



Institute for
New Economic Thinking
AT THE OXFORD MARTIN SCHOOL

Measuring excess mortality: the case of England during the Covid-19 Pandemic

Janine Aron and John Muellbauer

18th May 2020

INET Oxford Working Paper No. 2020-11



Measuring excess mortality: the case of England during the Covid-19 Pandemic

Janine Aron and John Muellbauer, INET@Oxford, updated and revised version, 18 May 2020

Summary:

Excess mortality data avoid miscounting deaths from under-reporting of Covid-19-related deaths and other health conditions left untreated. EuroMOMO tracks excess mortality for 24 European states. England had the *highest peak weekly excess mortality* in total, for the over-65s, and, most strikingly, for the 15-64 age-group. Research is needed into such divergent patterns. A simple measure of excess mortality is the P-score: excess deaths divided by 'normal' deaths. National statistical offices should publish P-scores for states and sub-regions, and permit EuroMOMO to publish P-scores as well as their less transparent Z-scores. This would aid comparability, better to inform pandemic policy and draw lessons across heterogeneous regions and countries.

Why is it important to examine excess mortality data?

Excess mortality is a count of deaths from all causes relative to what would normally have been expected. In a pandemic, deaths rise sharply, but causes are often inaccurately recorded. The death count¹ attributed to Covid-19 may thus be significantly undercounted. Excess mortality data overcome two problems in reporting Covid-19-related deaths. Miscounting from misdiagnosis or under-reporting of Covid-19-related deaths is avoided. Excess mortality data include 'collateral damage' from other health conditions, left untreated if the health system is overwhelmed by Covid-19 cases.

Excess mortality data can be used to draw lessons from cross- and within-country differences and help analyse the social and economic consequences of the pandemic and relaxing lockdown restrictions. For country comparisons (where under-recording may differ), policy-makers should examine *robust measures* expressed relative to the benchmarks of 'normal' deaths. Normal death rates reflect persistent factors that differ between countries or regions such as the age composition of the population, the incidence of smoking and air pollution, the prevalence of obesity, poverty and inequality, and the normal quality of health service delivery. Excess death rates therefore already take considerable account of heterogeneity between different places, which enhances their comparability in picking up the differential effects of a pandemic.

Estimating the R_0 is crucial for assessing the rate and nature of relaxation of lockdowns.² Excess death rates also help to avoid the measurement biases inherent in other data used to estimate R_0 used in epidemiological models.³

How is excess mortality measured and who measures it?

National statistical agencies publish actual weekly deaths and averages of past 'normal' deaths. The Office for National Statistics (ONS) reports 'normal' deaths for England and for Wales as the average

¹ See webpage: [COVID-19 Dashboard by the Center for Systems Science and Engineering \(CSSE\)](#), Johns Hopkins University.² R_0 is the virus reproduction rate, which needs to be kept below 1 to avoid exponential growth of infections.

² R_0 is the virus reproduction rate, which needs to be kept below 1 to avoid exponential growth of infections.

³ Details on this can be found in the evidence of Prof. John Edmunds to the UK Science and Technology Parliamentary Select Committee on 7th May.

of the previous five years' deaths. They do not publish benchmarks for more granular or disaggregated data, such as sub-regions or cities. Using the weekly historical data, researchers could calculate such benchmarks with some effort.

To obtain cross-European comparisons requires data collation from individual national agencies - unless the Z-scores compiled by EuroMOMO⁴ for 24 states are used, see Box 1. EuroMOMO's timely measures of weekly excess mortality in Europe allow comparisons of the mortality patterns between different time-periods and countries, and by age-groups. The Z-scores standardise data on excess deaths by scaling by the standard deviation of deaths. EuroMOMO are currently not permitted to publish actual excess death figures by country and do not publish the standard deviations used in their calculations. However, they graph the Z-scores and the estimated confidence intervals back to 2015 providing a useful visual guide to their variability. In the U.S., the National Centre for Health Statistics publishes data on excess deaths and P-scores (see Box 1), defining excess deaths as deviations from 'normal' deaths plus a margin adjusting for the uncertainty of the data.⁵ These data include counties and states, and are disaggregated by gender, age and ethnicity. The NCHS thus sets an international standard for statistical agencies.

At least three separate journalistic endeavours have recently engaged in the time-consuming effort of collecting and presenting more transparent excess mortality data, see Table 1. The Financial Times plots numbers of excess deaths, and the P-score or percentage of deaths that are above normal deaths. The Economist shows figures and graphics for excess deaths but not P-scores. However, the published estimates of P-scores in newspapers give only a recent snapshot, missing the context of historical variability provided by EuroMOMO. And we only have P-scores for some countries, regions and cities.

In contrast to the P-scores, the Z-scores are a measure that is less easily interpretable. Moreover, if the natural variability of the weekly data is lower in one country compared to another, then the Z-score could lead to exaggeration of excess mortality compared to the P-score. While this could be a problem, selective comparisons below between Z- and P-scores, reveal that England had the highest rate of excess deaths from *both measures* of all the European states covered. Therefore, comparisons of Z-scores remain highly relevant.

⁴ [EuroMOMO](#) is a European mortality monitoring entity, aiming to detect and measure excess deaths related to seasonal influenza, pandemics and other public health threats. Official national mortality statistics are provided weekly from the 24 European countries and regions in the EuroMOMO collaborative network, supported by the European Centre for Disease Prevention and Control (ECDC) and the World Health Organization (WHO).

⁵ See webpage: "[National Center for Health Statistics](#)", Centers for Disease Control and Prevention (CDC), US Government.

Table 1: Three journalistic endeavours to examine comparative excess mortality data for Europe, the UK and the US, and other countries (for ‘all age groups’ only).

| <i>Source</i> | <i>Measure</i> | <i>Period of measurement</i> | <i>Benchmark</i> | <i>Locations compared</i> | <i>First publication date</i> |
|---------------------|---|---|---|---|--|
| The Economist | Numbers of excess deaths. <i>[The raw data - published as of 15 May - allow P-scores for 17 countries to be calculated.]</i> | Weekly; approximately monthly in one table. | “Expected deaths”, averages ranging from 2 to 5 years | UK and its constituent nations and regions, 12 European countries, Departments in France, Ecuador, Indonesia, Turkey, U.S. including states. Selected cities. | 16 April 2020. <i>Regularly updated (raw country data, including expected deaths, published on GitHub).</i> |
| The Financial Times | Numbers of excess deaths, cumulative P-score. | Weekly and cumulative, from beginning of outbreak. | Historical average of deaths over 2015-2019. | UK and its constituent nations and regions, 12 European countries, Israel, U.S. and South Africa, 10 global cities or conurbations. | 26 April 2020. <i>Regularly updated.</i> |
| The New York Times | P-score (percentage above ‘normal’) and numbers of excess deaths. | Weekly or monthly, differs per country. For the U.S. states: 9 March- 11 April. | Average of deaths, data-dependent, & differing by country. 5-years for the U.S. | U.S. States, France, Spain, Germany, Ecuador, Netherlands, Belgium, Sweden, Switzerland, Norway, Denmark; New York City, Jakarta, Istanbul. | 30 April 2020. <i>Regularly updated.</i> |

Sources: The Economist (2020): “[Tracking Covid-19 excess deaths across countries](#)”, 16 April, 2020, and Tozer (2020): “[Measuring the true toll of the pandemic](#)”, 24 April, 2020. For the Economist, Tozer measures excess deaths from the week the first 50 Covid deaths were reported, to around April 12. As of 15 May, The Economist’s J. Tozer and M. González publish the raw country data on [GitHub](#). Also see “[Global coronavirus death toll could be 60% higher than reported](#)”, Financial Times, 26 April, 2020 and Wu et al. (2020): “[46,000 Missing Deaths: Tracking the True Toll of the Coronavirus Outbreak](#)”, The New York Times, 30 April, 2020.

