19 in this age class did not spike as much as it did for the under 40 or the 40-49. In fact, for the 80-89 the expected mortality is already very high in regular times. Still, in the first quarter of 2020, in Bergamo the 80-89 mortality increased by 350% for males and 250% for females. Similar figures are observed for the 90+ and for the 70-79. Note that for the age group 90+ the gender disparity in COVID-19 mortality reverses due to a naturally much higher number of women living until over 90 years of age (Niu et al. 2020). This is consistent with the outbreaks spreading in care homes, which are

usually undercounted in official COVID-19 mortality data of most countries (Comas-Herrera et al. 2020). In Brescia, the number and the relative increase in mortality for the 70+ are lower than in Bergamo, with female and male mortality increasing in 2020 roughly by 190% and 220% respectively. However, a much higher than expected number of under 40, both males (1300%) and females (infinite increase) died in Brescia municipalities.

Table 1: Distribution of the excess mortality by province, age and sex

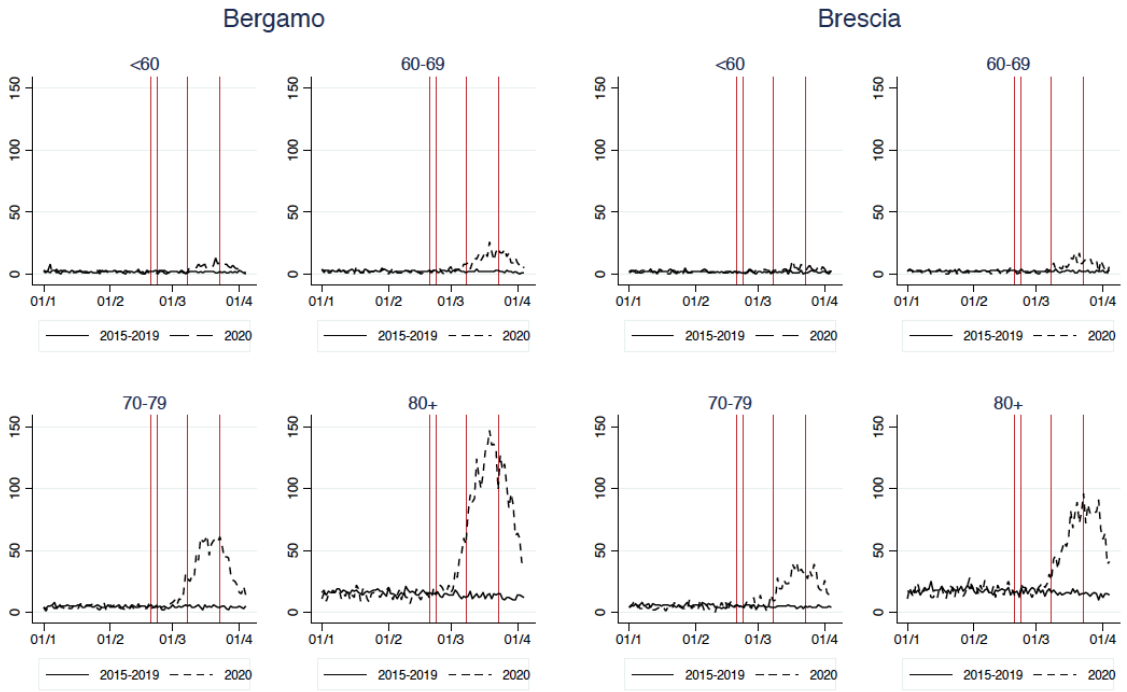
BERGAMO							
Age	**Population Males**	**Population Females**	**Excess deaths Males**	**Excess deaths Females**	**RR (M/F)**	**2020/(2015-19) Males**	**2020/(2015-19) Female**
<40	45988	44446	6	6	0.97	2.50	3.00
40-49	47729	45756	23	25	0.88	4.83	6.00
50-59	59654	58588	100	46	2.14	4.13	3.19
60-69	46749	48512	267	84	3.30	3.26	2.56
70-79	35135	40880	726	340	2.48	3.87	3.11
80-89	16383	26456	975	723	2.18	3.51	2.54
90+	2066	6619	267	607	1.41	3.32	2.64
BRESCIA							
Age	**Population Males**	**Population Females**	**Excess deaths Males**	**Excess deaths Females**	**RR (M/F)**	**2020/(2015-19) Males**	**2020/(2015-19) Female**
<40	58289	55936	25	11	2.18	13.50	-
40-49	52397	50987	21	11	1.86	2.62	2.83
50-59	62076	62175	56	30	1.87	2.51	2.25
60-69	48749	52003	123	37	3.55	1.92	1.59
70-79	39186	45880	411	160	3.01	2.44	1.86
80-89	18239	30035	516	394	2.16	2.16	1.76
90+	2406	8322	157	402	1.35	2.03	1.91

Legend
Excess deaths: Number of deaths 1 January- 4 April 2020- Average number of death 1 January- 4 April (2015-2019)
RR(M/F): Risk Ratio (Male/Female)=(excess death M/Pop M)/(excess death F/Pop F)
2020/(2015-19): Number of deaths 1 January - 4 April 2020/Average number of death 1 January – 4 April (2015-2019)

Figure 2 shows trends in daily mortality within each of the two provinces from raw data. It is evident how the epidemic curve inflates with age and how this age effect is particularly remarkable for Bergamo. It is also clear that our last day of observation is already on a downward pattern of daily mortality. Indeed, on the 4th of April the epidemic wave was already towards the end in both provinces. This allows us to rely on hard data for the most part of the infection wave, reducing our reliance on modelling in the estimation of the total number of deaths related to this first wave of infection.

From the graphs shown in Figure 2, it is also evident that the rumors of an unexpected peak of flu mortality in 2020 can be rejected. By mid-March the seasonal flu, which started in December 2019, was at the very final stages of its wave. However, during the first two months and a half of 2020, mortality across age classes did not show any significant unexpected peak. Hence, the observed trends do not support the hypothesis that the boost in mortality could be caused by the interaction of COVID-19 with seasonal flu.

