News from the front: Estimation of excess mortality and life expectancy in the major epicenters 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, Via Roentgen 1, 20136, Milan, Italy
[2] Wittgenstein Centre for Demography and Global Human Capital (Univ. Vienna, IIASA, VID/ÖAW), International Institute for Applied Systems Analysis, Schlossplatz 1, A-2361 Laxenburg, Austria
[3] Wittgenstein Centre for Demography and Global Human Capital (Univ. Vienna, IIASA, VID/ÖAW), Vienna Institute of Demography, Austrian Academy of Sciences, Vordere Zollamtsstrasse 3, 1030 Vienna, Austria

Abstract

Existing studies commonly rely on national official reports to estimate the impact of COVID-19 on human life. This is problematic because classification, estimation and reporting of COVID-19 mortality are not consistent across countries. To overcome these problems, this study exploits all cause daily death registrations data provided by the Italian Statistical Office (ISTAT) from January 1 to April 15, 2020. This allows us to: 1) calculate excess mortality in 2020 compared to the average of the years 2015 to 2019; and 2) estimate life expectancy on a seasonal and annual basis. Focusing on the five most severely hit provinces in Italy (Bergamo, Brescia, Cremona, Lodi and Piacenza) and Lombardy region, this analysis captures the impact of COVID-19 on mortality and life expectancy, which are likely to be *underestimated* when only national level data are considered. We find that seasonal life expectancy in 2020 in the five provinces reduced by 5.1 to 7.8 and 3.2 to 5.8 years for men and women, respectively. For annual life expectancy for the year 2020, in a scenario with no harvesting effect i.e. mortality rates resume to an average level of the years 2015-2019 after the end of the first epidemic wave, the years of life lost is equivalent to 2 to 3.5 years for men and 1.1. to 2.5 years for women in the five provinces. This represents the largest decline in life expectancy after the 1918 influenza pandemic and the Second World War.

Corresponding authors:

* Raya Muttarak (muttarak@iiasa.ac.at)
* Benedetta Scotti (benedetta.scotti@unibocconi.it)

Keywords

COVID-19, death registration, excess mortality, Italy, life expectancy

Author Contributions

S.G. and B.S. designed research; B.S. and M.S. performed research and analyzed data; S.G., R.M., M.S. and R.M. wrote the paper

Introduction

Italy is the first country outside China and in Europe to experience the largest-scale impact of COVID-19 pandemic on human life. As the COVID-19 outbreak spread aggressively to other European countries, country disparities in COVID-19 deaths became distinct. In particular, Italy, Belgium and the UK are shown to have the highest case fatality rates while Germany, Austria and Norway are among the European countries with the lowest rate (1). Comparing COVID-19 deaths across countries, however, is challenging. Most of the existing literature has relied on case-fatality rates (CFR) as a measure of mortality (2–5), 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 reliant on testing strategies and capacities (6).

Another issue with comparing COVID-19 deaths across countries is the absence of a uniform way in classifying, recording and reporting COVID-19 deaths (7). There is a variation across countries in counting mortality "from" and mortality "with" COVID-19. Italy, for instance, includes 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 (5). Likewise, while countries like Belgium and France (after April 2) 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 unlike in Germany and the Netherlands (8). 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 prioritizes emergency operations (9, 10). The overall number of COVID-19 deaths thus is most likely undercounted.

The COVID-19 pandemic can also have an indirect impact on human life. As medical resources are primarily allocated for the treatment of COVID-19 patients, non-COVID medical emergencies and cares for chronic illness are likely to be undermined (11). Depleted hospital capacity in providing routine medical treatments coupled with reluctance to visit healthcare facilities due to a fear of the spread of COVID-19 infection within the hospital could contribute to peaks in mortality for other medical conditions that are not directly related to the epidemic (12). Arguably, these deaths should also be considered as death indirectly related to COVID-19.

To overcome these problems, this study uses daily death registrations data for the period January 1 to April 15 2020 published by the Italian Statistical Office (ISTAT). This allows us to calculate excess all-cause mortality during the period of the epidemic and corresponding annual life expectancy without suffering from selection bias related to sampling in testing and reporting policies. This is the first study to provide a complete assessment of the impact of COVID-19 on human life.

In particular, this study focuses on specific geographical areas in Italy which are the most severely affected by the COVID-19 pandemic: four provinces in Lombardy (Bergamo, Lodi, Cremona, Brescia) and one province in Emilia Romagna (Piacenza). Indeed, Gatto et al. (13) in modelling the spread of COVID-19 in Italy highlight the importance of considering a spatial nature of the progress of the wave of infections. The selected provinces experienced the highest number of excess deaths in Italy in the period observed compared to the average level of mortality in the years 2015-2019. The highly clustered nature of local transmission results in high concentration of severe illnesses and deaths in one area (14). 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. In fact, our results show that even in Lombardy – the hardest hit region (15) – the years of life lost due to COVID-19 appear to be significantly lower than in Bergamo, the province which contributed to one-third of the total excess mortality in the Lombardy region. A spatial granularity thus needs to be warranted in order to assess a full scale of human life impact.

Based on daily death registrations data, the present study compares mortality rates in 2015-2019 and 2020 across age and gender categories and provides estimations of the changes in life expectancy following the COVID-19 pandemic. While measures such as mortality rates are no doubt useful, 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 significantly related to the overall wellbeing of population and could thus provide a simple, objective and immediate measure of the human casualties associated to unprecedented shocks such as the COVID-19 pandemic (16–18). 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

Institutional and geographical context of the hardest hit areas

In the early morning of February 21, 2020, the first severe case of local transmission was diagnosed in Europe, in a small hospital in Codogno, a municipality in the province of Lodi, southeast of Milan (19). Initially, authorities reacted by tracing the connections of *patient one* but finally failed to identify a *patient zero*. 11 municipalities within the province of Lodi were put under strict measures to contain the spread of the disease and declared the quarantine "Red Zone" as early as February 24, 2020. Meanwhile, another cluster of COVID-19 outbreak emerged in Alzano Lombardo and Nembro, two municipalities in the province of Bergamo, north east of Milan. Amidst the rapid rise in the number of detected cases, especially in the municipalities surrounding these two epicenters, on March 8 the Italian government imposed a (partial) nationwide lockdown from March 9, followed by a total lockdown of all non-essential activities on March 23, 2020 (20). Whilst the Italian government was praised by WHO for an implementation of such a drastic measure (not being employed in modern democratic nations since World War II), the virus has already spread undetected in the northern part of the country since late January 2020 (21–23). The containment measures thus might have been imposed a little too late (24). As a consequence, the outbreak has put unprecedented burden on the Italian healthcare system, resulting in an exceptionally high number of coronavirus deaths, currently the third highest in the world after the US and the UK (1).

Geographically, Lodi and Codogno are close to other two provinces included in our sample, Cremona and Piacenza (see Fig. S1. and S2. in SI Appendix for a geographical location of the provinces being studied). The epidemic wave involving these provinces is thus considered as part of the Lodi-Codogno cluster. Bergamo and Brescia are located north-east of Milan and, despite the first COVID-19 severe cases being detected just one day after *patient one* in Lodi, experienced a week delay in the arrival of the epidemic wave (20). Lombardy, on the other hand, is the most populated region in Italy and the one with the highest Gross Domestic Product (GDP). Overall, it accounts for one sixth of the Italian population and one fifth of its GDP. Lombardy is relevant for our analysis because it is by far the hardest hit region by the COVID-19 pandemic in Italy, accounting for almost 50% of the human casualties of the entire country (6). Indeed, with the exception of Piacenza (locating in Emilia Romagna region), a bordering province with Lombardy region, all other four provinces included in the analysis are in Lombardy. In the following, we will thus also present statistics also for the region of Lombardy.

Excess mortality

The geographical distribution of excess deaths in the observed period across Italy (Fig. 1) matches with the distribution of confirmed cases (which comprise the deceased, recovered individuals and active cases) provided by the Italian Civil Protection Department, the official surveillance data on COVID-19 (25). This geographical pattern warrants that the excess mortality observed in our data represents mortality directly and indirectly related to COVID-19.

Bergamo, Brescia, Cremona, Lodi and Piacenza represent the top five provinces in terms of increase in mortality in Italy. The largest percentages increase in excess mortality is observed in Bergamo and Cremona. Clearly, Bergamo province shows the highest number of "red" areas – municipalities where excess mortality increased by over 300%.

When considering the number of deaths, the hardest hit provinces are also the most populated ones: Bergamo and Brescia.

FIG. 1. ABOUT HERE

Compared to the average number of people that died in the same period in the previous five years (2015-2019), the excess number of deaths (over 40) between January 1 and April 15, 2020 sums to 5686 in Bergamo, 3671 in Brescia, 1756 in Cremona, 827 in Lodi and 906 in Piacenza. For the entire region of Lombardy the excess number of deaths are approximately 19,000 (See Table S1 in SI Appendix). Bergamo and Brescia have the highest number of excess deaths. In fact, the death toll in Bergamo alone accounts for almost one third of the excess mortality in Lombardy, underlining the clustering nature of the epidemic. Age clearly represents a risk factor for excess mortality in a similar manner as age-gradient in COVID-19 CFR found in Italy and elsewhere (26). For instance, among the excess deaths observed in Bergamo, mortality rate is much higher among older men aged 70 years and over (18 times higher than those aged below 70). A similar ratio is found in other provinces.