Box 1: Measures of excess deaths: comparing and contrasting the Z-score and the P-scores

Denote the number of weekly deaths by x .

The P-score is defined as follows:

$(x$ minus the expected value of x for the population), divided by the expected value of x for the population

A variant P-score (U.S. National Center of Health Statistics) is defined as follows:

$(x$ minus the upper threshold for the expected value of x for the population), divided by the upper threshold for the expected value of x for the population.

- *The upper threshold* is defined as the expected value plus the 2.5% confidence interval for this expected value. This takes into account uncertainty created by the natural variability of x .

The Z-score is defined as follows:

$(x$ minus the expected value of x for the population), divided by the standard deviation for the population of x around its expected value.

- EuroMOMO estimate the expected value of each country's weekly deaths using data for the previous five years, taking seasonal factors and trends into account, and adjust for delays in registration.
- For count data, like weekly deaths, a Poisson distribution is a reasonable approximation to the underlying probability, taken into account in the estimated Z-scores of EuroMOMO. *
- Graphs published for each country show the weekly Z-scores since 2015 compared to their usual range of -2 to +2, the approximate 95% confidence interval. Around 2.5% of observations would thus usually have a Z-value over 2. The Z-score equals 4 line is also shown, corresponding to a 'substantial increase': under usual conditions, the Z-value would exceed 4 only around 0.003% of the time.
- The graphs show more deviations of Z-scores 'exceeding 2' and 'exceeding 4', than one would expect. The main reason is that to fit the baseline, EuroMOMO chose only the period of the year when additional processes (e.g. Winter influenza and Summer heat waves) leading to excess deaths are not likely to happen. Normal variability is thus measured after excluding these seasons. **

* The Poisson is a discrete probability distribution that expresses the probability of a given number of events occurring in a fixed interval of time if these events occur with a known constant mean rate and independently of the time since the last event. The calculation is described in Farrington et al. (1996). As the data are probably not time-independent, the claimed probabilities associated with different Z-values are likely to be only approximate.

** See EuroMOMO webpage: "[Methods](#)".

Key issues for comparing rates of excess mortality

To interpret large differences in excess mortality between nations requires consideration of three main factors, and the within-nation deviations in these factors: the average infection rates in preceding weeks, average mortality risk from Covid-19 and constraints on Covid-19-specific health capacity. Suppose for simplicity's sake that every adult has the same chance of being infected and that effect of Covid-19 is to increase every infected adult's initial risk of dying in the next year by the same proportion. According to UK data examined by Spiegelhalter (2020), for an infected adult, but not for children, Covid-19 roughly doubles their initial risk.

Compare two countries or regions with the same average Covid-mortality risk where 1% of all adults are infected in A, while 5% are infected in B. Then the rate of excess deaths measured by the P-score will be about 5 times as large in B in the weeks following the incidence of the infection. Countries that locked down early and had effective test, trace and isolate procedures kept down the average infection rate and hence the excess death rate.⁶ Within countries, infection rates can differ. London's higher excess mortality, see FT May 12, was influenced by higher initial imports of infections and a higher virus reproduction number given its high density and hard-to-avoid close physical contact on public transport and at work. Thus, countries that have a higher fraction of adults in locations or occupations where the virus can more easily spread will tend to have higher excess death rates.

Second, mortality risks for infected adults can differ between and within countries. For example, the percentage increase in mortality risk may be greater for some ethnic groups, or for some co-morbidities such as diabetes or pre-existing lung conditions. Then country differences in the prevalence of obesity and smoking will influence comparative excess mortality.

Finally, a country's P-score is further driven up, and potentially *much* further, by limited Covid-specific health capacity. The death rate among infected adults depends on capacity constraints on numbers of hospital beds and staff, numbers of ventilators, PPE, testing and logistical failures in delivery e.g. to care homes. Given similar initial capacities, a country with a higher average infection rate will be more likely to run into these constraints. By the same logic, given the same high infection rate, a country with lower health capacity would have a higher rate of excess mortality. This is why there is such a focus on 'flattening the pandemic curve'.

Different capacity constraints can have different implications for different groups. For example, lack of PPE and testing facilities in care homes will have disproportionately larger effects on mortality for the oldest individuals and this could affect country comparisons.

Covid-19, therefore, interacts with the age distribution, the nature of health service delivery, poverty and inequality, ethnic and occupational structures, air pollution, the relative size of major conurbations and so on. Comparing rates of excess mortality statistics within countries by age groups, by city size and by occupational, social and ethnic groups should generate important insights for future pandemic policy.

⁶ Transmission and hence rates of infection are also influenced by factors like the nature of social distancing, availability and use of face masks, and cultural differences in the exercise of self-discipline and following of advice.

Table 2: EuroMOMO Z-scores for poor performers showing peak weeks of excess mortality by age-group