Figure 2: Trends in daily mortality in the provinces of Bergamo and Brescia (1Jan-4Apr)



Timing of the events
20 February: Patient 1 in Codogno
23 February: Red zones in Codogno. Schools and Universities in affected Regions are closed.
8 March: Lombardy and Piacenza 'orange zones'
23 March: Closures of all non-essential economic activities.

The graphs plot the total daily death count between January 1 and April 4 (2020 vs 2015-2019 average), disaggregated by four age classes (males and females combined). The red vertical lines show relevant dates for the evolution of the epidemic.

Vertical lines show four relevant dates for the evolution of the epidemic. The misalignment between drastic containment measures ("orange zone", 8 March, and closure of all non-essential activities, 23 March) is striking. If one considers that the time between infection and the development of severe and critical conditions could be as long as three weeks, it is evident that the lockdowns were imposed too late.

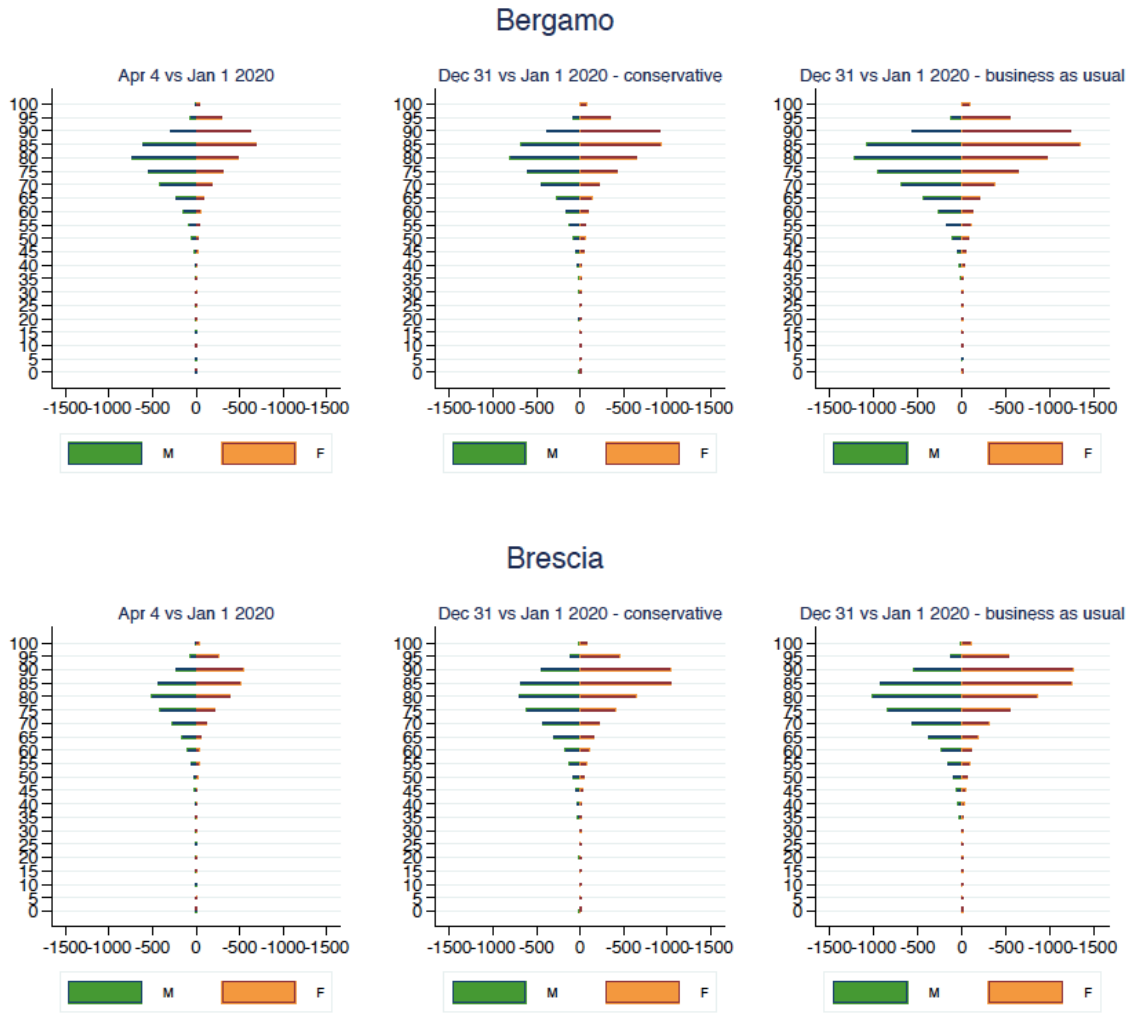
In order to measure the human life impact of this epidemic on the two provinces, we show two set of results: 1) changes in the age structure; and 2) changes in life expectancy. As stated in the methods section, when predicting mortality for the rest of the year, two scenarios can be defined. In the 'conservative' scenario, harvesting is complete and the overall mortality rates by age at the end of the year are maintained constant with respect to the previous years (unless the number of deaths by 4 April already exceeds the overall expected deaths for 2020). In the 'business as usual' scenario, the after COVID mortality drops at its regular values observed in the previous years. Harvesting is thus set to zero.

To the purpose of estimation, the cumulative number of people that died for COVID compared to the expected annual death by age is an important information. Figures A2a and A2b in the Appendix report its evolution by age classes in both provinces, for men and women respectively. It is impressive to notice that in the province of Bergamo the number of 70-79 men that died by 4

April already *exceeded the annual number of age specific deaths* in the same municipalities in 2017.

It has been argued that the impact of the COVID epidemic should be understood also by looking at the age structure of a population (Dowd et al. 2020). From the evidence provided so far, it is clear that age is the main risk factor for COVID related mortality. Figure 3 reports the changes in the age pyramid between 1 January 2020 and: a) 4 April 2020 b) 31 December 2020, in the two scenarios.