When simply considering the distribution of excess mortality without adjusting for population size in each age-gender category, we observe slightly more excess mortality in men than in women (56% of excess deaths involve male subjects). This figure is not far from the death toll reported in official COVID-19 fatality data, where 63% of COVID-19 deaths are males (27). However, when considering the mortality risk ratio between sexes, the excess mortality for males is consistently higher than that of females across all age groups and provinces. For example, between January 1 and April 15, men 70-79 years old in Bergamo are 3.75 times more likely to die than women of the same age. It is evident that male sex is a risk factor associated with severe illness and deaths from COVID-19 in Italy (28) and elsewhere (29, 30). Our data on excess mortality represent similar sex disparities in COVID-19 related deaths.

FIG. 2. ABOUT HERE

Fig. 2 shows trends in daily mortality for the selected provinces and Lombardy. Plotting mortality distribution by age groups allows us to fully capture the progression of the epidemic in terms of mortality. It is evident how the epidemic curve inflates with age across all provinces. It is also clear that by April 15 the daily mortality in all selected provinces was approaching the pre-pandemic values. Hence, the wavelength of the epidemic in these provinces was between 6 and 8 weeks, with the peak happening around two weeks after the onset of the outbreak.

Vertical lines show four relevant dates for the evolution of the epidemic. With the case of *patient one* being first identified in Codogno located in the province of Lodi, the authorities quickly locked down 11 municipalities in the area on February 24, 2020. The containment measures through lockdown were only implemented in other provinces from March 8 on. Although the earlier lockdown enabled Lodi to flatten the curve more effectively than other severely hit provinces (Fig. 2), the province still experienced a notable increase in excess mortality. Considering that the incubation period – the time between exposure and symptom onset – can be up to 24 days (31, 32), it is evident that the lockdown were imposed too late in these provinces. Apart from political reasons preventing the authorities to introduce the lockdown earlier in the provinces where the number of cases had been risen rapidly like Bergamo, the recent evidence shows that COVID-19 has already been circulated undetected in northern Italy since January 2020 (21–23). Our study thus proxies the impact of COVID-19 outbreak in the absence of containment interventions.

Results

In order to measure the human life impact of this epidemic on the selected five provinces and in Lombardy, we present estimated life expectancy for the year 2020 compared to that observed in the period 2015 to 2019 (Fig.3 and Fig.4). To obtain annual life expectancy, we need to predict mortality for the rest of the year. Here two scenarios are defined (see Materials and Methods for details about the approach). In the 'conservative' scenario, harvesting (i.e. the potential reduction in mortality rates following a temporary peak in mortality associated to negative shocks) is complete and the overall mortality rates by age at the end of the year are kept constant with respect to the previous five years. This approach implies by construction that everyone who died during COVID-19 would have died anyway during the year. In this setting, the main effect of COVID-19 is thus to anticipate mortality to the first quarter of the year 2020. Although this scenario is probably unlikely, it represents a clearly defined upper bound for life expectancy in 2020.

In the 'business as usual' scenario, on the contrary, 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 five years. There are good reasons to think that mortality will not reduce substantially after the end of the first wave of COVID-19 infections. For example, since 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, premature mortality after the first epidemic wave might actually ***increase*** as a consequence of the lockdowns and the following economic crises, as has been observed, for example, in Greece (33) and elsewhere in Europe after the 2008 global economic crisis (34, 35). Furthermore, in some provinces like Bergamo, the number of deaths of males in some age groups by April 15 has already *exceeded the annual number of age-sex specific deaths* in the same municipalities in 2017 (Fig. S3A and S3B in SI Appendix). Therefore, if anything, years of life lost due to COVID-19 for the year 2020 may in fact be underestimated.

FIG. 3. ABOUT HERE

When considering seasonal life expectancy, it is evident that life expectancy has reduced significantly for men and women in all provinces (Fig.3). When compared to the average life expectancy of the 2015-2019 period, the years of life lost for men range from 5.1 in Brescia to 7.8 in Bergamo and for women from 3.2 in Piacenza to the 5.8 in Bergamo. The higher years of life lost for men is due to sex differentials in COVID-19 mortality risk, as consistently found in both the official case fatality data (27, 36) and our death registration data. Indeed, when decomposing the loss in life expectancy to identify which age groups mainly contribute to a reduction in life expectancy (Fig. S4 in SI Appendix), it is clear that the older populations, especially men aged 60-79 years play a major role.

FIG. 4. ABOUT HERE

When life expectancy is extrapolated for the whole year (Fig.4), the loss in life expectancy is naturally diluted over a longer period. The drop in life expectancy due to COVID-19-related excess mortality thus is less sharp than that observed in seasonal life expectancy. In the conservative scenario of full harvesting (Fig. 4), not much changes are expected with respect to the previous years. This rather unlikely situation represents a potential upper bound of life

expectancy. Note however that in some cases life expectancy in 2020 drops even in this conservative scenario (for example, for men in Bergamo). This is because some age groups experienced already more deaths than in the past years. For these groups, the mortality for 2020 is thus likely to be higher than what experienced in the past, even if full harvesting is assumed (see Fig. S3A and S3B in SI Appendix).

If mortality goes back to regular times ("business as usual" scenario) (Fig. 4), in the most severely hit province of Bergamo, life expectancy will drop by 3.55 and 2.48 years for men and women, respectively when compared to the years 2015-2017. In slightly less affected provinces of Brescia, Cremona, Lodi and Piacenza, the years of life lost for men are between 1.88 in Piacenza and 2.53 in Lodi and for women between 1.17 in Piacenza to 1.62 in Lodi. As expected, Lombardy shows a smaller reduction in life expectancy of 1.15. and 0.82 years for males and females, respectively.

Discussion

Avoiding the inconsistencies in classification of cause of deaths and in testing practices, this study estimates excess mortality directly and indirectly linked to COVID-19. By focusing on the five most severely affected areas in Italy and the most affected region of Lombardy, we are able to capture the full impact of COVID-19 on population health and human life.

Two empirical regularities clearly emerge regarding demographic differentials. First, the age gradient in excess mortality is steep and age represents the most evident risk factor for COVID-19 mortality (26, 37). In Lombardy, men and women aged over 70 are 25 times and 66 times more likely to die than their counterparts aged under 70 and under 60, respectively. These patterns are replicated in all five provinces. Therefore, an area with a high proportion of older populations (e.g. 17% of the population was over 70 in Lombardy region in 2019) would suffer a higher burden of COVID-19 mortality (38). Second, the risk of dying for men within each province is between 1.5 to 3.5 higher than that for women, depending on the age class and the province considered. Higher mortality rates consequently translate to a larger reduction in life expectancy for men than for women.

In terms of life expectancy, we have shown two main set of results. First, when focusing on the period Jan 1 to April 15, 2020, a reduction in seasonal life expectancy, compared with the average 2015-2019, can be as high as 7.8 years for men and 5.8 years for women in Bergamo, respectively. One way to interpret these figures is to consider mortality and corresponding life expectancy as seasonal (39).

Moving forward, we can gauge how the COVID-19 pandemic would affect annual life expectancy for the year 2020. In the case of anticipated mortality caused by heat waves or cold spells, we typically observe harvesting effect where the frailest population subgroups die while strong individuals survive and later contribute to the number of deaths which are lower than expected (40). In our case, although we do provide a scenario with harvesting effect in the calculation of annual life expectancy for the year 2020, this scenario is unlikely. In the five provinces considered, the no harvesting scenario implies a loss in life expectancy of around 2 to 3.5 years and 1.1 to 2.5 years for males and females, respectively. In fact, there remain many uncertainties regarding the longer-term effects of the pandemic on health conditions; for instance, regarding patients recovering from COVID-19 with important co-morbidities and mental health issues (41, 42) or pregnant women (43, 44). Indirect physical and mental health consequences following changing socioeconomic conditions might also affect mortality patterns in 2020 (45). Therefore, if anything, our estimates of annual life expectancy for 2020 in the business-as-usual scenario with no harvesting effect might represent a scenario which is not too pessimistic.

The COVID-19 pandemic is often compared with the 1918 influenza pandemic (commonly known as "Spanish flu") – the most severe pandemic in the last century. The COVID-19 pandemic brought about a large-scale loss in life expectancy from pandemics, second only to the 1918 Spanish flu, where life expectancy for Italy dropped by more than

15 years (See Fig. S5 in SI Appendix) and for the USA, by 11.8 years as estimated by Noymer and Garenne (46). Fortunately, the direct impact of COVID-19 on human life seems to be smaller than the 1918 Spanish flu for many reasons. First of all, since the Spanish flu typically killed younger people, it reduces life expectancy much more than the COVID-19 which is particularly fatal for older populations. 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 adequately implemented. The coexistence of war and pandemic made the Spanish flu exceptionally lethal. Finally, during the period of the 1918 influenza pandemic, the medical knowledge and technology were 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 observed in the most severely affected areas of Italy, 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.

The timing and the explicit focus on a local context can be considered as the strength of this analysis. 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. The present analysis represents the first attempt to provide an evidence-based assessment of the human life impact from 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. 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 the most severely affected provinces in northern Italy

At the same time, providing evidence right at the end of the first wave of COVID-19 epidemic in It comes at a price. that scenarios are needed regarding mortality for the remaining months of 2020. When it comes to annual life expectancy, 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. Hence, this paper focuses on the *direct impact* of the first wave of the pandemic and, if anything, may underestimate the years of life lost due to the COVID-19.

Human life in itself is a solid indicator of human development because it represents health and wellbeing of a population (16, 17), which is one key priority of the Sustainable Development Goals. So many years of life lost due to COVID-19 pandemic is worrisome since it sets the affected areas back to their level of human development 15 years ago. As shown in the cases of Bergamo, Brescia, Cremona, Lodi, and Piacenza, 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.