| EuroMOMO Z-scores | | | | | | | |
|--------------------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| For week ending: | 22-Mar-20 | 29-Mar-20 | 5-Apr-20 | 12-Apr-20 | 19-Apr-20 | 26-Apr-20 | 3-May-20 |
| All age-groups | <i>Week 12</i> | <i>Week 13</i> | <i>Week 14</i> | <i>Week 15</i> | <i>Week 16</i> | <i>Week 17</i> | <i>Week 18</i> |
| England | 5.58 | 15.83 | 33.8 | 42.75 | 37.33 | 30.46 | 22 |
| Spain | 12.89 | 31.27 | 34.41 | 26.7 | 15.7 | 6.62 | 1.48 |
| Belgium | 5.17 | 14.18 | 26.84 | 29.91 | 20.53 | 11.74 | 5.02 |
| Italy | 11.75 | 18.46 | 22.16 | 15.52 | 9.47 | 6.86 | 4.36 |
| Netherlands | 6.37 | 14.82 | 21.13 | 20.61 | 14.53 | 10.91 | 5.12 |
| France | 5.78 | 12.71 | 21.7 | 18.59 | 12.41 | 3.81 | -1.74 |
| The over-65s | <i>Week 12</i> | <i>Week 13</i> | <i>Week 14</i> | <i>Week 15</i> | <i>Week 16</i> | <i>Week 17</i> | <i>Week 18</i> |
| England | 4.5 | 13.48 | 29.6 | 37.97 | 33.72 | 27.74 | 20.27 |
| Spain | 12.55 | 30.39 | 33.77 | 26.59 | 15.43 | 6.94 | 1.7 |
| Belgium | 4.98 | 15.15 | 28.12 | 31.27 | 22.1 | 12.18 | 5.19 |
| Italy | 11.86 | 19.28 | 22.89 | 15.7 | 9.98 | 6.97 | 4.45 |
| Netherlands | 6.24 | 14.9 | 21.69 | 21.33 | 15.29 | 11.05 | 5.08 |
| France | 5.58 | 12.52 | 21.95 | 19.03 | 12.99 | 4.25 | -1.16 |
| 65-74 age-group | <i>Week 12</i> | <i>Week 13</i> | <i>Week 14</i> | <i>Week 15</i> | <i>Week 16</i> | <i>Week 17</i> | <i>Week 18</i> |
| England | 4.9 | 12.64 | 22.32 | 27.57 | 21.27 | 15.89 | 11.4 |
| Spain | 8.77 | 24.82 | 20.53 | 14.03 | 6.43 | 2.78 | 0.47 |
| Belgium | 4.43 | 8.74 | 12.96 | 10.05 | 9.54 | 3.86 | 1.98 |
| Italy | 7.84 | 9.49 | 11.47 | 5.66 | 3.66 | 2.33 | 1.91 |
| Netherlands | 4.08 | 7.92 | 9.98 | 9.08 | 4.98 | 4.29 | 2.38 |
| France | 6.29 | 9.36 | 13.83 | 10.65 | 6.48 | 1.96 | -0.99 |
| 15-64 age-group | <i>Week 12</i> | <i>Week 13</i> | <i>Week 14</i> | <i>Week 15</i> | <i>Week 16</i> | <i>Week 17</i> | <i>Week 18</i> |
| England | 6.06 | 6.06 | 13.59 | 24.84 | 28.02 | 20.9 | 15.67 |
| Spain | 4.02 | 4.02 | 10.03 | 8.82 | 4.75 | 4.33 | -0.55 |
| Belgium | 2.56 | 2.56 | 2.28 | 6.23 | 6.8 | 2.15 | 2.49 |
| Italy | 2.19 | 3.82 | 3.66 | 5.3 | 5.58 | 1.57 | 2.01 |
| Netherlands | 3.82 | 2.19 | 4.51 | 4.83 | 3.97 | 2.32 | 3.13 |
| France | 2.95 | 2.95 | 5.43 | 7.17 | 5.12 | 2 | -0.27 |

Source: Z-scores extracted from the EuroMOMA graphs, 16-May-2020.

Notes: (i) The peak weeks for different countries are in **bold**. (ii) The country ordering is by peak mortality. (iii) The ONS defines a week as ending on Friday; Public Health England (PHE) and EuroMOMO define a week as ending on Sunday.

A first look at data for England: total figures and the over 65s

England eclipses all 24 countries covered by EuroMOMO in excess mortality scores. At the peak, England's total Z-score based on *actual* deaths was 42.8 (week 15 in Table 2). The ONS records 21,182 *registered* deaths for the comparable week compared to a normal number of 9787 (averaging the previous 5 years).⁷ This gives excess *registered* deaths of 11,395, and a P-score of 1.164. For the same week, the ONS registered 8335 deaths as Covid-19-related, accounting for 73% of excess deaths. Data on *actual* deaths, reported by The Economist, give a peak P-score of 1.134.

⁷ This week, ending in 17 April, corresponds roughly with EuroMOMO's peak week 15, see definitional differences in Table 2. The ONS reports the *registration* of deaths, which lags the *actual* deaths. See the webpage: "[Deaths registered weekly in England and Wales](#)", Office for National Statistics.

England’s peak rate of excess deaths for the most vulnerable age-group, the over-65s, is also the highest (Table 2, and Figure 1 for the next worst affected European countries). England’s far higher excess mortality scores than the rest of the UK are shown in Figure 2; these countries’ scores are lower than Spain, Belgium, Italy, Netherlands, and France. Figure 2 shows that not only is England distinctive in the rate of excess mortality in the peak week, but the same is true in comparisons of the three weeks before the peak and the subsequent three weeks. Italy initially dominated the headlines for Covid-19-related deaths but ranked fourth for peak excess mortality figures for the over-65s, below Spain and Belgium. In contrast, Germany, throughout the nine weeks in Figure 1 showed excess mortality well within the -2, +2 normal range.⁸

As a spot-check, P-scores were calculated from *actual* deaths and normal deaths, reported by The Economist. Peak P and Z scores are compared in Table 3. Within Europe, the rankings almost coincide.⁹

Table 3: Peak week of excess mortality: country P-scores and Z-scores compared

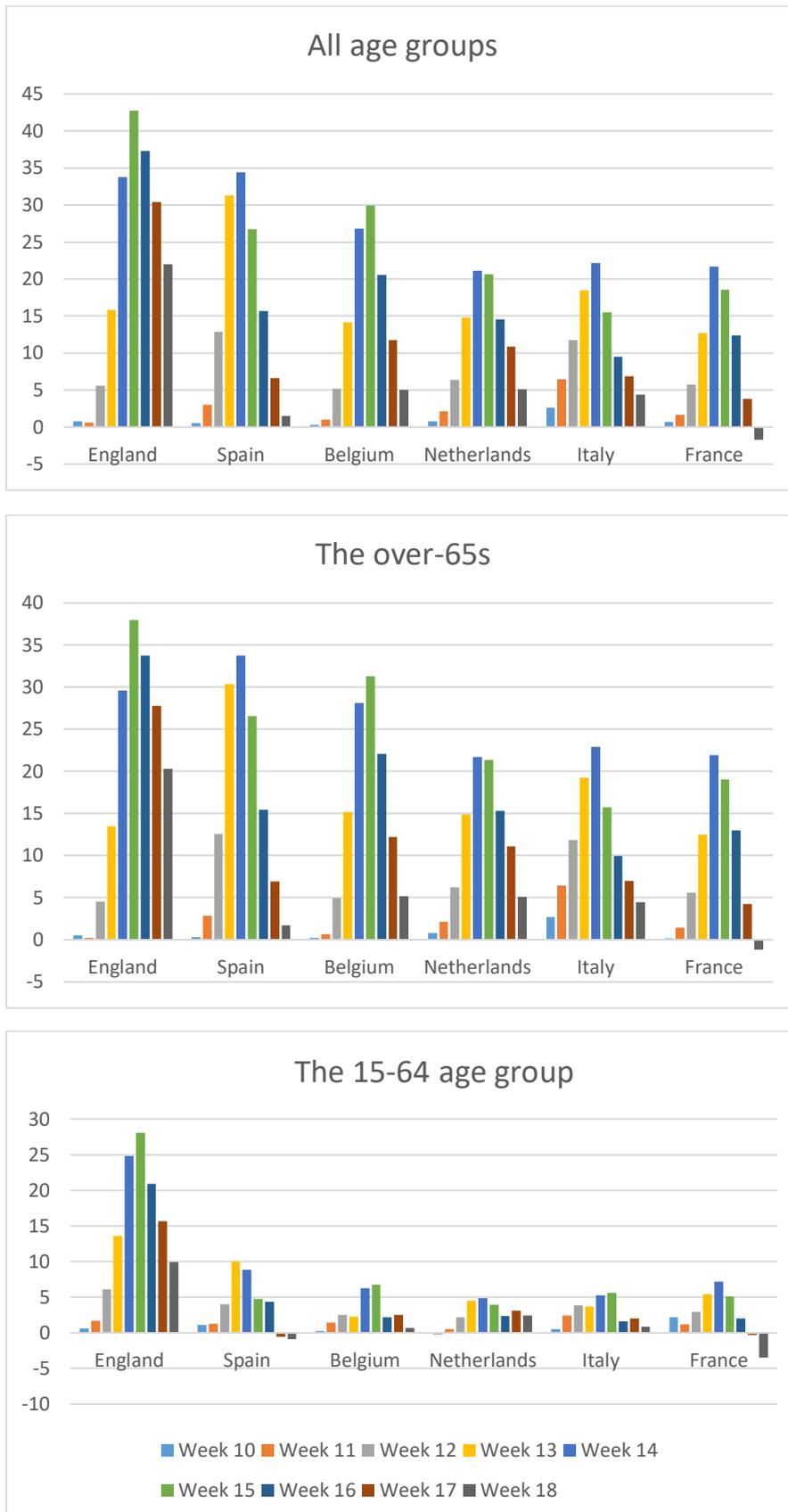
| <i>Excess mortality scores</i> | | |
|--------------------------------|---------|---------|
| <i>All age groups</i> | P-score | Z-score |
| England | 1.13 | 42.75 |
| Spain | 1.05 | 34.41 |
| Belgium | 0.94 | 29.91 |
| Italy | 0.81 | 22.16 |
| Netherlands | 0.71 | 21.13 |
| France | 0.64 | 21.70 |
| Rest of UK | | |
| Wales | 0.76 | 19.71 |
| Scotland | 0.81 | 15.57 |
| N. Ireland | 0.58 | 8.90 |

