



There are deaths, excess deaths, and statistics ...

By John Roberts and Matthew Edwards

COVID-19 Actuaries Response Group – Learn. Share. Educate. Influence.

Summary

In this bulletin we note the recent news about the unreliability of daily COVID-19 death figures in England¹, and re-state our belief that the metric of excess deaths is the most reliable way of gauging the pandemic's impact. Hence, nothing has changed – it remains the case that over 60,000 have (very likely) died from COVID-19.

We have analysed the difference between ONS and PHE-reported deaths, tracing the growing difference over time. Although the figures are different, the total erroneous reporting to date changes nothing. However, although the 'quantitative' aspect is almost immaterial, we are concerned that the 'qualitative' aspect – of reduced public trust in the numbers, or Government advice – may have unwarranted consequences.

Lies, damned lies and statistics

Many have heard the phrase 'there are lies, damned lies, and statistics', coined (perhaps) by Disraeli. Some may feel the same about the reporting of COVID-19 deaths.

The debate has exploded again, with the announcement by the Secretary of State for Health and Social Care into potential misclassification of COVID-19 as a cause of death in the daily death figures that are so widely reported. It seems that in its attempt to ensure that all COVID deaths have been included, PHE has been checking for a match with a previous positive test, and including all of those in the figures. That might have been reasonably accurate in April when "out of hospital" deaths were added, but with much greater volumes of testing, and longer elapsed times between the test and subsequent death, there will be many cases now where COVID will have been incidental or even irrelevant in terms of the cause of death.

Problems with cause of death

We like to imagine that our public health systems are reliable repositories of perfect knowledge about something so apparently obvious as to how people have died. But registering the correct cause of death has always been problematic in the very old (the main COVID-19 age band). So there is room for error in two senses: first, the problem of distinguishing one out of possibly many causes in the common old-age event of 'total systems failure' of the body. Secondly, the problem that a meaningful proportion (but we think a minority) of deaths of the very old, that may have been triggered 'now' by COVID-19, would have been deaths later in the year in any case.

And on top of that, deaths outside hospitals – so in a COVID-19 context, we're mostly talking care homes – will have a degree of subjectivity in terms of cause. The most objective measurement will still be the GP's recording of the reason on the death certificate, which will be recorded in the ONS figures.

Even here, there is likely to be more uncertainty, as GP access to end of life patients will have been limited, but it is going to be better than a rather crude match against a positive COVID test that may have been taken several weeks earlier.

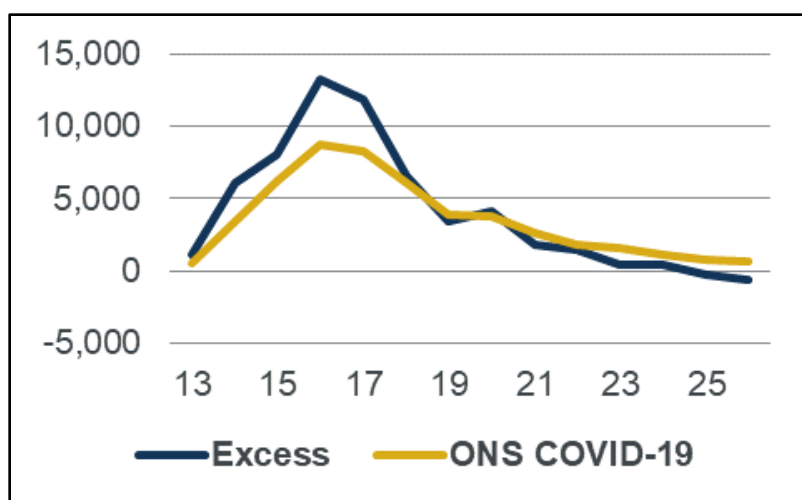
So, causes have always been unreliable. But how can we gauge the pandemic's deaths if we can't rely on the official death registration process?

Excess death reporting

We can obtain a very accurate (but not perfect) understanding of deaths attributable to the pandemic by looking at 'excess deaths', and this is something that various bodies have been doing for many months. At its simplest, we could calculate total deaths this year, and compare them with what we would expect – for instance, look at last year's over the same period. In the absence of any other known public health disasters this year, or known reasons why last year might have been 'light' (and barring a bit of statistical 'noise'), this difference will constitute deaths attributable to COVID-19.

That sounds too simple and it is – in reality, we should adjust for the slightly different age/gender mix of the population now compared with last year, and we could also allow for typical yearly mortality improvements (as we would expect age-standardised deaths to be lower this year than last year, all other things being equal). We could also do the comparison with a starting point in March, to avoid any meaningless (in this context) differences in January and February mortality.