Materials and Methods

Estimate procedures for excess mortality

We first estimate the resident population by age and sex for each of the selected five provinces on January 1, 2020 in 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 January 1, 2020 by means of linear extrapolation based on the 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 each province.

Excess mortality is measured at any time *t* in 2020 as the difference between the observed and the expected number of deaths in 2020 in *t*. The expected number of deaths in *t* is defined as the average number of deaths observed in *t* over the period 2015-2019:

$$ {}_{n}D_{x}^{excess}(t_{2020}) = {}_{n}D_{x}^{observed}(t_{2020}) - {}_{n}D_{x}^{expected}(t_{2020}) $$

with

$$ {}_{n}D_{x}^{expected}(t_{2020}) = \frac{1}{k}\sum_{i=1}^{k} {}_{n}D_{x}\,(t_{2020-i}) $$

where the number of deaths in the age interval *x* to *x+n* at time *t* are defined as ${}_{n}D_{x}(t)$.

Estimate procedures for life expectancy

(A) Seasonal life expectancy

We calculate seasonal life expectancy for the period January 1 and April 15 for the years 2015-2020 (separated for men and women). To do so, seasonal age-specific mortality rates are calculated for each year. We aggregate the daily death counts (the numerators) over the period January 1 and April 15 at the provincial level (and at the regional level for Lombardy).

The corresponding exposures (i.e., the denominators) were estimated as follows. Starting from the estimated resident population on January 1 in 2020, we counted age-specific person-days up to April 15. 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 throughout the month. The effect of ageing, i.e. individuals might be in 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 assume no migration because the population being studied is under quarantine.

Formally, 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*:

$$ {}_{n}E_{x}\,(t) = {}_{n}P_{x}(t-1) + {}_{n}Age_in_{x}(t) - {}_{n}D_{x}(t) - {}_{n}Age_out_{x}(t) $$

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 January 1 to April 15 by diving ${}_nD_x(t)$ by ${}_nE_x(t)$. Finally, life tables are built following standard procedures outlined by the Human Mortality Database protocol (47). Life expectancy at birth (seasonal or annual) can be expressed as:

$$e_0(t) = \frac{1}{l_0}\sum_{x=0}^{\infty} {}_nL_x\,(t)$$

with ${}_nL_x$ being the number of person-years lived between age *x* and *x* + *n* in a life table for the period t and l_0 the life table radix (i.e., 100,000 newborns).

We estimate confidence intervals for life expectancies by bootstrapping using Monte Carlo simulation methods, assuming death counts follow a binomial distribution (48, 49).

(B) Annual life expectancy

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 (40). Here life expectancy for the entire 2020 is estimated considering two alternative scenarios: 1) conservative scenario; and 2) business-as-usual scenario.

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. In the business-as-usual 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.

Based on the evolution of daily mortality observed until April 15, we posit that the epidemics in our area of interest should be over by April 30. To estimate life expectancy for the entire 2020, we need to project the evolution of deaths after April 15. Cubic spline extrapolation is employed to project the number of daily deaths in each age class from April 15 until April 30. After April 30, we project mortality until the end of the year based on the 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[1]. 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, 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 April 30 as the difference between the annual death counts and deaths registered up to April 30 for the years 2015-2017. As we do not know the daily distribution of deaths after April 30, we make two assumptions based on empirical evidence (39). For age groups <50, we assume that deaths occur uniformly over the period May 1 –December 31, 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.

[1] 2017 is the last year for which data on yearly death counts disaggregated by age and sex at the municipality level is available.

We then proceed with the estimation of population exposure for each day between January 1 and December 31, 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.

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 five-year age class. They cover the period between January 1 and April 30 for the years 2015-2019, while for 2020 the observation period extends up to April 15. Importantly, Istat does not collect and release data on daily death counts for all Italian municipalities. These data are available for the 6,866 municipalities, out of a total of 7,904 municipalities in Italy, which have been included in the recently established National Register of Resident Population (ANPR). In our geographical areas of interest, the number of municipalities included in the ANPR ranges from 91% in the province of Piacenza to 99% in the province of Cremona. For Lombardy, the ANPR coverage is 96%. As of today, Istat has released daily death counts for all the municipalities included in the ANPR up to March 31, 2020. For a large subset of them (around 70% in our areas of interest), daily death counts are also available up to April 15, 2020. We summarize coverage of data on daily death counts in Table S2 in the Supplemental Material.

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 five-year age group. 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 on January 1 of for the years 2015-2019. We reclassify age classes to five-year age group so as to match the ones used by Istat for daily and annual death counts 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.

Acknowledgments

We would like to thank the Italian National Statistical Institutes (ISTAT) for providing invaluable assistance with data access and enquiries.

References

1. E. Dong, H. Du, L. Gardner, An interactive web-based dashboard to track COVID-19 in real time. *The Lancet Infectious Diseases* **20**, 533–534 (2020).

2. CDC COVID-19 Response Team, Severe Outcomes Among Patients with Coronavirus Disease 2019 (COVID-19) - United States, February 12-March 16, 2020. *MMWR Morb. Mortal. Wkly. Rep.* **69**, 343–346 (2020).

3. G. Giangreco, Case fatality rate analysis of Italian COVID-19 outbreak. *Journal of Medical Virology* **n/a** (2020).

4. M. A. Khafaie, F. Rahim, Cross-Country Comparison of Case Fatality Rates of COVID-19/SARS-COV-2. *Osong Public Health Res Perspect* **11**, 74–80 (2020).

5. G. Onder, G. Rezza, S. Brusaferro, Case-Fatality Rate and Characteristics of Patients Dying in Relation to COVID-19 in Italy. *JAMA* (2020) https:/doi.org/10.1001/jama.2020.4683 (March 25, 2020).

6. A. Odone, D. Delmonte, T. Scognamiglio, C. Signorelli, COVID-19 deaths in Lombardy, Italy: data in context. *The Lancet Public Health* **0** (2020).

7. A. L. García-Basteiro, *et al.*, Monitoring the COVID-19 epidemic in the context of widespread local transmission. *The Lancet Respiratory Medicine* **0** (2020).

8. M. Roser, H. Ritchie, E. Ortiz-Ospina, J. Hasell, "Coronavirus Disease (COVID-19)" (2020).

9. G. Iacobucci, Covid-19: Care home deaths in England and Wales double in four weeks. *BMJ* **369** (2020).

10. A. O'Dowd, Covid-19: Care home deaths in England and Wales rise sharply. *BMJ* **369** (2020).

11. The Lancet Oncology, Safeguarding cancer care in a post-COVID-19 world. *The Lancet Oncology* **21**, 603 (2020).

12. J. Appleby, What is happening to non-covid deaths? *BMJ* **369** (2020).

13. M. Gatto, *et al.*, Spread and dynamics of the COVID-19 epidemic in Italy: Effects of emergency containment measures. *PNAS* (2020) https:/doi.org/10.1073/pnas.2004978117 (April 26, 2020).

14. J. Jia, *et al.*, Epidemiological Characteristics on the Clustering Nature of COVID-19 in Qingdao City, 2020: A Descriptive Analysis. *Disaster Med Public Health Prep*, 1–5 (2020).

15. G. Sebastiani, M. Massa, E. Riboli, Covid-19 epidemic in Italy: evolution, projections and impact of government measures. *Eur. J. Epidemiol.* **35**, 341–345 (2020).

16. A. Sen, Mortality as an indicator of economic success and failure. *The Economic Journal* **108**, 1–25 (1998).

17. S. Ghislandi, W. C. Sanderson, S. Scherbov, A Simple Measure of Human Development: The Human Life Indicator. *Population and Development Review* **45**, 219–233 (2019).

18. J. M. Aburto, F. Villavicencio, U. Basellini, S. Kjærgaard, J. W. Vaupel, Dynamics of life expectancy and life span equality. *PNAS* **117**, 5250–5259 (2020).

19. M. Paterlini, On the front lines of coronavirus: the Italian response to covid-19. *BMJ* **368** (2020).

20. M. M. Galizzi, S. Ghislandi, Bergamo's response to the coronavirus pandemic. *Cambridge Core blog* (2020) (May 7, 2020).

21. D. Cereda, *et al.*, The early phase of the COVID-19 outbreak in Lombardy, Italy. *arXiv:2003.09320 [q-bio]* (2020) (April 27, 2020).

22. L'Eco di Bergamo, «Il virus già qui da gennaio, in incognito» I dati di uno studio: primo caso a Curno. *L'Eco di Bergamo* (2020) (April 26, 2020).

23. Reuters, Italy's coronavirus epidemic began in January, study shows. *Reuters* (2020) (April 26, 2020).

24. C. Signorelli, T. Scognamiglio, A. Odone, COVID-19 in Italy: impact of containment measures and prevalence estimates of infection in the general population. *Acta Bio Medica Atenei Parmensis* **91**, 175–179 (2020).

25. Italian Civil Protection Department, M. Morettini, A. Sbrollini, I. Marcantoni, L. Burattini, COVID-19 in Italy: Dataset of the Italian Civil Protection Department. *Data in Brief* **30**, 105526 (2020).

26. R. E. Jordan, P. Adab, K. K. Cheng, Covid-19: risk factors for severe disease and death. *BMJ* **368** (2020).

27. Global Health 5050, “COVID-19 sex-disaggregated data tracker” (Globale Health 5050, 2020).

28. A. Remuzzi, G. Remuzzi, COVID-19 and Italy: what next? *The Lancet* **395**, 1225–1228 (2020).

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

30. P. Mo, *et al.*, Clinical characteristics of refractory COVID-19 pneumonia in Wuhan, China. *Clin Infect Dis* (2020) https:/doi.org/10.1093/cid/ciaa270 (May 3, 2020).