Sources: EuroMOMO Z-scores for the peak week of excess mortality, see Table 2. The peak weeks for the all age group category is the same for the UK countries. The P-Scores for the peak week of excess mortality are calculated by the authors from newly-released country data by The Economist [GitHub](#). The peak week timing for P-scores and Z-scores coincide for all countries.

⁸ German data are for Berlin and the state of Hesse and are available from EuroMOMO.

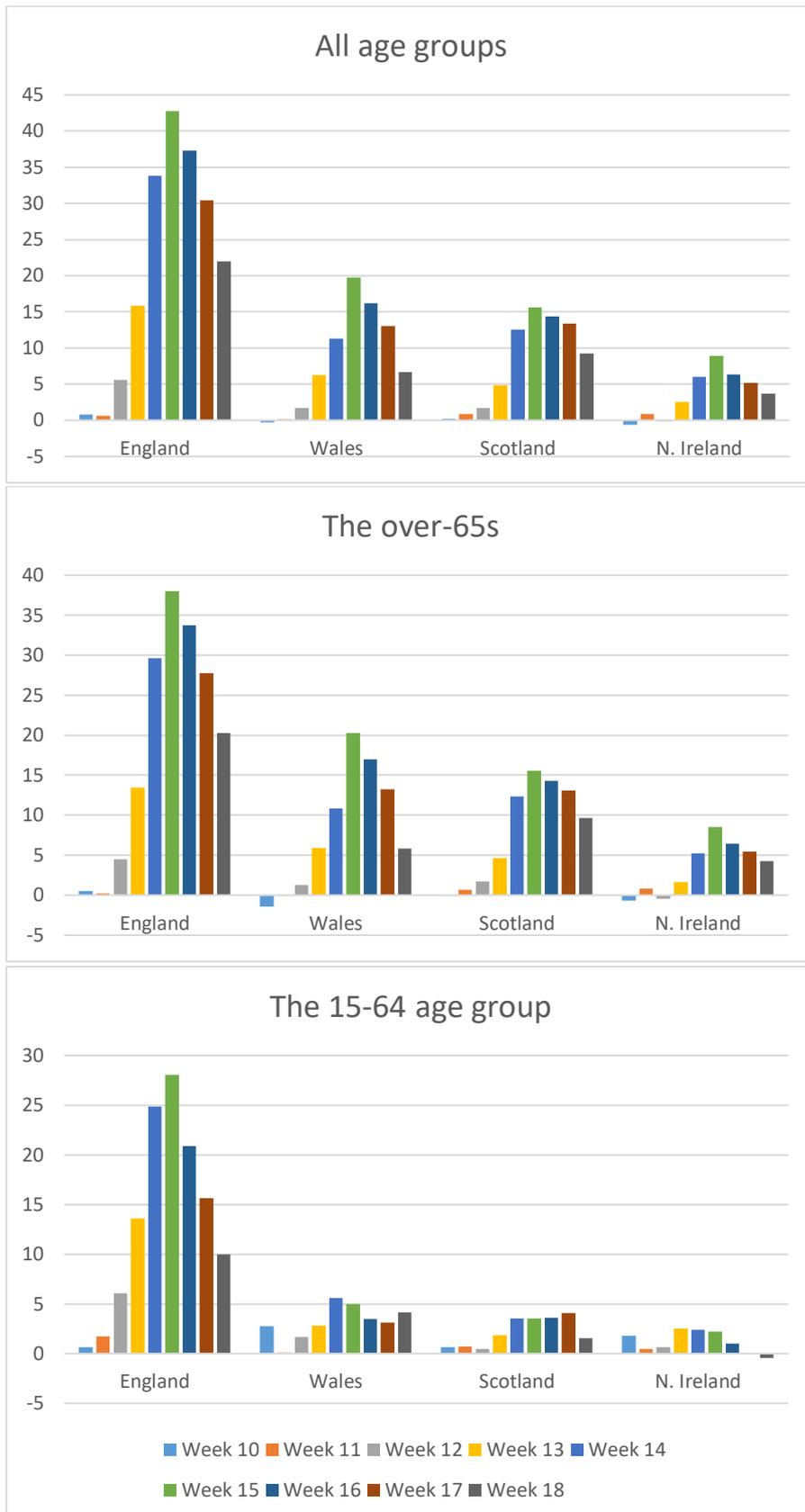
⁹ Scotland’s peak P-score is higher than that of Wales, though if cumulated over a few weeks, it would be lower.

Figure 1: Recent weeks of Z-scores for poor performers showing peak weeks of excess mortality by age-group



Source: Z-scores extracted from the EuroMOMA graphs, 16-May-2020. Their week 19 data are omitted, as the most recent week's data tend to be heavily revised. Country ordering is by peak mortality for all age groups.

Figure 2: Recent weeks of Z-scores for the UK: England, Scotland, Wales and N. Ireland



Source: Z-scores extracted from the EuroMOMA graphs, 16-May-2020. Their week 19 data are omitted, as the most recent week's data tend to be heavily revised.

Excess mortality for the 15-64 and 65-74 age-groups

Most disturbing is the comparative story for the 15-64 age-group, where England's relative record in excess mortality in the Covid-19 era is strikingly higher than in the European countries. The 15-64 age-group includes the mass of the working age population. At its peak in week 15, it is 2.8 times worse than the weekly peak in next worst country, Spain, around 4 times worse than France and Belgium, and more than 5 times worse than in Italy, Table 2. Within the UK, excess deaths for this age group are also strikingly worse for England than for the other nations, see Figure 2 and Table 4. England is the *only* country in Europe, for which Z-scores for the 15-64 group had not decreased below about 2 by week 18, ending 3 May.