The CMI (part of the Institute and Faculty of Actuaries) has been publishing an 'excess deaths' calculation of this type since the start of the pandemic. This shows significantly higher excess deaths than the official COVID-19 death figures. The most recent publication, for instance, showed total excess deaths to 26 June of around 62,500² (eg graph below). During the peak of deaths in the UK excess deaths were much higher than official figures; in recent weeks they have been slightly lower, even turning negative.



As the year develops, this 'excess death' reporting can then also pick up a point noted above, that some of the deaths seen so far would have happened later in the year in any case. So if the current 'excess' figure starts to decline, that decline would relate to that point of COVID-19 having 'accelerated' some deaths. Indeed, the most recent publication reduced the excess deaths by 1,000 for just that reason.

Overall, then, what has changed from this news about mis-reporting? From our perspective, nothing. The registered death numbers have never been particularly useful – indeed we cautioned against reliance on the PHE daily numbers in one of our first bulletins³ back in April. Excess death numbers are the most reliable indicator.

“Excess deaths” is not a perfect measure. For instance, we may have had some cancer deaths recently that might not have happened if all the normal cancer screening and treatment processes were working normally (which they are not). These deaths would be implied by the ‘excess’ figures to be attributable to the pandemic – which they are in a sense, but not to COVID-19 directly so much as the reshaping of our world to counter the pandemic.

COVID-19 deaths: are any of the numbers reliable?

It’s been some time now that the daily reported figures have looked odd to those of us keeping track of the data. How could hospital deaths be falling so much quicker than other deaths, and if that was true, why weren’t more people being admitted to hospital for treatment, given the NHS clearly had the capacity?

Let’s remind ourselves what makes up the published figures. There are two sources of data, the NHS deaths data (published like clockwork by 2 pm), and then the additional “non-hospital” deaths added by PHE to give the total, published later in the day. Both these sources include deaths going back several days and sometimes weeks, though that aspect of the data is well understood, not unreasonable, and not under scrutiny. Neither are the NHS deaths at issue.

But it wasn’t always like that. In the early days of the pandemic, only the hospital deaths were reported, until the extent of deaths in care homes became apparent and the government came under intense pressure to include all deaths from COVID-19. So in late April, PHE was tasked with a method of including other deaths, and that’s where the problem seems to have arisen.

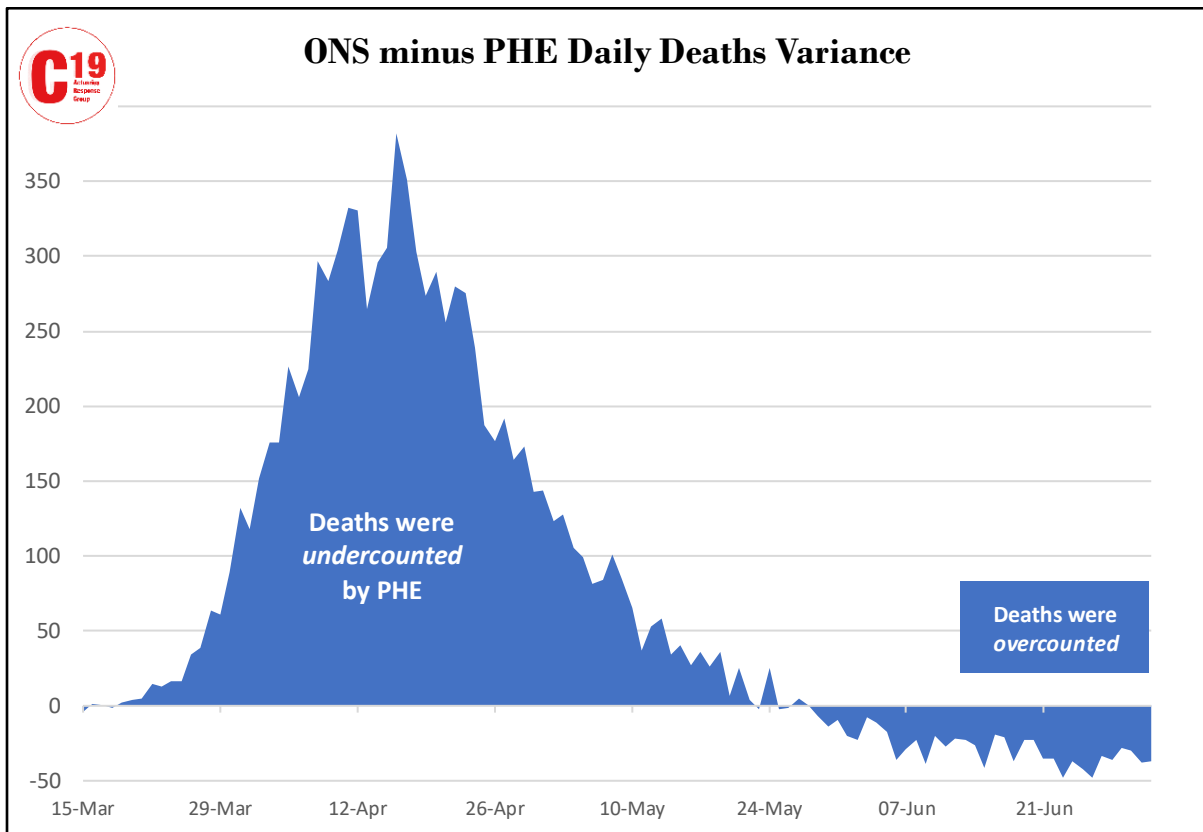
In hindsight, the method PHE used was reasonable as a quick fix. It compared deaths with those where there was a positive test. Almost by definition, given the (then) low volumes of testing and the relatively short duration of the pandemic, any such deaths had a very high likelihood of being caused by COVID-19, either directly or indirectly.