31. Y. Wang, Y. Wang, Y. Chen, Q. Qin, Unique epidemiological and clinical features of the emerging 2019 novel coronavirus pneumonia (COVID-19) implicate special control measures. *Journal of Medical Virology* **92**, 568–576 (2020).

32. Y. Yan, *et al.*, The First 75 Days of Novel Coronavirus (SARS-CoV-2) Outbreak: Recent Advances, Prevention, and Treatment. *Int J Environ Res Public Health* **17** (2020).

33. I. Laliotis, J. P. A. Ioannidis, C. Stavropoulou, 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**, e56–e65 (2016).

34. D. Parmar, C. Stavropoulou, J. P. A. Ioannidis, Health outcomes during the 2008 financial crisis in Europe: systematic literature review. *BMJ* **354**, i4588 (2016).

35. D. Stuckler, A. Reeves, R. Loopstra, M. Karanikolos, M. McKee, Austerity and health: the impact in the UK and Europe. *Eur J Public Health* **27**, 18–21 (2017).

36. L. A. Walter, A. J. McGregor, Sex- and Gender-specific Observations and Implications for COVID-19. *Western Journal of Emergency Medicine: Integrating Emergency Care with Population Health* **21** (2020).

37. F. Zhou, *et al.*, Clinical course and risk factors for mortality of adult inpatients with COVID-19 in Wuhan, China: a retrospective cohort study. *The Lancet* **0** (2020).

38. J. B. Dowd, *et al.*, Demographic science aids in understanding the spread and fatality rates of COVID-19. *PNAS* (2020) https:/doi.org/10.1073/pnas.2004911117 (April 27, 2020).

39. T. Ho, A. Noymer, Summertime, and the livin’ is easy: Winter and summer pseudoseasonal life expectancy in the United States. *Demographic Research* **37**, 1445–1476 (2017).

40. M. Luy, P. Di Giulio, V. Di Lego, P. Lazarevič, M. Sauerberg, Life Expectancy: Frequently Used, but Hardly Understood. *Gerontology* (2019) https:/doi.org/10.1159/000500955.

41. F. Dutheil, L. Mondillon, V. Navel, PTSD as the second tsunami of the SARS-Cov-2 pandemic. *Psychol Med*, 1–2 (2020).

42. J. Torales, M. O’Higgins, J. M. Castaldelli-Maia, A. Ventriglio, The outbreak of COVID-19 coronavirus and its impact on global mental health. *Int J Soc Psychiatry*, 20764020915212 (2020).

43. E. Mullins, D. Evans, R. M. Viner, P. O’Brien, E. Morris, Coronavirus in pregnancy and delivery: rapid review. *Ultrasound Obstet Gynecol* **55**, 586–592 (2020).

44. M. Zaigham, O. Andersson, Maternal and perinatal outcomes with COVID-19: A systematic review of 108 pregnancies. *Acta Obstet Gynecol Scand* (2020) https:/doi.org/10.1111/aogs.13867.

45. M. McKee, D. Stuckler, If the world fails to protect the economy, COVID-19 will damage health not just now but also in the future. *Nature Medicine*, 1–3 (2020).

46. A. Noymer, M. Garenne, The 1918 Influenza Epidemic's Effects on Sex Differentials in Mortality in the United States. *Popul Dev Rev* **26**, 565–581 (2000).

47. J. R. Wilmoth, K. Andreev, D. Jdanov, A. Glei, T. Riffe, "Methods Protocol for the Human Mortality Database" (The Human Mortality Database, 2019).

48. E. M. Andreev, V. Shkolnikov, "Spreadsheet for calculation of confidence limits for any life table or healthy-life table quantity" (Max Planck Institute for Demographic Research, 2010) (April 28, 2020).

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

Figures and Tables

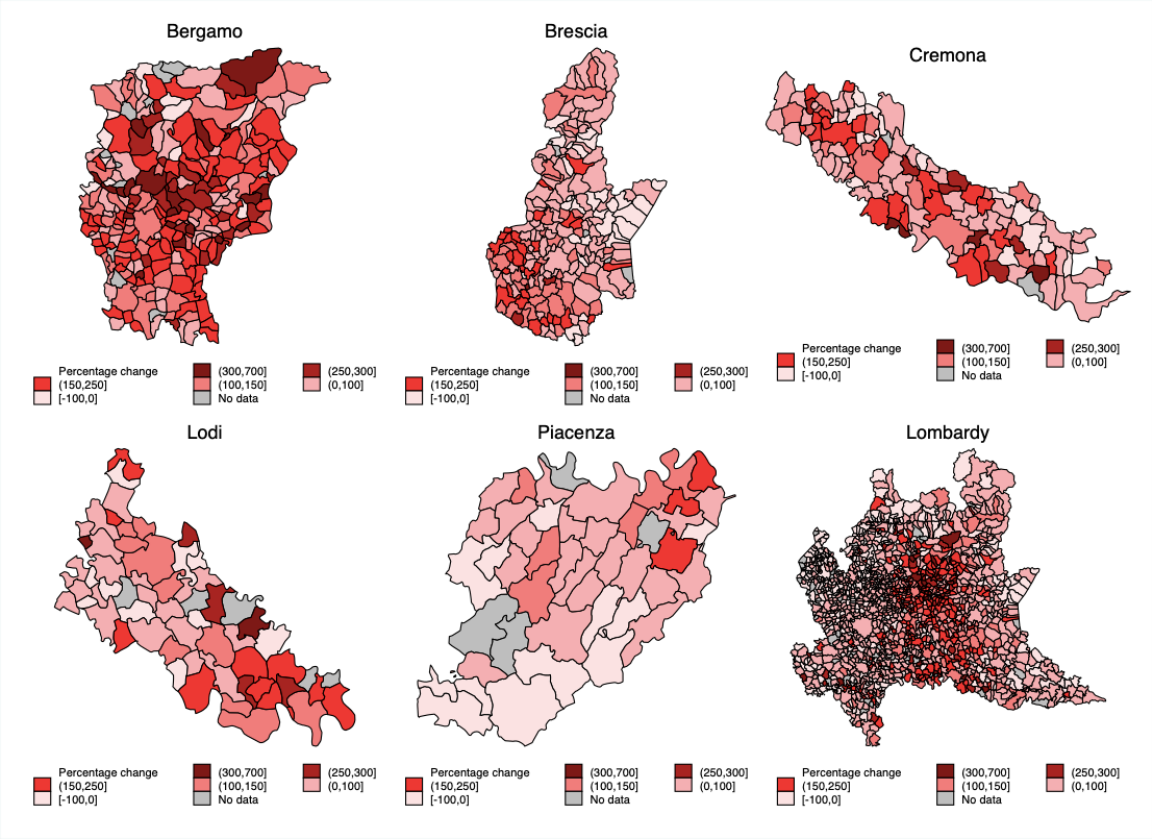
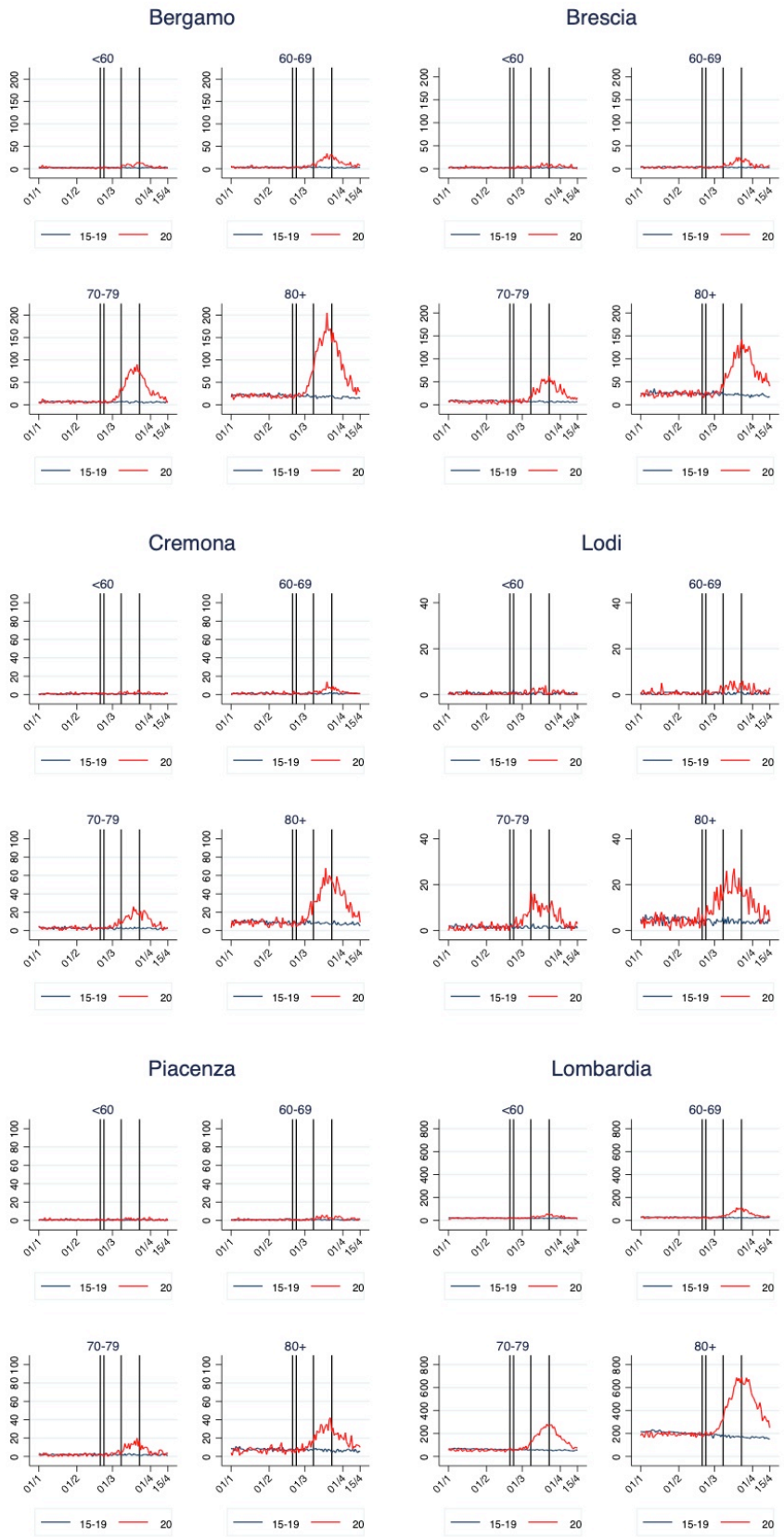