Puzzling too, is that Z-scores in the 65-74 age-group for England, see Table 2, are similar to the 15-64 age-group. By contrast, in the five European countries, excess deaths in the 65-74 age-group are about twice as high as for the 15-64 age-group, though still below the 65+ age-group. Might this imply that more vulnerable individuals in England retire later than their counterparts in other places, and so resemble more closely the patterns of the 15-64 age group? Excess mortality in other residential institutions could also explain part of excess mortality in the age group 15-64 in England, with similar factors driving up infection rates as operated in English care homes. A greater degree of inequality in incomes and in access to decent housing in England - compared to many European countries - could also be partly culpable. These factors would be particularly pertinent to the 15-64 age group.

As explained above, it is also possible that the infection rates for England in the weeks preceding high excess mortality were much higher than elsewhere, which would have affected all age groups. Until systematic sampling using antibody-based tests of the population in different countries at a considerable scale has occurred¹⁰, it will not be known how the comparative rates of infection in the population differed between countries. Only then will it be clear how far the mean vulnerability for the English population had shifted compared to that in other countries.

¹⁰ See "[A Workable Strategy for Covid-19 Testing: Stratified Periodic Testing rather than Universal Random Testing](#)", INET, Oxford, 22 April, 2020.

Table 4: EuroMOMO Z-scores for the UK: England, Scotland, Wales and N. Ireland

| EuroMOMO Z-scores | | | | | | | |
|--------------------------|----------------|----------------|----------------|----------------|----------------|----------------|----------------|
| For week ending: | 22-Mar-20 | 29-Mar-20 | 5-Apr-20 | 12-Apr-20 | 19-Apr-20 | 26-Apr-20 | 3-May-20 |
| All age groups | <i>Week 12</i> | <i>Week 13</i> | <i>Week 14</i> | <i>Week 15</i> | <i>Week 16</i> | <i>Week 17</i> | <i>Week 18</i> |
| England | 5.58 | 15.83 | 33.8 | 42.75 | 37.33 | 30.46 | 22 |
| Wales | 1.71 | 6.22 | 11.3 | 19.71 | 16.19 | 13.07 | 6.69 |
| Scotland | 1.73 | 4.83 | 12.54 | 15.57 | 14.35 | 13.4 | 9.27 |
| N. Ireland | -0.15 | 2.56 | 6.02 | 8.9 | 6.36 | 5.14 | 3.71 |
| The over-65s | <i>Week 12</i> | <i>Week 13</i> | <i>Week 14</i> | <i>Week 15</i> | <i>Week 16</i> | <i>Week 17</i> | <i>Week 18</i> |
| England | 4.5 | 13.48 | 29.6 | 37.97 | 33.72 | 27.74 | 20.27 |
| Wales | 1.28 | 5.89 | 10.81 | 20.29 | 16.96 | 13.24 | 5.81 |
| Scotland | 1.74 | 4.6 | 12.33 | 15.56 | 14.3 | 13.11 | 9.61 |
| N. Ireland | -0.44 | 1.65 | 5.24 | 8.51 | 6.46 | 5.47 | 4.26 |
| 15-64 age group | <i>Week 12</i> | <i>Week 13</i> | <i>Week 14</i> | <i>Week 15</i> | <i>Week 16</i> | <i>Week 17</i> | <i>Week 18</i> |
| England | 6.06 | 13.59 | 24.84 | 28.02 | 20.9 | 15.67 | 9.96 |
| Wales | 1.67 | 2.79 | 5.59 | 5 | 3.47 | 3.14 | 4.13 |
| Scotland | 0.47 | 1.83 | 3.54 | 3.53 | 3.57 | 4.06 | 1.54 |
| N. Ireland | 0.63 | 2.5 | 2.42 | 2.19 | 0.98 | 0.06 | -0.45 |

Source: Z-scores extracted from the EuroMOMA graphs, 16-May-2020.

Notes: (i) The peak weeks for different countries are in **bold**. (ii) The country ordering is by peak mortality. (iii) The ONS defines a week as ending on Friday; Public Health England (PHE) and EuroMOMO define a week as ending on Sunday.

What can be learned from the within-UK comparisons of the Z-scores?

To interpret large differences in excess mortality between nations requires consideration of the three main factors, mentioned above: the average infection rates in preceding weeks, average mortality risk from Covid-19 and constraints on Covid-19-specific health capacity. London's international connectedness and the timing and London-centric location of the spread of the infection help explain England's worse performance than Wales, Scotland and Northern Ireland. Regional data show a clear rise in London's excess deaths ahead of other regions, and a peak P-score of 2.37, far above England's 1.13. London's population density and public transport system mean physical distancing is harder to achieve. The undeniably late and initially unclear application of social distancing and delayed lock-down measures, discussed further below, then had a far worse outcome. Generally, there was a collective failure in preparedness across the public health system, especially for testing capability and adequate supplies and distribution logistics of personal protection equipment (PPE) for health workers. The late recognition of the need to provide care-homes with PPE and tests has received recent attention. Regions outside London mostly fared better, though the West Midlands and the North West, the next largest conurbations, eventually had the next highest excess death rates (data from the FT and The Economist). This underlines the roles of timing and urban density. Infection rates outside London may have been at lower levels, so when social distancing and lock-down measures were introduced, they were the more effective.

Data from the ONS on age-corrected mortality rates by location show much higher Covid-19-related death rates in places with the greatest economic deprivation.¹¹ Underlying health is likely worse in these areas and low-paid key workers, more exposed to potential infection, may live there in disproportionate numbers. This is particularly pertinent to the 15-64 age-group. High comparative levels of excess mortality in England may also have been affected by ethnic differences in the incidence of Covid-19-related deaths.¹²