Figure 3: Changes in the age pyramid structure by sex



Overall, the age pyramid is going to be affected significantly at the top levels. The number of "missing" people among the elderlies at the end of 2020 is going to be significant, especially in the extreme scenario. Note also that elderly women seem to be the most affected by the epidemic in terms of COVID. This reflects the fact that among the 80+, 63% are females.

Life expectancy has reduced accordingly, as shown in Figure 4. The drop at the end of the observation period is significant, especially for the men in Bergamo. When seasonal life expectancy is calculated, the reduction in 2020 for males reaches 8 years. For women, the reduction is less marked, although seasonal life expectancy at 4 April 2020 is around 5 years lower than in 2019. We refer to the Discussion part for the interpretation of these figures. When life expectancy is extrapolated to the whole year, the loss in life expectancy is of course diluted over a longer period and hence reduced. In the conservative case of full harvesting, nothing

really changes with respect to the previous years. This is a very unlikely situation, but one that represents a potential upper bound. If mortality goes back to regular times ("business as usual"), life expectancy for males and females will drop by respectively 3.5 and 2.5 years in Bergamo, and by 2.3 and 1.5 years in Brescia.

Figure 4: Life expectancy at birth estimates

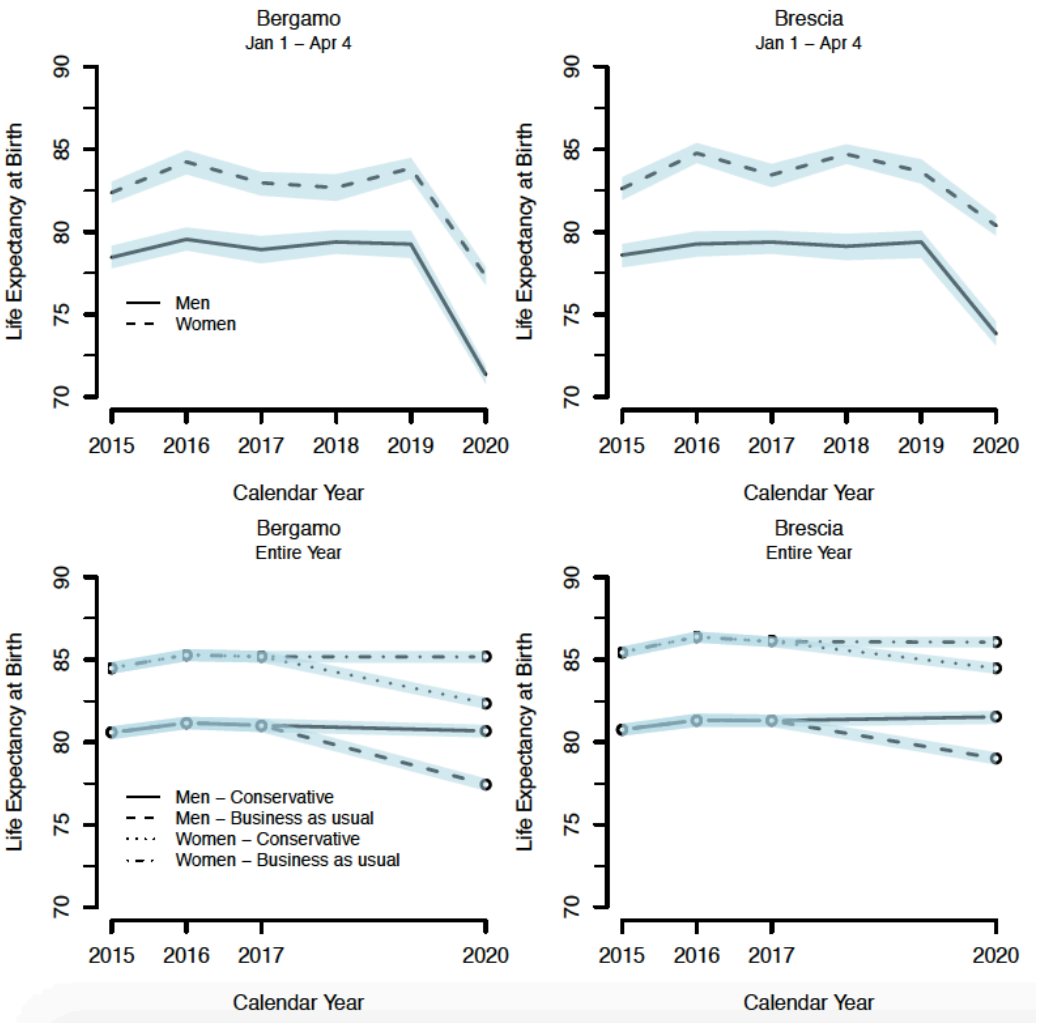
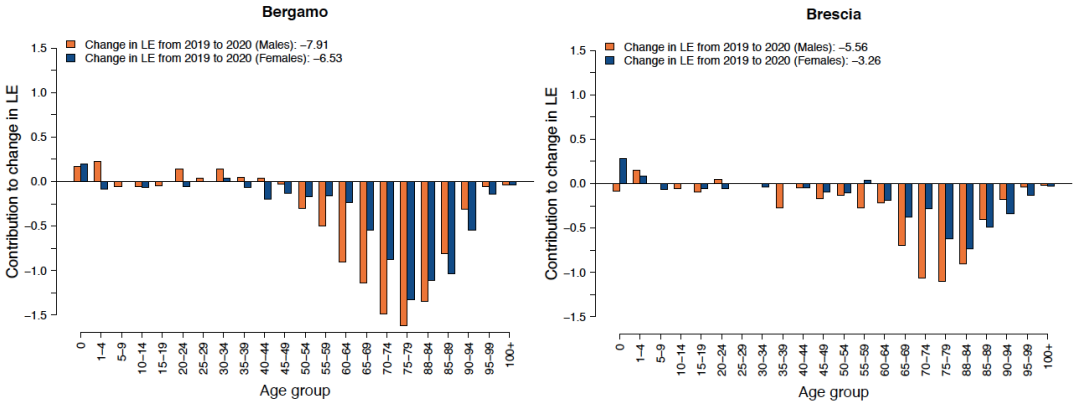


Figure 5 shows the decomposition associated to the loss of life expectancy by April 2020, by age and sex. Clearly the major role is played by older populations. The role of mortality in 60-79 years old men is particularly evident. At the same time, the 90+ women in Bergamo, despite their much older age, are still responsible of more than half a year loss in life expectancy.