Fig. 1. Excess mortality in Italy by municipality in (A) Bergamo, (B) Brescia, (C) Cremona, (D) Lodi, (E) Piacenza and (F) Lombardy. The maps plot the percentages change in total number of deaths recorded between Jan 1 - Apr 15, 2020 with respect to the 2015-2019 average.



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.

>

Fig. 2. Trends in daily mortality by age in selected five provinces and in Lombardy: (A) Bergamo, (B) Brescia, (C) Cremona, (D) Lodi, (E)Piacenza and (F) Lombardy. The graphs plot the total daily death count between January 1 and April 15 (2020 vs 2015-2019 average), males and females combined. The vertical lines show relevant dates for the evolution of the epidemic. Vertical lines indicate relevant days: February 20 =Patient 1 found in Codogno; February 23= Red zones in Codogno. Schools and Universities in affected Regions are closed; March 8= Lombardy and Piacenza 'orange zones; March 23= Closures of all non-essential economic activities.

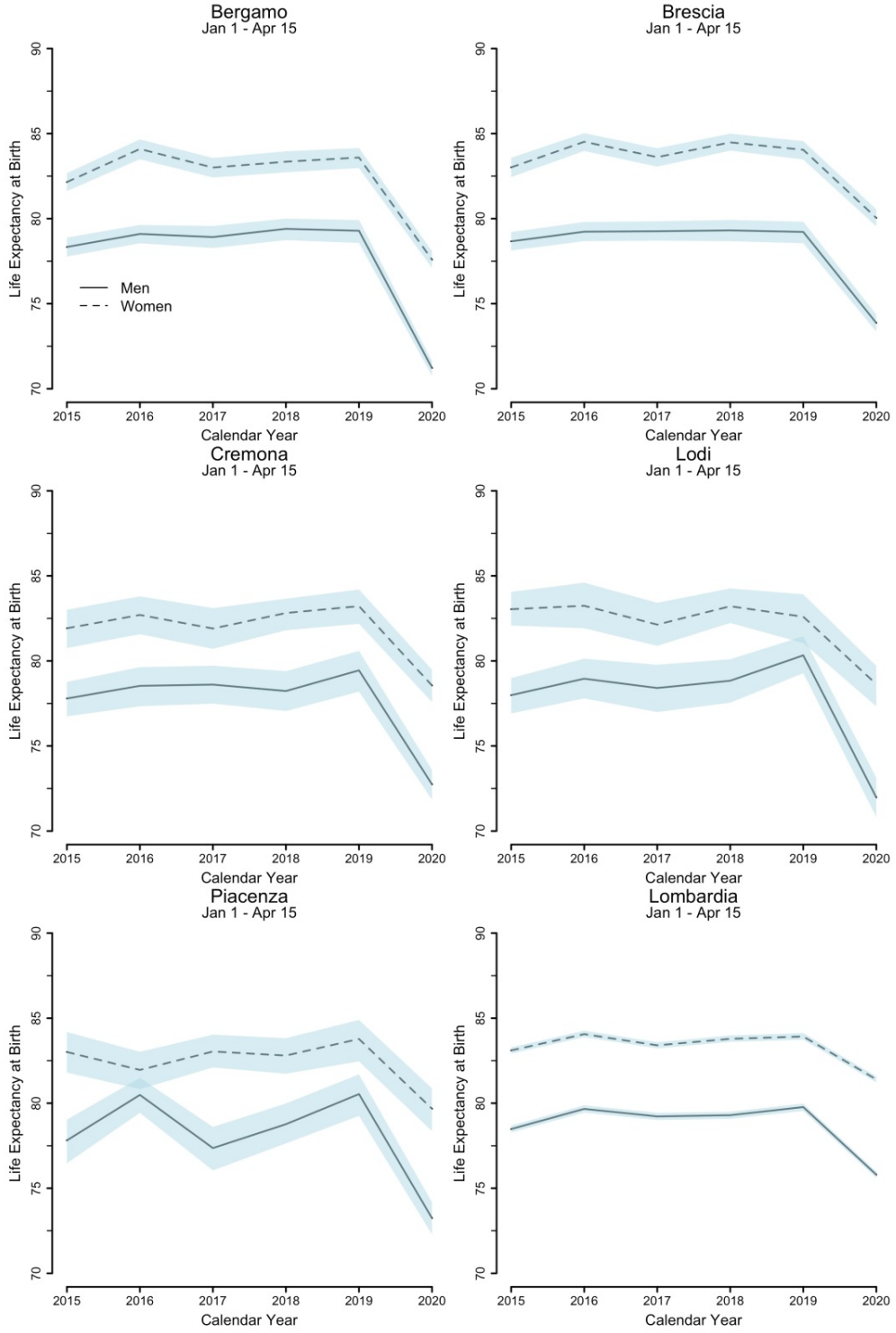


Fig. 3. Estimates of seasonal (January 1- April 15) life expectancy at birth by sex in selected provinces and in Lombardy. Confidence intervals (95%) for life expectancies are estimated by bootstrapping using Monte Carlo simulation methods, assuming death counts follow a binomial distribution.

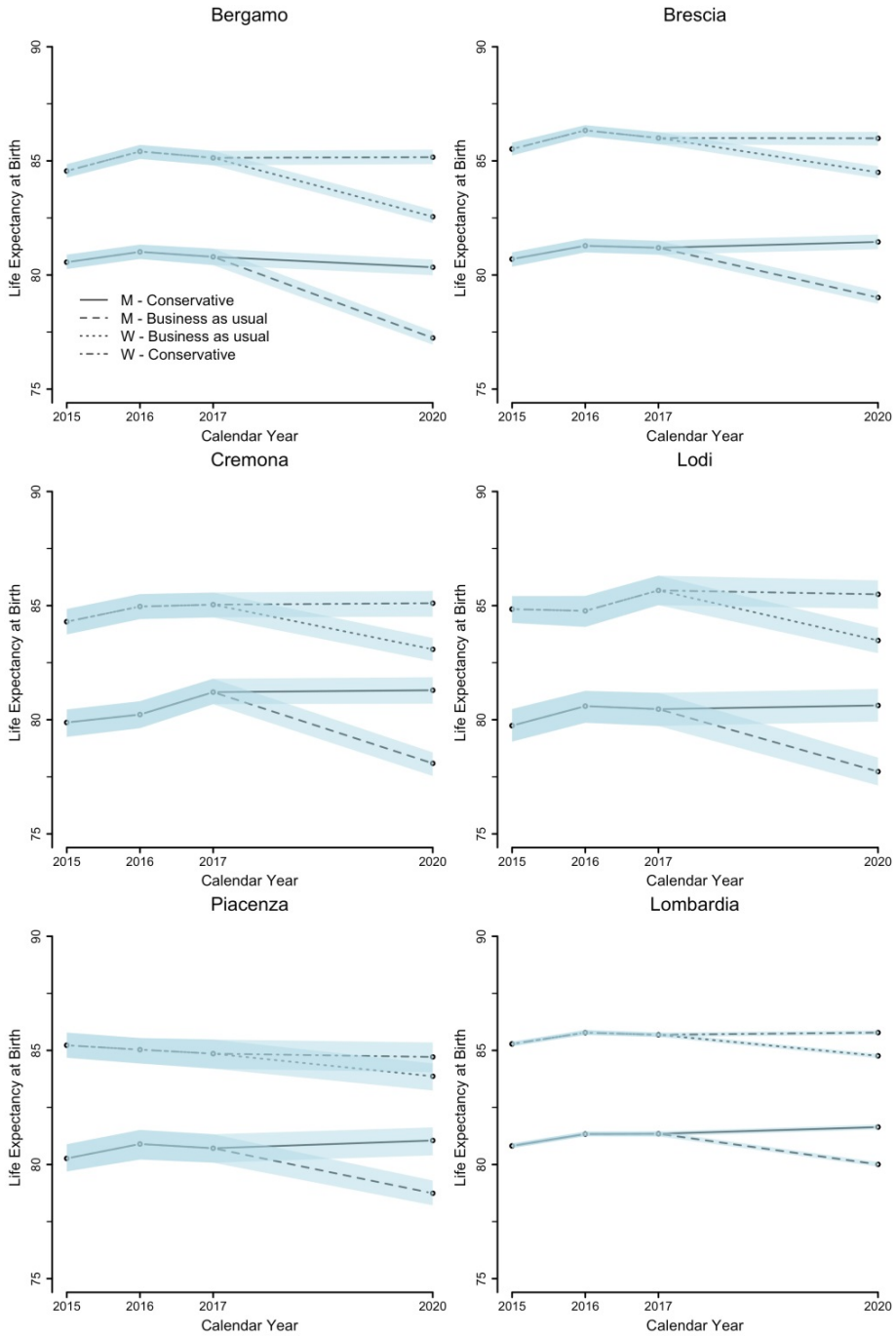


Fig.4. Estimates of annual life expectancy at birth by sex in selected provinces and in Lombardy (95% confidence intervals). Two scenarios: (A) conservative scenario; and (B) business as usual scenario. Confidence intervals (95%) for life expectancies are estimated by bootstrapping using Monte Carlo simulation methods, assuming death counts follow a binomial distribution