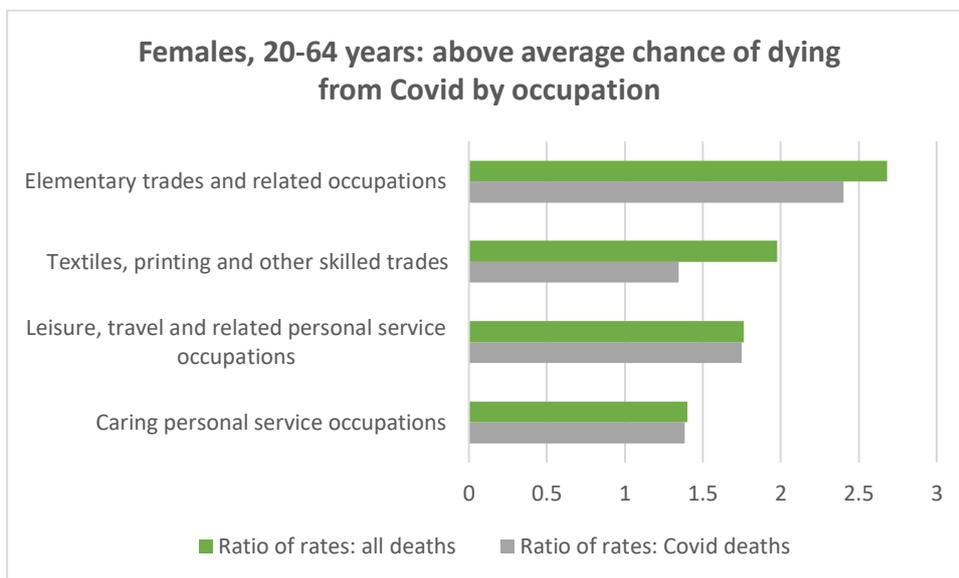
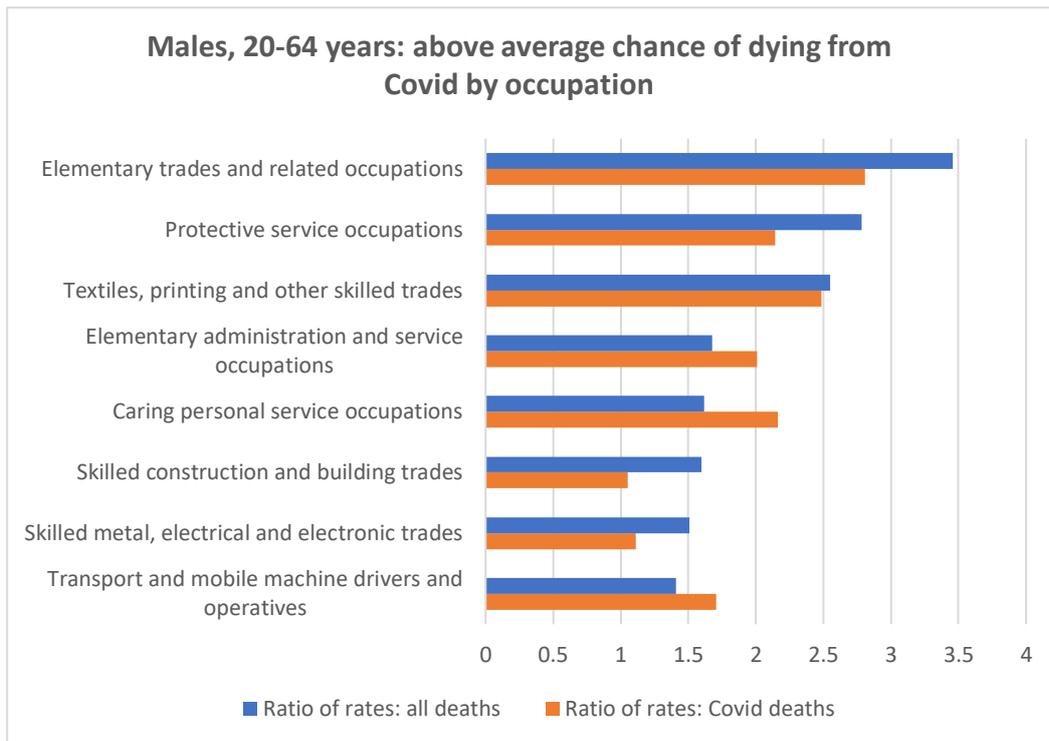
Further clues stem from ONS data on death rates by occupational groups. It is possible to examine the deviation from the average Covid death rate per 100,000 by occupation. Figure 3 shows death rate ratios to the average rate for the worst affected, mostly low-wage occupations, both for Covid-19 deaths and deaths from all causes, during the pandemic period to 20 April. The two ratios are strongly correlated and are a good indirect indicator of the corresponding rates of excess mortality.

For male transport workers, above-average mortality rates are pronounced for taxi drivers followed by bus drivers. Security guards and male care workers fare even worse. Among women, care workers have around 2.4 times the average death rates for working age women. For nurses, the ratio is 1.3, suggesting that staff in care-homes were particularly badly protected. Further work is needed to elucidate whether differences in the occupational composition of different regions could explain regional differences in excess mortality. Also, to check whether there are differences between regions in the occupation-specific excess mortality rates.

¹¹ See webpage: [“Deaths involving COVID-19 by local area and socioeconomic deprivation: deaths occurring between 1 March and 17 April 2020”](#), ONS.

¹² See Khunti et al. (2020): [“Is ethnicity linked to incidence or outcomes of Covid-19?”](#), BMJ, 20 April, 2020, and Barr et al. (2020): [“Ethnic minorities dying of Covid-19 at higher rate”](#), The Guardian, 22 April, 2020.

Figure 3: Deaths involving COVID-19 and all causes among (selected) occupation groups by gender (aged 20 to 64 years), England and Wales.



Source: ONS release, Coronavirus (COVID-19) related deaths by occupation, England and Wales, 11-May-2020.

Notes: (i) Age-standardised rates of death per 100,000 population. (ii) Figures for the most recent death registrations, deaths involving COVID-19 registered between 9th March and 20th April.

What can be learned from the cross-country comparisons of the Z-scores?

There is heterogeneity between nations and, as seen above, within nations. Ideally, to draw valid lessons for health policy and easing lockdown requires like-for-like comparisons, controlling for differences in population density, average age, prevalence of diabetes and smoking, wealth factors,

amongst others. This reinforces the need for disaggregated excess mortality data to sub-regional levels across Europe and the UK.

Nevertheless, one can draw lessons from the aggregate EuroMOMO Z-scores. Excess mortality rates rose first in Italy, then Spain and closely followed by England, Belgium, France and the Netherlands. Germany never had weekly Z-scores outside the -2, +2 normal range.¹³ Germany's better comparative record helps explain outcomes in the high excess mortality countries. Important were the early large-scale testing and contact tracing, trust in the government's social distancing guidelines and the fact that many of the initially infected, returning from skiing holidays were young and fit.¹⁴ Germany developed the first coronavirus test in mid-January¹⁵ and ensured that capacity to manufacture large quantities of the test kits was rapidly rolled out. Germany managed largely to keep the virus out of care-homes with their high mortality risk, evidenced by the lower median age of those infected. Higher capacity in health service delivery, measured by higher numbers of hospital beds, ICUs, and ventilators per head of the population, combined with the lower rate of infection, meant that Germany's health system could deliver better care, reducing mortality. Patients were admitted into hospitals at a much earlier stage if they were unable to shake off the symptoms within the course of a week. Many cities introduced 'coronavirus taxis', enabling medical staff to visit patients at home and take a test. If positive, the GP would then be in touch every other day to enquire about the development of the disease. If someone experienced breathing problems, they were sent straight to hospital where beds and, if need be, ventilators, were available. It seems that Germany built up stocks of personal protective equipment (PPE) for health workers earlier than other countries.¹⁶

The *interaction of multiple factors* is likely to account for Germany's better record. If health workers are effectively protected, not only do fewer of them become sick, affecting the delivery capability of the health system, but many fewer patients are infected. People are then less fearful of the risk of catching the virus from medical staff when seeking medical attention for other conditions. In the UK, there has been a sharp reduction in people seeking medical attention for heart attacks, strokes and other life-threatening conditions accounting for part of the surge in excess deaths in the UK. Reinforced by the much-repeated 'protect our NHS' message, patients showing symptoms, including of breathlessness,¹⁷ were encouraged to manage at home, when in Germany or South Korea they would have been routinely admitted. This policy difference helps account for the far higher death rates among Covid-19 patients admitted to UK hospitals. A further worry is that some extra deaths caused by late medical attention will show up in future years.

¹³ Complete but less timely data for all of Germany, tracked by the FT, confirm Germany and Denmark as having the lowest cumulative excess mortality rates of the 13 European countries compared. The EuroMOMO data for Berlin suggest a dramatic contrast with London.

¹⁴ See Bennhold (2020): "A German exception?", The New York Times, 4 April, revised 6 May 2020.