Figure 5: Decomposition of seasonal life expectancy by age



Discussions

We are able to show excess mortality directly and indirectly linked to COVID-19 without any bias in testing coverage and classification of cause of deaths. By focusing on the two hardest hit areas in Lombardy and Italy, Bergamo and Brescia, we are able to capture the full impact of COVID-19 on population health and human life.

The first important piece of evidence refers to the duration of the epidemic wave. By 4 April 2020 the daily mortality in both Bergamo and Brescia was approaching the regular values observed in 2019. Hence, the time between the first sign of an epidemic wave (first three days of March) and the (possible) return to normality is likely to be between 6 and 8 weeks, with the peak happening around two weeks from the start. It is likely that the lockdown policy might have contributed to the reduction of mortality in the late stage of the epidemic wave. However, containment policies were not implemented during the initial spread of the infections, when the virus circulated undetected in Lombardy; hence they played no role in reducing mortality in the first part of the epidemic wave.

Regarding the demographic characteristics, two empirical regularities emerge clearly. First, the age gradient in mortality is steep and age represents the most evident risk factor for COVID-19 mortality. In our sample, men aged over 70 are 10 times and 37 times more likely to die than their counterparts aged under 70 and under 60, respectively. Second, gender disparities in the percentages distribution of deaths are not as large (56% deaths involve male subjects) as reported for the whole nation of Italy (65%). However, this figure is misleading: the risk of dying for a male is between 1.5 to 3.5 higher than that for a female, depending on the age class and the province considered. Men are thus ***more*** at risk of mortality, but since COVID-19 affects the elderly the most and women are overrepresented among those aged 70+, absolute figures show a high number of deaths also among the females. This age gradient of mortality affects the age pyramid as well, with elderly women being associated to the sharpest reduction in their number by the end of 2020.

In terms of life expectancy, we have shown two main set of results. First, when focusing on the period 1 Jan to 4 April 2020, life expectancy in Bergamo reduced by around 8.1 and 6.5 years compared to 2019 for men and women respectively (4.5 and 3.4 years in Brescia). One way to understand these figures is to imagine that they represent what would happen if the 1 Jan to 4 April period were to be stretched for the entire year, so that the epidemic would last around four months over twelve. Fortunately, this is not the case because the epidemic wave had a limited

duration. Nevertheless, these figures give us an idea of the power of the epidemic wave that crushed the provinces.

Return to normality in Bergamo and Brescia is not going to be straightforward. The near future is hard to predict. For this reason, when calculating life expectancy at the end of the year we focus on very simple assumptions, i.e. focusing only on the impact of the first epidemic wave. Only harvesting effect is considered as possible consequences of COVID-19 epidemic on mortality. Two scenarios of full harvesting (conservative) and no harvesting (business as usual) are introduced. By construction, the conservative scenario does not show significant changes at the end of the year. This is because, although the number of deaths by 4 April exceeded the expected yearly number of deaths in some age class, it did so only by a small number.

On the other hand, the no harvesting scenario implies a loss in life expectancy of around 4.5 and 2.8 years for males and females, respectively in Bergamo (2.2 and 2.4, respectively in Brescia). Although the conservative scenario represents a logical extreme upper bound to any estimation of life expectancy, the business as usual is a more realistic one.

But how can we interpret these results? One good feature of the life expectancy indicator is that it provides a historical perspective over the human consequence of relevant shocks (Ghislandi et al. 2019). Looking at Figure A3 in the Appendix, it appears evident that such a sharp decrease in life expectancy has not been experienced in modern history since the Second World War. In terms of the pandemic impact on years of life lost, we are observing a large-scale loss in life expectancy second only to the Spanish flu in 1918, where, according to Figure A3, life expectancy for Italy dropped by more than 15 years. For the USA, Noymer and Garenne (2000) estimate reduction in life expectancy of 11.8 years. The differences between the COVID-19 pandemic and the 1918 flu pandemic are very distinct. First of all, the Spanish flue typically killed the young, which of course, all else equal, reduces life expectancy much more. Second, the outbreak of the Spanish flu occurred at the same time as the First World War when no public health and preventive measures were implemented. Finally, the medical technology and knowledge was much less advanced, with penicillin and other crucial pharmaceutical products not yet being available. It thus can be said that the years of life lost in Bergamo and Brescia, although not at the level achieved after the Spanish flu, are of a substantial scale given a much more advanced medical technology and the absence of war in the present day.

Note that our analysis focuses on a local context. Given that the COVID-19 outbreaks are geographically concentrated, looking at a country level life expectancy is misleading and underestimates the actual impact of the pandemic. Indeed, from our results, it is fair to conclude that this is a lifetime unprecedented shock for most of the residents in Bergamo and Brescia.

The main limitation of the present study is related to the limited coverage of the death registrations data used, both geographically and temporally. All the information provided here has been retrieved from a number of the most affected municipalities in Bergamo and Brescia. So, the estimation represents an attempt to describe and measure the human cost of COVID-19 among the populations that were hit the hardest in a country. This study does not represent the experiences of Italy and Lombardy region but also by choice, we intend to investigate the impact of COVID-19 on human life in the areas where little containment interventions were put in place. From the point of view of the timing, we have enough information to actually picture the entire evolution of the epidemic wave. However, when it comes to life expectancy for the year 2020, we need to rely on harvesting assumptions and we exclude the possible negative impact of the present and future lockdown policies and global economic recessions on health and human life.