Supplementary Information for

News from the front: Estimation of excess mortality and life expectancy in the major epicenters 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, Via Roentgen 1, 20136, Milan, Italy
[2] Wittgenstein Centre for Demography and Global Human Capital (Univ. Vienna, IIASA, VID/ÖAW), International Institute for Applied Systems Analysis, Schlossplatz 1, A-2361 Laxenburg, Austria
[3] Wittgenstein Centre for Demography and Global Human Capital (Univ. Vienna, IIASA, VID/ÖAW), Vienna Institute of Demography, Austrian Academy of Sciences, Vordere Zollamtsstraße 3, 1030 Vienna, Austria

Corresponding authors:

* Raya Muttarak (muttarak@iiasa.ac.at)
* Benedetta Scotti (benedetta.scotti@unibocconi.it)

Supplementary text

Methods

Estimate of daily deaths up to April 15, 2020

As specified in the Data section, daily death counts are available until April 15, 2020 for about 70% of municipalities in the geographical areas of interest. For the remaining 30%, daily death counts are available up to March 31, 2020. For municipalities in the latter subset, we extrapolate daily death counts until April 15, 2020 by assuming that daily deaths in each age class evolve at the same daily rate recorded by municipalities in the former subset between April 1 and April 15. Extrapolation is carried out at the provincial level.

Supplementary figures

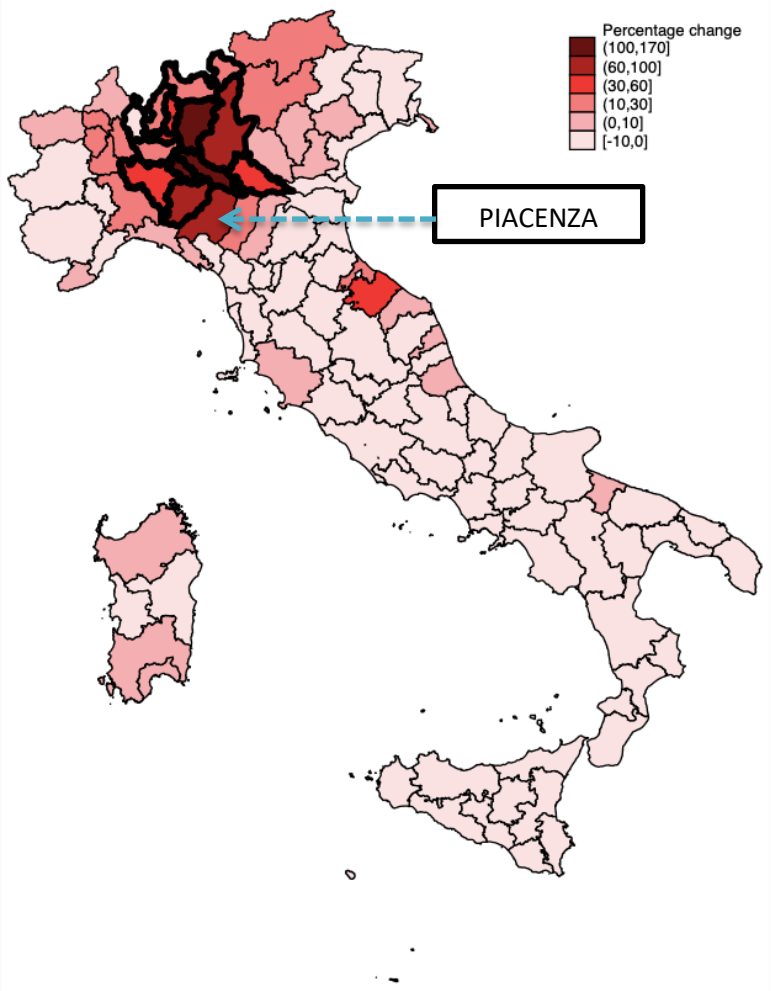


Fig. S1. Excess mortality distribution across Italian provinces. The provinces in Lombardy are highlighted by the bold black line. The province of Piacenza is indicated by the blue arrow.

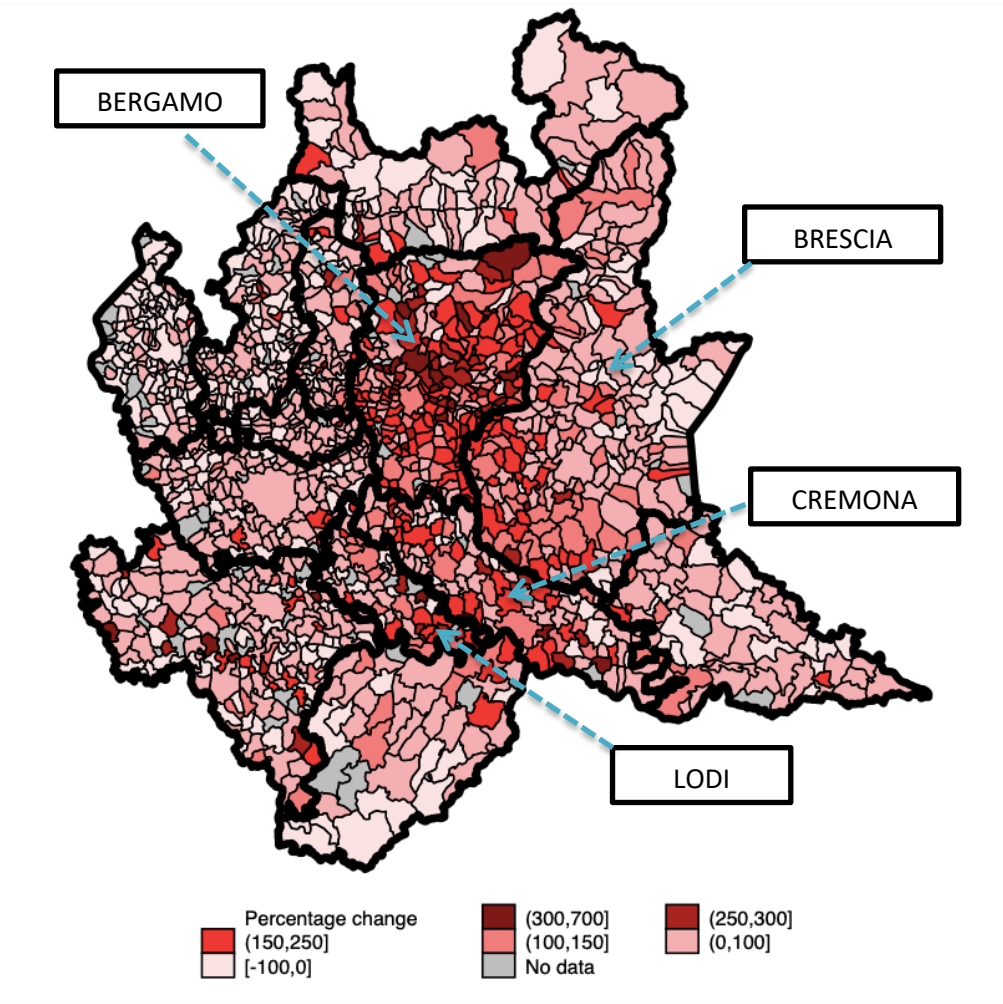


Fig. S2. Excess mortality distribution in Municipalities in Lombardy. In bold, the borders of the provinces in Lombardy. The provinces of Bergamo, Brescia, Cremona and Lodi (included in the analysis) are highlighted.