In fact, any criticism would have been that it still didn’t catch all deaths “due to COVID-19”, as many deaths would still have been for people who had never been tested, whereas the death certificate mentioned COVID-19 as a cause. The data shows that to be the case.

But as time has gone on, testing has become much more extensive, and, well, time has gone on, meaning that the elapsed time between a test and death can now be much greater. So the possibility of a death being registered in England where COVID-19 was not recorded as a reason, but a positive test had been obtained months earlier becomes much greater. Scotland, Wales and Northern Ireland only attribute the death to COVID-19 if it happens within 28 days of the positive test, for precisely this reason.

It’s this misreporting which is now becoming material and has prompted the investigation. PHE’s mistake was not to review the methodology over time and recognise that its limitations would become much more relevant as time and the pandemic progressed.

The following graph shows the situation, showing the variance between ONS and PHE daily death figures. A positive (above the axis) shows undercounting by PHE – the total area of this section completely overshadows the negative area (below the axis) which shows the recent overcounting.



Comparing the ONS data with the daily figures gives the pattern that we might broadly expect. In the early weeks the daily figures significantly underreported deaths, even once PHE had started to include deaths in other settings (which was backdated). By the middle of May this shortfall was around 9,000 deaths in total, and is a large component of the difference between reported deaths and the excess deaths that the CMI has been reporting in recent weeks.

The over-reporting of deaths has only started to become material from the beginning of June, and to date just under 1,000 deaths have been erroneously reported in the daily figures. In the last couple of weeks that ONS data is available for comparison the error has been around 250 a week – up to a third of the total reported.

Does it matter?

So does this matter? After all, the ONS publishes accurate data, albeit in arrears, and as we have consistently stated, excess deaths is a much better measure of the overall impact of the pandemic, not only covering COVID deaths, but others, such as those caused by the inability (perceived or real) to access timely healthcare for other conditions.

We think it does matter. Any issue around the credibility of figures reported can affect public confidence in whether the government has a firm handle on the situation. It can also enable those with “an agenda” to use the confusion as proof as to why their view of the pandemic is more convincing. With the importance of social media in influencing behaviour this shouldn’t be underestimated given the need to persuade the public to continue to take various COVID measures that may be perceived as restricting basic freedoms.

The other behavioural aspect of overstating the reported figures is more direct. It is clear that the error has meant the daily reported figures have not fallen as steeply as they should have done. The majority of the population will only see the daily figures and may have been encouraged to comply better with government guidelines as they remain concerned at the number of deaths, a positive, albeit unwitting impact.

Conversely, at a time when the government is now looking for the economy to open up and for people to “enjoy summer safely”, news coverage highlighting uncertainties in the situation might discourage people from venturing out and bring life back into the hospitality sector, thus delaying the recovery further. And as previously noted, individuals requiring urgent healthcare may be denying themselves the treatment they need due to a fear of hospitals or perception that it is not available.

So in summary, the daily reported deaths have been overstated in recent weeks, and this is regrettable. However, overall deaths reported through the same mechanism still remain substantially understated, and this point should not be lost when considering the mis-statement. Excess deaths remains the best measure of the overall impact of the pandemic, and the best estimate of this remains above 60,000 currently.

19 July 2020

¹ <https://www.bbc.co.uk/news/health-53443724>

² <https://www.actuaries.org.uk/learn-and-develop/continuous-mortality-investigation/other-cmi-outputs/mortality-monitor> - note the graph shown above is © Continuous Mortality Investigation Limited from <https://www.actuaries.org.uk/system/files/field/document/Mortality-monitor-Week-26-2020-v01-2020-07-07.pdf> and subject to the disclaimer on page 9 of that report

³ <https://www.covid-arg.com/bulletins> - “Interpreting the death data”