In this sense, this paper only focuses on the direct impact of the pandemic and, if anything, may underestimate the years of life lost due to the COVID-19.

Nevertheless, the timing of this paper is also its strength. Indeed, the present analysis represents the first attempt to provide an evidence-based assessment of the human life impact of a frontline COVID-19 outbreak. In doing so, we raise a warning for the scientific and public health community around the world: the epidemic, when undetected and not contained, can easily get out of control and cause relevant social and human consequences. In defining the lockdown strategies governments should balance the costs of limiting social and economic activities with the human costs of possible epidemic waves. Vulnerable populations should be protected with a particular attention and constant monitoring of the epidemic progression. This should be given a priority in order not to get caught by surprise by devastating epidemic waves like the ones experienced in Bergamo and Brescia.

Conclusion

Human life in itself is a solid indicator of human development because it represents health and wellbeing of a population, which is one key priority of the Sustainable Development Goals. So many years of life lost due to COVID-19 pandemic is worrying since it sets the affected areas back to their level of human development 20 years ago. As shown in the case of Bergamo and Brescia, the cost of human life from a delay in public interventions to reduce the virus transmission is disturbingly high. Amidst countries' preparation to ease the lockdown restrictions and social distancing measures, it is important to keep in mind the conceivable risk of viral reintroduction and a potential direct and indirect cost it can pose on human life. Well-planned government measures to prevent a second wave of infections along with public collaboration in keeping physical distancing and practicing proper hygiene until a vaccine for the novel coronavirus is available are key to achieve a balance between health and economic protection.

References

Andreev, E. M., & Shkolnikov, V. (2010). *Spreadsheet for calculation of confidence limits for any life table or healthy-life table quantity* (No. MPIDR Technical Report 2010-005). Rostock: Max Planck Institute for Demographic Research. https://www.demogr.mpg.de/en/publications_databases_6118/publications_1904/mpidr_technical_reports/spreadsheet_for_calculation_of_confidence_limits_for_any_life_table_or_healthy_life_table_quantity_3853. Accessed 28 April 2020

Arriaga, E. E. (1984). Measuring and Explaining the Change in Life Expectancies. *Demography*, *21*(1), 83–96. https://doi.org/10.2307/2061029

CDC COVID-19 Response Team. (2020). Severe Outcomes Among Patients with Coronavirus Disease 2019 (COVID-19) - United States, February 12-March 16, 2020. *MMWR. Morbidity and mortality weekly report*, *69*(12), 343–346. https://doi.org/10.15585/mmwr.mm6912e2

Cereda, D., Tirani, M., Rovida, F., Demicheli, V., Ajelli, Poletti, P., et al. (2020). The early phase of the COVID-19 outbreak in Lombardy, Italy. *arXiv:2003.09320 [q-bio]*. http://arxiv.org/abs/2003.09320. Accessed 27 April 2020

Chiang, C. L. (1984). *The Life Table and Its Applications* (Original edition.). Malabar, FL: Krieger Publishing Company.

Comas-Herrera, A., Zalakaín, J., Litwin, C., Hsu, A. T., Lane, N., & Fernández, J.-L. (2020). *Mortality associated with COVID-19 outbreaks in care homes: early international evidence*. International Long Term Care Policy Network, CPEC-LSE. https://ltccovid.org/wp-content/uploads/2020/04/Mortality-associated-with-COVID-26-April-2.pdf