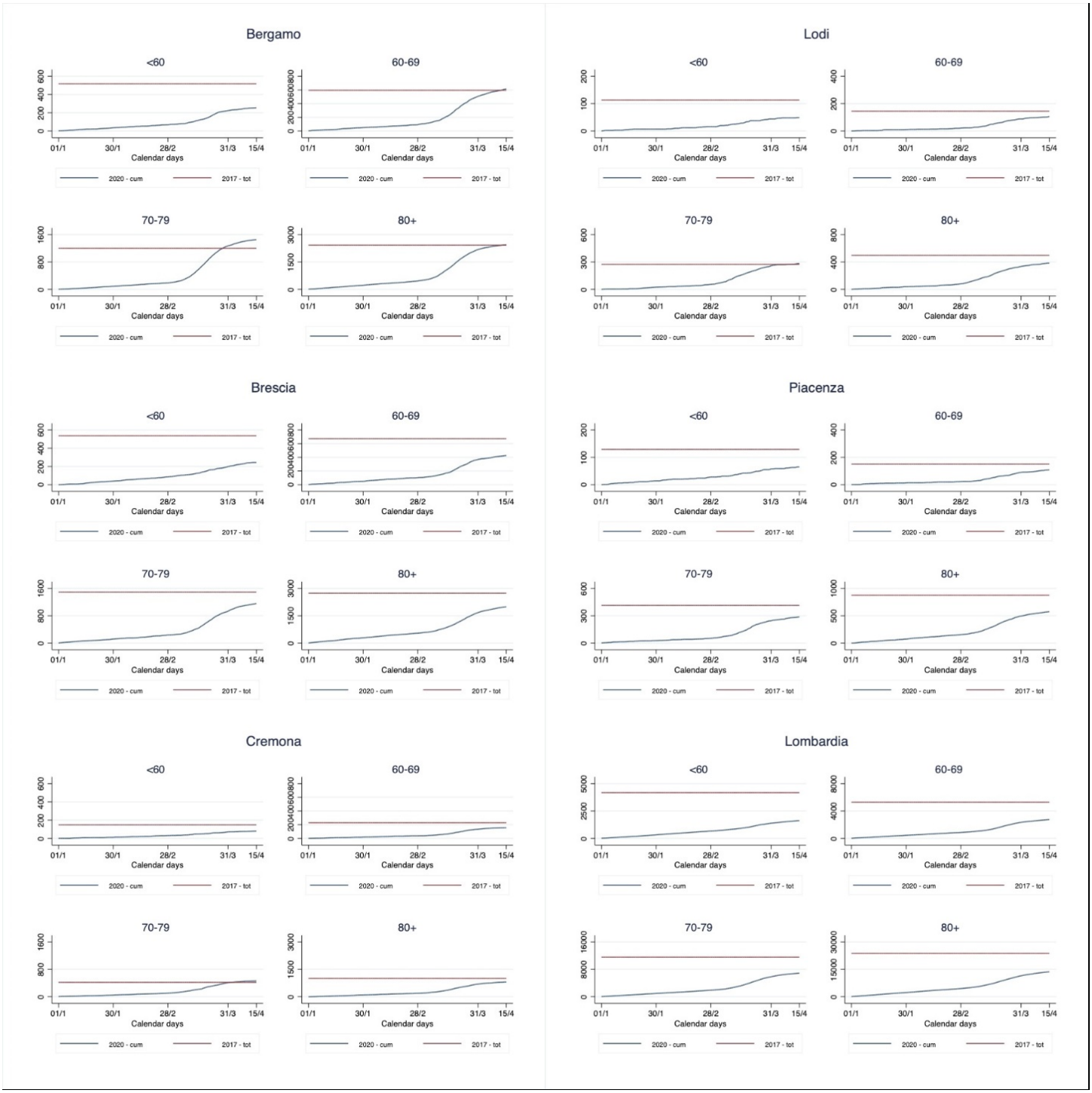


Figure S3A: Cumulative number of deaths (January 1-April 15 2020) in the five provinces and in Lombardy for men. The red line represents the total number of death in 2017.

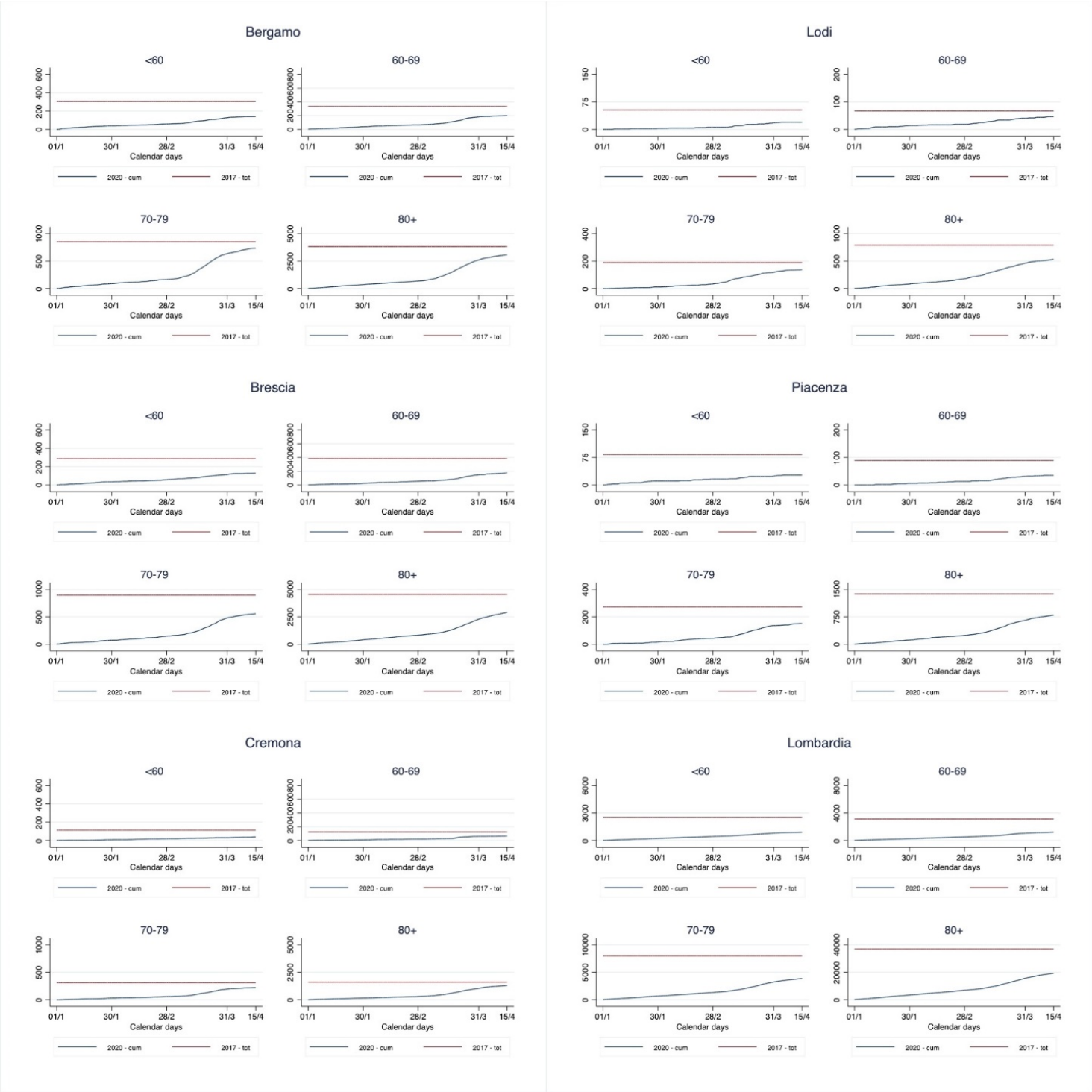


Figure S3B: Cumulative number of deaths (January 1- April 15 2020) in the five provinces and in Lombardy for men. The red line represents the total number of death in 2017.

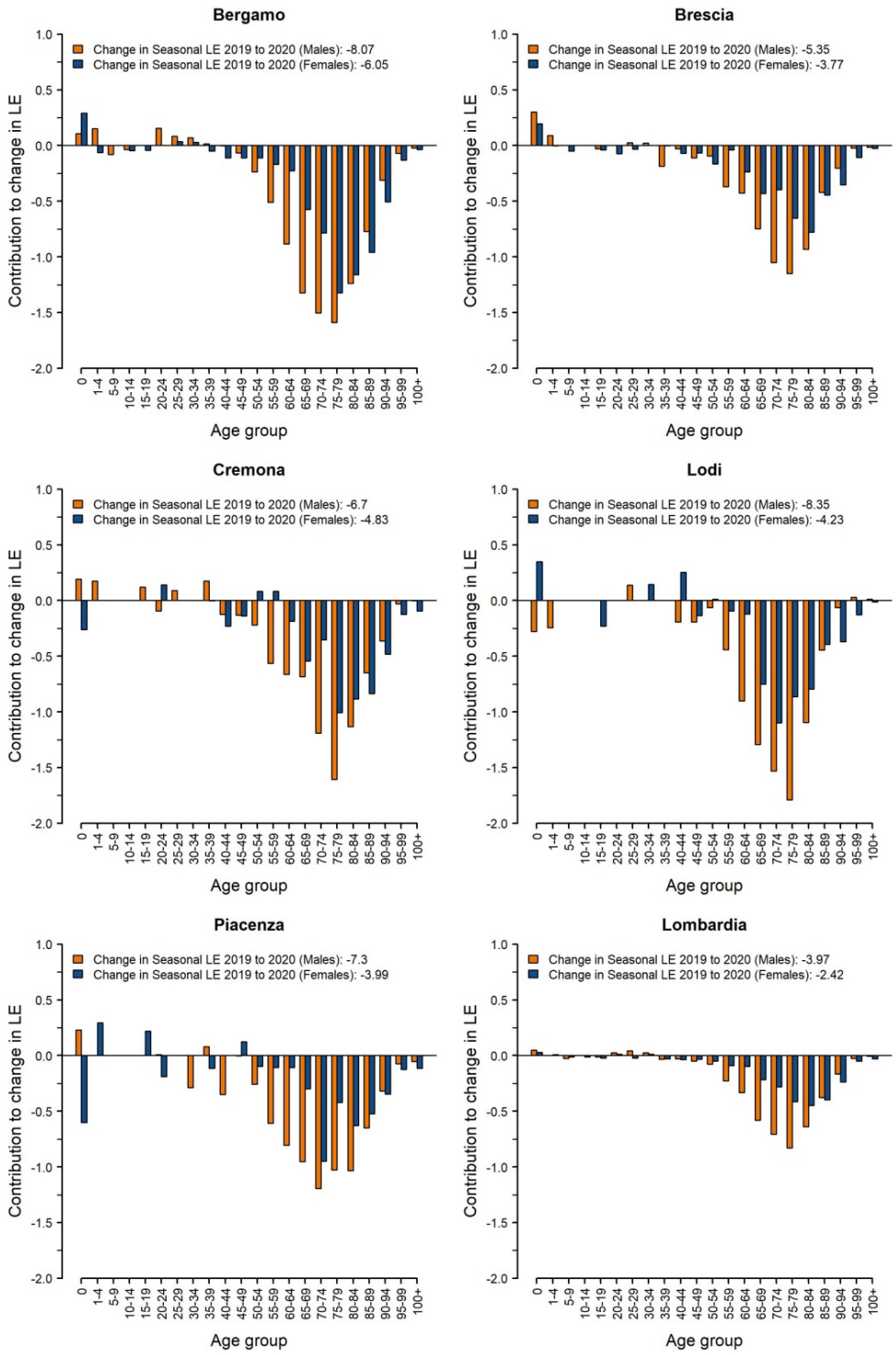


Fig. S4. Decomposition of the seasonal loss of life expectancy, by age and sex and province.