¹⁵ See Stafford (2020): "Covid-19: Why Germany's case fatality rate seems so low", British Medical Journal, 7 April, 2020.

¹⁶ For the U.S. record, see Lopez (2020): "Why America ran out of protective masks", Vox.com, 27 March, 2020.

¹⁷ See Levitan (2020): "The infection that's silently killing coronavirus patients", The New York Times, 20 April, 2020.

Insufficient attention to warnings from critics and scientific experts

There is widespread agreement among critics of the UK government's policy response that two key failures were not investing in adequate supplies and distribution logistics of PPE¹⁸ and in testing facilities at the beginning of the crisis. The late recognition of the need to provide care homes with PPE and tests has received much recent attention. A third factor was the late and initially unclear application of social distancing and delay in lock-down measures, see Sridhar (23 March)¹⁹ and Horton (28 March)²⁰ and the comprehensive account by Guardian journalists.²¹ As early as 2 March, an alarming consensus report was sent to the government's Scientific Advisory Group for Emergencies, SAGE, by its subcommittee of modellers of epidemics.²² Perhaps the most puzzling decision of all was the abandonment of testing, tracing and isolation on 12 March to follow temporarily the false trail of counting on 'herd immunity' to slow the spread of the virus (Sridhar, 2020). The argument advanced is that the number of infections was overwhelming the capacity to test and trace. If that was the case, it should have been even more urgent to implement stringent social distancing measures, instead of which major crowd events such as the 5-day Cheltenham Festival were allowed to proceed in that week.

Anderson et al. (online, 6 March)²³ had by that stage already analysed the key factors in the dynamics of the spread of Covid-19. To quote: "What has happened in China shows that quarantine, social distancing, and isolation of infected populations *can contain the epidemic*. [Our italics.] This impact of the Covid-19 response in China is encouraging for the many countries where Covid-19 is beginning to spread. However, it is unclear whether other countries can implement the stringent measures China eventually adopted. Singapore and Hong Kong, both of which had severe acute respiratory syndrome (SARS) epidemics in 2002–03, provide hope and many lessons to other countries. In both places, Covid-19 has been managed well to date, despite early cases, by early government action and through social distancing measures taken by individuals". A widely-circulated article by Pueyo (10 March) drew convincingly on an array of international evidence to call for early action to flatten the curve of the pandemic, following the example of Asian countries.²⁴ Indeed, he argued, citing evidence from China, that even *one day* of delay could cost as many as 40 percent more excess deaths. The mass gatherings in England permitted in the week of 12 March are thus likely to have cost many lives. Criticism of the herd immunity idea, allowing mass gatherings and delaying a lockdown was immediate, see Financial

¹⁸ See Foster and Neville (2020): "[How poor planning left the UK without enough PPE](#)", The Financial Times, 1 May, 2020.

¹⁹ See Sridhar (2020): "[Britain had a head start on Covid-19, but our leaders squandered it](#)", The Guardian, 23 March, 2020.

²⁰ See Horton (2020): "[COVID-19 and the NHS—"a national scandal"](#)", The Lancet, Comment, 28 March, 2020.

²¹ See Conn et al. (2020): "[Revealed: the inside story of the UK's Covid-19 crisis](#)", The Guardian, 29 April, 2020.

²² To quote: "It is highly likely that there is sustained transmission of Covid-19 in the UK at present," the report said. The coronavirus was noted to be highly contagious, with each infected person infecting two to three more. If "stringent measures" were not imposed, "it would correspond to around 80% of the population [53 million people] becoming infected", Conn et al. (2020).

²³ See Anderson et al. (2020): "[How will country-based mitigation measures influence the course of the COVID-19 epidemic?](#)" The Lancet, 9 March 2020.

²⁴ See Pueyo (2020): "[Coronavirus: Why You Must Act Now](#)" 2 March 10, revised March 19, 2020.

Times²⁵ and Guardian,²⁶ followed by a letter to the Lancet on 17 March by 36 distinguished medical and public health experts.²⁷ The notion that following a laissez-faire policy to protect the economy was the right choice was also condemned in CEPR's second e-book on 18 March: "Covid-19 economic crisis: Act fast and do whatever it takes", and in articles within, some of which had been circulated earlier (Baldwin and Weder Di Mauro, 2020). There was a rapid change of course, after the alarming implications of the pandemic modelling at Imperial College were released on 16 March by Ferguson et al. (2020). On March 16, people were told to avoid all non-essential contact with others. On March 18, the closure was announced of schools and childcare facilities, except for the children of key workers, from 23 March, and on March 20, all bars, pubs, cafes, restaurants and gyms were closed, with effect from the next day.

In a televised address to the nation on March 23, Prime Minister Boris Johnson announced unprecedented limits on where and how people can meet and gather during the continuing coronavirus. The lockdown officially started the next day.

EuroMOMO and national statistical agencies should publish improved measures of excess mortality

Forecasting P-scores from epidemiological models for different scenarios on ending lockdown measures should be an important aid to formulating policy.²⁸ Granular data by location within and between countries must be produced and made accessible for research and forecasting. An example using granular Italian death registry data is Ciminelli and Garcia-Mandicó (2020).²⁹ Belloc et al. (2020) caution against drawing simplistic conclusions from cross-country correlations; they too stress the need for granular, comparable data.

National statistical offices should publish weekly P-scores of excess mortalities for the constituent countries, regions and broad social groupings such as care-home residents, to help understand the pandemic and inform policy.³⁰ P-scores are more salient and interpretable than Z-scores. They avoid the possibility that Z-scores may under-represent excess mortality where death data are noisier.

We argue that EuroMOMO should be permitted to produce P-scores as well as Z-scores to aid comparability across countries. EuroMOMO's five-year graphs of Z-scores visualise the natural weekly variability, helping to interpret the confidence intervals. Similar practice should be followed for published P-scores, including at national statistical agencies. There is no reason why EuroMOMO could not report confidence intervals for P-scores similar to the confidence intervals it reports for Z-scores.

²⁵ See FT (2020): "[Johnson under fire as coronavirus enters dangerous phase](#)", The Financial Times 12 March, 2020.

²⁶ See Stewart and Busby, 2020): "[Coronavirus: science chief defends UK plan from criticism](#)", The Guardian, 13 March, 2020.

²⁷ See Alwan et al (2020): "[Evidence informing the UK's COVID-19 public health response must be transparent](#)", The Lancet, 17 March, 2020.

²⁸ A study which forecasts the one year ahead mortality is Denaxas et al. (2020).

²⁹ They analyse daily death registry data for over 1000 Italian municipalities, which suggest that deaths registered as Covid capture only about half of excess deaths. They find strong evidence that locations where mass testing, contact tracing, and at-home care provision was introduced experienced lower numbers of excess deaths.

³⁰ At more granular levels, the weekly data can become noisy. Averages over longer periods are more informative.