Dowd, J. B., Andriano, L., Brazel, D. M., Rotondi, V., Block, P., Ding, X., et al. (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
Galeotti, A., & Surico, P. (2020). *Why testing a representative sample of the population must be done now*. https://voxeu.org/article/why-testing-representative-sample-population-must-be-done-now
García-Basteiro, A. L., Chaccour, C., Guinovart, C., Llupià, A., Brew, J., Trilla, A., & Plasencia, A. (2020). Monitoring the COVID-19 epidemic in the context of widespread local transmission. *The Lancet Respiratory Medicine*, *0*(0). https://doi.org/10.1016/S2213-2600(20)30162-4
Ghislandi, S., Sanderson, W. C., & Scherbov, S. (2019). A Simple Measure of Human Development: The Human Life Indicator. *Population and Development Review*, *45*(1), 219–233. https://doi.org/10.1111/padr.12205
Giangreco, G. (2020). Case fatality rate analysis of Italian COVID-19 outbreak. *Journal of Medical Virology*, *n/a*(n/a). https://doi.org/10.1002/jmv.25894
Global Health 5050. (2020). *COVID-19 sex-disaggregated data tracker*. Globale Health 5050. https://globalhealth5050.org/covid19/
Jia, J., Hu, X., Yang, F., Song, X., Dong, L., Zhang, J., et al. (n.d.). Epidemiological Characteristics on the Clustering Nature of COVID-19 in Qingdao City, 2020: A Descriptive Analysis. *Disaster Medicine and Public Health Preparedness*, 1–5. https://doi.org/10.1017/dmp.2020.59
Khafaie, M. A., & Rahim, F. (2020). Cross-Country Comparison of Case Fatality Rates of COVID-19/SARS-COV-2. *Osong Public Health and Research Perspectives*, *11*(2), 74–80. https://doi.org/10.24171/j.phrp.2020.11.2.03
Laliotis, I., Ioannidis, J. P. A., & Stavropoulou, C. (2016). Total and cause-specific mortality before and after the onset of the Greek economic crisis: an interrupted time-series analysis. *The Lancet Public Health*, *1*(2), e56–e65. https://doi.org/10.1016/S2468-2667(16)30018-4
L'Eco di Bergamo. (2020, March 31). «Il virus già qui da gennaio, in incognito» I dati di uno studio: primo caso a Curno. *L'Eco di Bergamo*. Bergamo, Italy. https://www.ecodibergamo.it/stories/bergamo-citta/il-virus-gia-qui-da-gennaio-in-incognitoi-dati-di-uno-studio-primo-caso-a-cu_1347299_11/. Accessed 26 April 2020
Leon, D. A., Shkolnikov, V. M., Smeeth, L., Magnus, P., Pechholdová, M., & Jarvis, C. I. (2020). COVID-19: a need for real-time monitoring of weekly excess deaths. *The Lancet*, *0*(0). https://doi.org/10.1016/S0140-6736(20)30933-8
Niu, S., Tian, S., Lou, J., Kang, X., Zhang, L., Lian, H., & Zhang, J. (2020). Clinical characteristics of older patients infected with COVID-19: A descriptive study. *Archives of Gerontology and Geriatrics*, *89*, 104058. https://doi.org/10.1016/j.archger.2020.104058
Noymer, A., & Garenne, M. (2000). The 1918 Influenza Epidemic's Effects on Sex Differentials in Mortality in the United States. *Population and development review*, *26*(3), 565–581.
Odone, A., Delmonte, D., Scognamiglio, T., & Signorelli, C. (2020). COVID-19 deaths in Lombardy, Italy: data in context. *The Lancet Public Health*, *0*(0). https://doi.org/10.1016/S2468-2667(20)30099-2
Onder, G., Rezza, G., & Brusaferro, S. (2020). Case-Fatality Rate and Characteristics of Patients Dying in Relation to COVID-19 in Italy. *JAMA*. https://doi.org/10.1001/jama.2020.4683
Porcheddu, R., Serra, C., Kelvin, D., Kelvin, N., & Rubino, S. (2020). Similarity in Case Fatality Rates (CFR) of COVID-19/SARS-COV-2 in Italy and China. *The Journal of Infection in Developing Countries*, *14*(02), 125–128. https://doi.org/10.3855/jidc.12600
Reuters. (2020, April 24). Italy's coronavirus epidemic began in January, study shows. *Reuters*. Rome. https://www.reuters.com/article/us-health-coronavirus-italy-study/italys-coronavirus-epidemic-began-in-january-study-shows-idUSKCN2262B1. Accessed 26 April 2020
Roser, M., Ritchie, H., Ortiz-Ospina, E., & Hasell, J. (2020). *Coronavirus Disease (COVID-19)* (Our World in Data). https://ourworldindata.org/coronavirus
Sen, A. (1998). Mortality as an indicator of economic success and failure. *The Economic Journal*, *108*(446), 1–25. https://doi.org/10.1111/1468-0297.00270

Signorelli, C., Scognamiglio, T., & Odone, A. (2020). COVID-19 in Italy: impact of containment measures and prevalence estimates of infection in the general population. *Acta Bio Medica Atenei Parmensis*, *91*(3-S), 175–179. https://doi.org/10.23750/abm.v91i3-S.9511

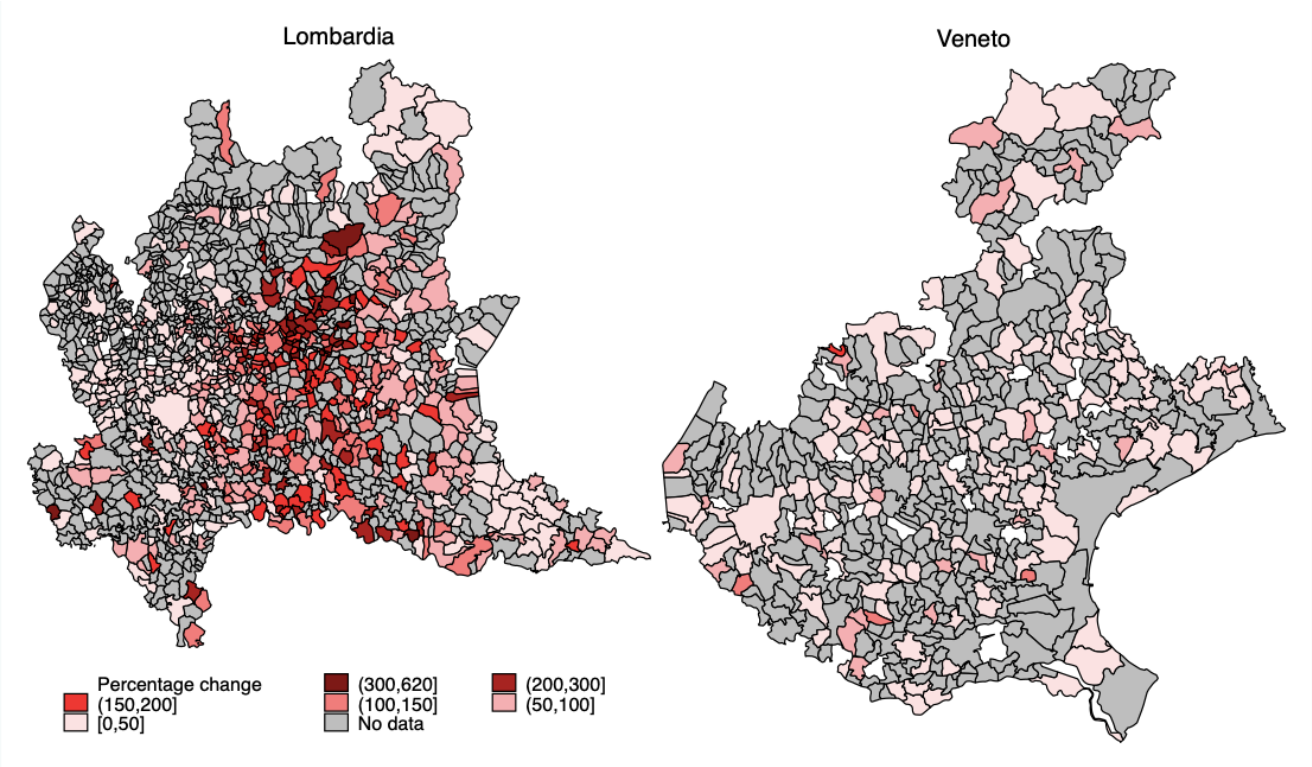
Wang, L., Wang, Y., Ye, D., & Liu, Q. (2020). Review of the 2019 novel coronavirus (SARS-CoV-2) based on current evidence. *International Journal of Antimicrobial Agents*, 105948. https://doi.org/10.1016/j.ijantimicag.2020.105948