Note: For decomposing changes in life expectancy into age-specific contributions, the method proposed by Arriaga (1) is applied. This approach is used to assess which age-groups have primary contributed to the change in seasonal life expectancy between 2019 and 2020.

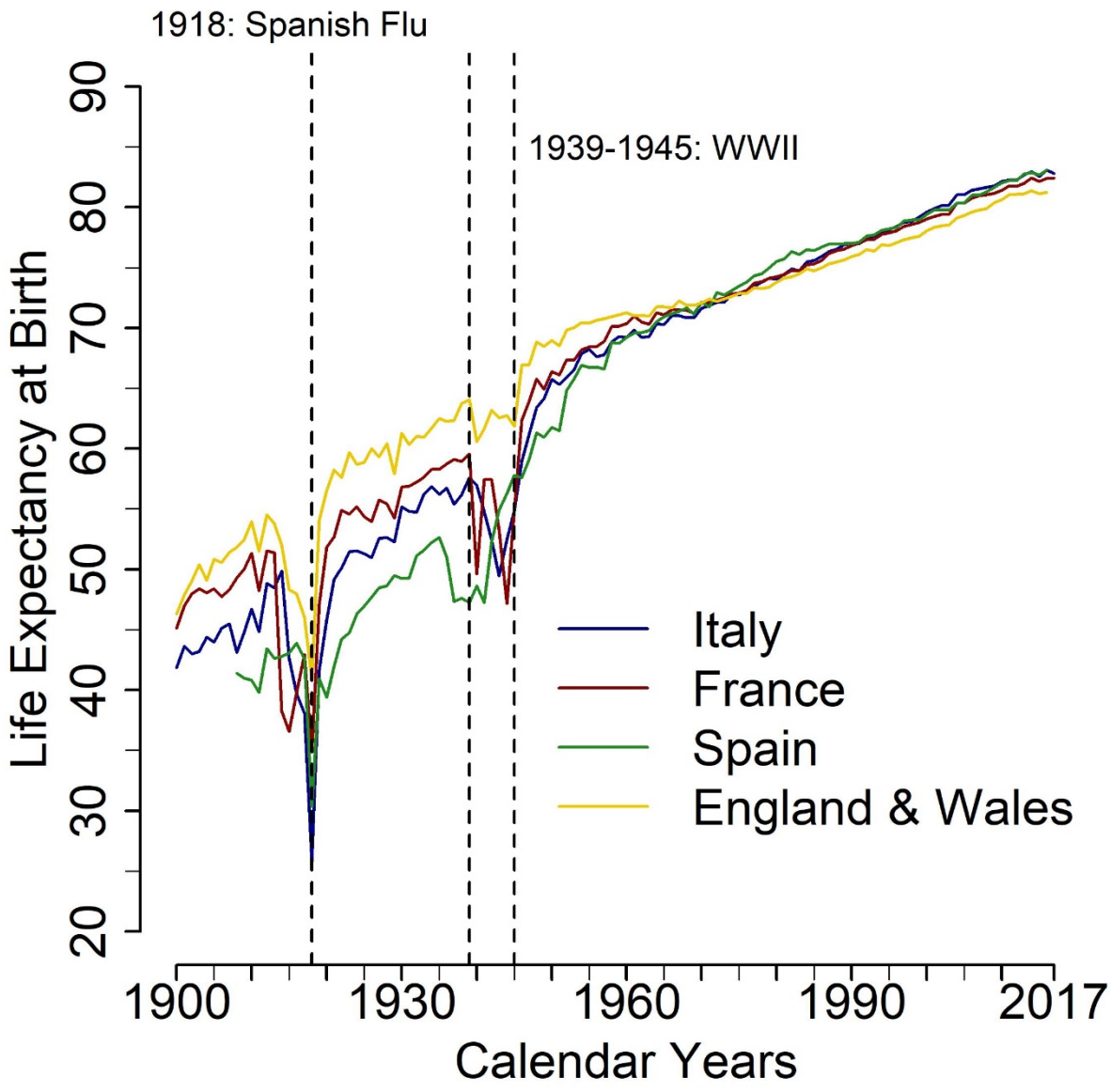


Fig. S5. Trends in life expectancies by countries since 1900.

Note: Own elaborations from the Human Mortality Database.

Table. S1. Total population, excess mortality and male to female relative risk (RR) 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**
40-49	83509	88371	8	12	1.21	1.21	1.50
50-59	84680	87182	104	32	2.02	2.20	1.53
60-69	65702	63970	408	96	3.10	2.97	1.89
70-79	53943	47212	1067	464	2.25	3.75	2.69
80-89	34366	21564	1313	1013	1.79	3.23	2.48
90+	8521	2647	356	813	1.28	2.86	2.44
BRESCIA							
Age	**Population Males**	**Population Females**	**Excess deaths Males**	**Excess deaths Females**	**RR (M/F)**	**2020/(2015-19) Males**	**2020/(2015-19) Female**
40-49	96787	101696	6	2	1.99	1.14	1.10
50-59	95941	97257	75	30	1.85	1.83	1.52
60-69	74275	71060	203	62	2.54	1.91	1.54
70-79	63616	55272	699	250	2.40	2.52	1.82
80-89	41332	25220	808	662	1.69	2.18	1.84
90+	11317	3289	239	725	1.19	1.93	2.01
CREMONA							
Age	**Population Males**	**Population Females**	**Excess deaths Males**	**Excess deaths Females**	**RR (M/F)**	**2020/(2015-19) Males**	**2020/(2015-19) Female**
40-49	26859	27935	4	6	1.28	1.33	2.00
50-59	27747	28265	31	1	2.48	2.15	1.05
60-69	22991	22158	83	22	2.47	2.15	1.51
70-79	19847	17233	298	110	2.43	2.80	2.00
80-89	14157	8413	334	378	1.40	2.43	2.24
90+	3894	1198	147	342	1.29	2.58	2.29
LODI							
Age	**Population Males**	**Population Females**	**Excess deaths Males**	**Excess deaths Females**	**RR (M/F)**	**2020/(2015-19) Males**	**2020/(2015-19) Female**
40-49	17808	18599	4	0	2.30	1.50	1.00
50-59	17423	18077	14	0	2.52	1.70	1.00
60-69	13633	13116	65	23	2.40	2.59	2.00
70-79	11515	9874	191	72	2.38	3.05	2.07
80-89	7611	4446	173	118	1.85	2.34	1.73
90+	1838	549	36	131	1.08	1.78	2.07

PIACENZA							
Age	**Population Males**	**Population Females**	**Excess deaths Males**	**Excess deaths Females**	**RR (M/F)**	**2020/(2015-19) Males**	**2020/(2015-19) Female**
40-49	20715	20861	4	-3	5.46	1.57	0.40
50-59	21480	21466	22	6	2.35	1.88	1.43
60-69	17056	16336	60	3	3.25	2.22	1.09
70-79	15661	13350	157	70	2.23	2.21	1.86
80-89	11442	7282	215	156	1.59	2.06	1.61
90+	3366	1144	65	151	1.21	1.71	1.66
LOMBARDY							
Age	**Population Males**	**Population Females**	**Excess deaths Males**	**Excess deaths Females**	**RR (M/F)**	**2020/(2015-19) Males**	**2020/(2015-19) Female**
40-49	761233	780876	35	10	1.74	1.11	1.05
50-59	761806	756097	328	108	1.83	1.43	1.22
60-69	595072	550240	1007	240	2.42	1.58	1.24
70-79	525130	439579	3178	1269	2.14	1.85	1.49
80-89	351948	219348	4471	3521	1.60	1.79	1.53
90+	90685	29070	1427	3437	1.21	1.68	1.60

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)

Supplementary material: Data

Table S2. Data coverage (number of municipalities) by province and for Lombardy

Province / Region	# of municipalities	# of municipalities in the ANPR	Data until March 31 2020			Data until April 15 2020		
			# of municipalities	% of municipalities	% of population	# of municipalities	% of municipalities	% of population
Bergamo	243	237	237	97,5	98,4	171	70,4	77,7
Brescia	205	201	201	98,0	98,9	147	71,7	76,8
Cremona	113	112	112	99,1	99,8	79	69,9	76,2
Lodi	60	57	57	95,0	98,1	53	88,3	89,4
Piacenza	46	42	42	91,3	95,5	32	69,6	81,3
Lombardia	1505	1441	1441	95,8	97,1	1066	70,8	78,3

Note: ANPR refers to the Anagrafe Nazionale della Popolazione Residente (National Register of Resident Population).

SI References

1. E. E. Arriaga, Measuring and Explaining the Change in Life Expectancies. *Demography* **21**, 83–96 (1984).