To end on a cautionary note, excess mortality should also be examined in a longer-term perspective. Spiegelhalter (2020) argues the main impact of Covid-19 may be to shift forward the date of death by a few months for those close to death because of underlying poor health. Then, a peak in weekly deaths should be followed by a trough, see Table 2 for hints that this may be occurring in some countries. However, total years of life lost is a better measure of the pandemic's social toll. Even in the extreme case envisaged by Spiegelhalter, if the 12-month moving average of excess mortality showed no deviation outside the -2, +2 normal range, total years of life lost could still show an upturn. If national statistical agencies regularly published monthly, 3-month, 6-month and 12-month moving averages, and weekly P-scores, this would greatly assist our ability to interpret the pandemic data.³¹

A lucid analysis by Gourinchas (2020), online on 15 March, of the benefits of flattening the Covid-19 curve to bring it within the capacity of the health system to cope, to give time to build up that capacity and to develop medical treatments to improve health outcomes, has been widely accepted. It has been followed by much recent work on how to end lockdowns and develop a roadmap to pandemic resilience (e.g. Favero et al. (2020) and Edmond J. Safra Center (2020)). Policymakers almost everywhere are struggling with these issues, and in some countries, cautiously beginning to feel their way. Provision of timely, regularly updated and comparable granular data on excess mortality by national and international statistical agencies should be high on the agenda. It is not enough to leave this to hard-working journalists.

Acknowledgements

We are grateful to Eric Beinhocker, Gerry Kennally, Max Roser and David Vines for comments. An abbreviated version of this article appeared on VoxEU, see [Aron and Muellbauer \(2020\)](#).

References

- Anderson, R. M., H. Heesterbeek, D. Klinkenberg and T. D. Hollingsworth. 2020. "How will country-based mitigation measures influence the course of the Covid-19 epidemic?" *The Lancet* 395 (10228): 931-934, 21 March.
- Alwan, N. et al. 2020. "[Evidence informing the UK's COVID-19 public health response must be transparent.](#)" *The Lancet*, 17 March, 2020.
- Aron, J. and J. Muellbauer. 2020. "[Measuring excess mortality: England is the European outlier in the Covid-19 pandemic.](#)" VOXEU, Centre for Economic Policy Research, London, 18 May, 2020.
- Baldwin, R. and B. Weder di Mauro (Eds.) 2020. [Mitigating the Covid Economic Crisis: Act Fast and Do Whatever It Takes](#), a VoxEU.org eBook, CEPR Press, Centre for Economic Policy Research, London, 18 March 2020.
- Barr, C., N. Kommenda, N. McIntyre and A. Voce. 2020. "[Ethnic minorities dying of Covid-19 at higher rate.](#)" *The Guardian*, 22 April, 2020.
- Belloc, M., P. Buonanno, F. Drago, R. Galbiati and P. Pinotti. 2020. "Cross-country correlation analysis for research on Covid-19." VOXEU, Centre for Economic Policy Research, London, 28 March 2020.

³¹ In the case of the ONS, the latest update of their research plans, dated 21 April, 2020, suggests only a limited agenda for investigating excess deaths: see the webpage, [Statement of upcoming analysis on deaths and coronavirus \(COVID-19\)](#).

Bennhold, K. 2020. "[A German exception?](#)", The New York Times, 4 April, revised 6 May 2020.

Burn-Murdoch, J., V. Romei and C. Giles. 2020. "[Global coronavirus death toll could be 60% higher than reported.](#)" Financial Times, 26 April, 2020.

Ciminelli, G. and S. Garcia-Mandicó. 2020. "[COVID-19 in Italy: An analysis of death registry data.](#)" VOXEU, Centre for Economic Policy Research, London, 22 April 2020.

Conn, D., F. Lawrence, P. Lewis, S. Carrell, D. Pegg, H. Davies and R. Evans. 2020. "[Revealed: the inside story of the UK's Covid-19 crisis.](#)" The Guardian, 29 April, 2020.

Denaxas, S., H. Hemingway, L. Shallcross, M. Noursadeghi, B. Williams, D. Pillay, L. Pasea, A. González-Izquierdo, C. Pagel, S. Harris, A. Torralbo, C. Langenberg, W. Wong, and A. Banerjee. 2020. "[Estimating excess 1- year mortality from COVID-19 according to underlying conditions and age in England: a rapid analysis using NHS health records in 3.8 million adults.](#)" 10.13140/RG.2.2.36151.27047.

Edmond J. Safra Center. 2020. "[Roadmap to Pandemic Resilience: Massive Scale Testing, Tracing, and Supported Isolation \(TTSI\) as the Path to Pandemic Resilience for a Free Society.](#)" The Edmond J. Safra Center for Ethics.

Farrington, C.P., N.J Andrews, A.D. Beale and M.A. Catchpole. 1996. "A statistical algorithm for the early detection of outbreaks of infectious disease." Journal of the Royal Statistical Society A 159: 547-563.

Favero, C., A. Ichino, and A. Rustichini. 2020. "Restarting the economy while saving lives under Covid-19." DP14664, Centre for Economic Policy Research, London, April.

Ferguson et al. (2020): "[Report 9: Impact of non-pharmaceutical interventions \(NPIs\) to reduce COVID-19 mortality and healthcare demand.](#)" Imperial College COVID-19 Response Team, 16 March, 2020.

Foster, P. and S. Neville. 2020. "[How poor planning left the UK without enough PPE.](#)" The Financial Times, 1 May, 2020.

FT. 2020. "[Johnson under fire as coronavirus enters dangerous phase.](#)" The Financial Times, 12 March, 2020.

Gourinchas, P. 2020. "Flattening the Pandemic and Recession Curves." In Richard Baldwin and Beatrice Weder di Mauro (Eds.) *Mitigating the Covid Economic Crisis: Act Fast and Do Whatever It Takes*, CEPR e-book, Centre for Economic Policy Research, London, 18 March 2020. See also "[Flattening the Pandemic and Recession Curves.](#)"

Horton, R. 2020. "[COVID-19 and the NHS—“a national scandal.](#)" The Lancet, Comment by editor-in-chief of The Lancet, 28 March, 2020.

Khuntj, K., A. K. Singh, M. Pareek, W. Hanif. 2020. "[Is ethnicity linked to incidence or outcomes of Covid-19?](#)", British Medical Journal, 20 April , 2020.

Levitan, R. 2020. "[The infection that’s silently killing coronavirus patients.](#)" The New York Times, 20 April, 2020.

Lopez, G. 2020. "[Why America ran out of protective masks.](#)" Vox.com, 27 March, 2020.

Pueyo, T. 2020. "[Coronavirus: Why You Must Act Now.](#)" Blog, 10 March, revised 19 March, 2020.

Spiegelhalter, D. 2020. "[How much ‘normal’ risk does Covid represent?](#)" Winton Centre for Risk and Evidence Communication, Cambridge, 21 March, 2020.

Sridhar, D. 2020. "[Britain had a head start on Covid-19, but our leaders squandered it.](#)" The Guardian, 23 March 2020.

Stafford, N. 2020 "[Covid-19: Why Germany’s case fatality rate seems so low.](#)" British Medical Journal 369:m1395, 7 April, 2020.

Stewart, H. and M. Busby. 2020. "[Coronavirus: science chief defends UK plan from criticism.](#)" The Guardian, 13 March 2020.

The Economist (2020): "[Tracking Covid-19 excess deaths across countries.](#)" The Economist, 16 April, 2020,

Tozer, J. (2020). "[Measuring the true toll of the pandemic.](#)" The Economist, 24 April, 2020.

Wu, J., A. McCann, J. Katz and E. Peltier. 2020. "[46,000 Missing Deaths: Tracking the True Toll of the Coronavirus Outbreak.](#)" The New York Times, 30 April, 2020.