WHO. (2020). *Laboratory testing strategy recommendations for COVID-19*. Geneva: World Health Organization. https://apps.who.int/iris/bitstream/handle/10665/331509/WHO-COVID-19-lab_testing-2020.1-eng.pdf

Winfield, N. (2020, April 26). Many failures combined to unleash death on Italy's Lombardy. *AP News*. https://apnews.com/de2794327607a3a67ed551f0b6b71404. Accessed 26 April 2020

Appendix

Figure A1: Excess mortality in Lombardy and Veneto



The maps plot the % change in total number of deaths recorded between Jan 1-Apr 4, 2020 with respect to the 2015-2019 average.

Figure A2a: Cumulative deaths - Men

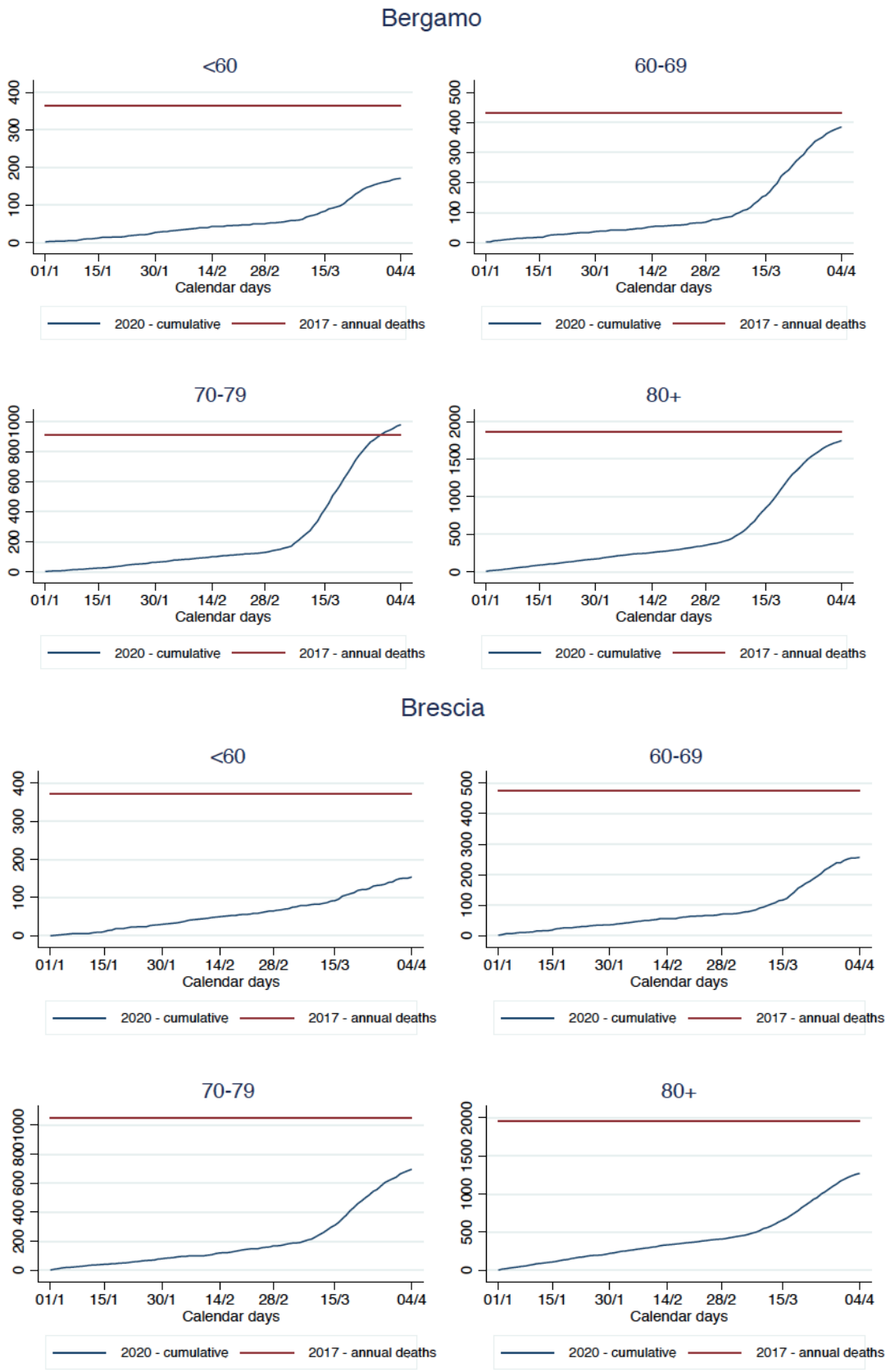


Figure A2b: Cumulative deaths - Women

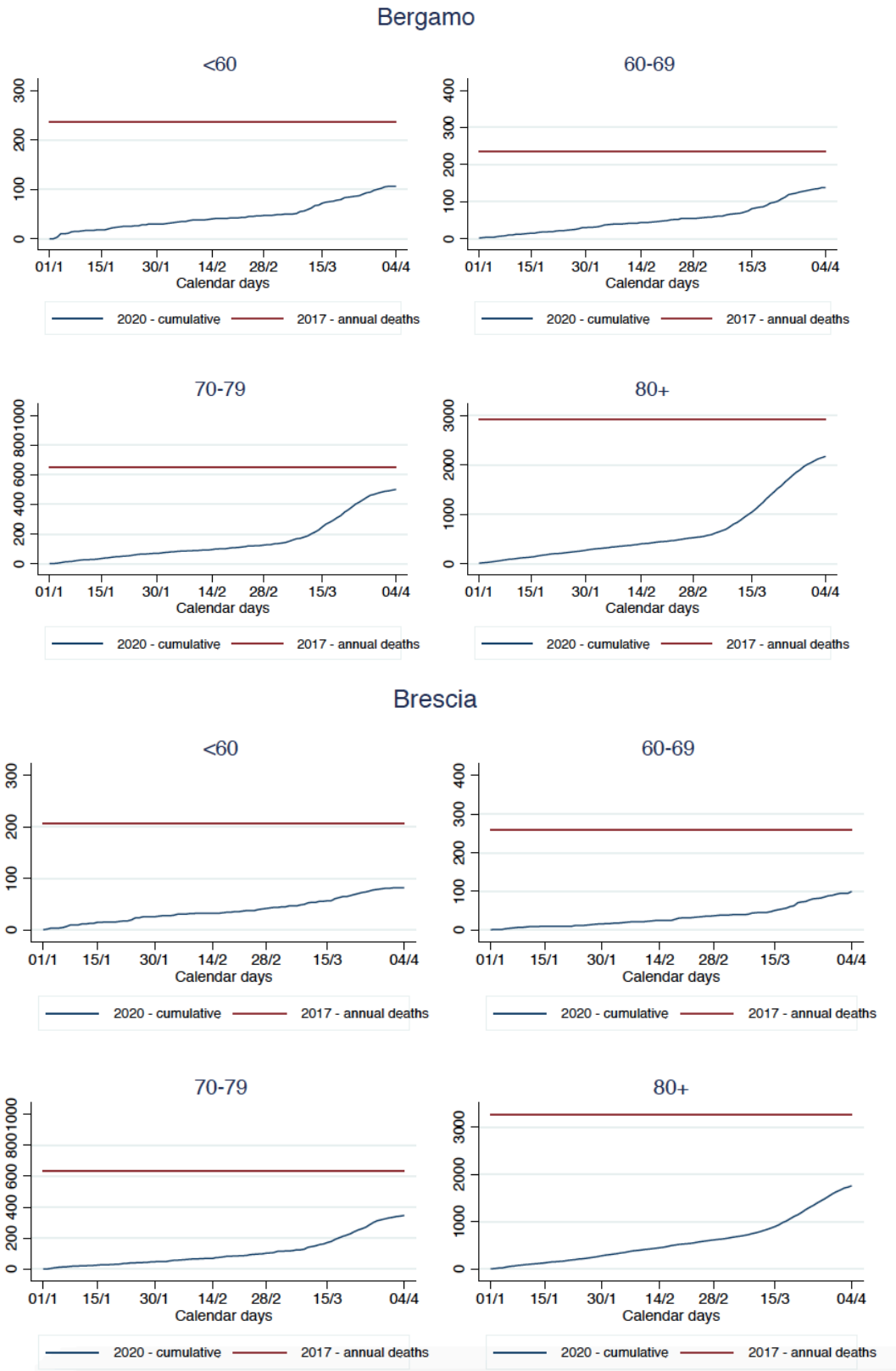
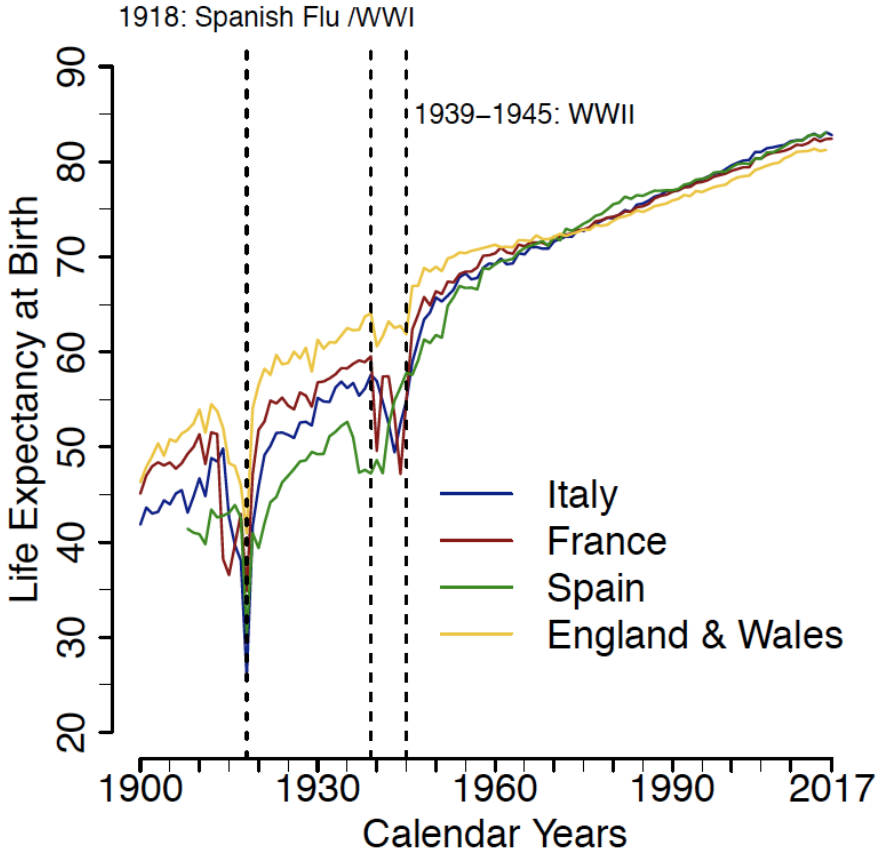


Figure A3: Trends in life expectancy at birth



Own elaborations from the Human Mortality Database, UC Berkeley, and Max Planck Institute for Demographic Research 2017.