



To handle Covid-19 we need better data

Winton's David Harding says unconscious bias in test data means epidemic is poorly understood

David Harding

27 Mar 2020

Until very recently we would probably never have known of the existence of a virus such as the novel coronavirus. The miracles of modern science have brought us the ability to diagnose and to track such a disease. That knowledge, though, can be a blessing and a curse.

We have up-to-the-minute data on cases and deaths, from multiple countries around the world. Yet, for the responsible authorities, there are pitfalls in the interpretation of this rich but fast-unfolding picture.

At present the true growth trajectories of the multiple epidemics are not well enough understood. The growth in the number of reported cases from day to day does not measure these trajectories because the data is muddied by the growth in the number of tests and by changes in who is being tested.

Simple comparisons of diagnoses and death rates across countries are impossible due to different medical conventions and unconscious biases introduced in the data gathering and recording process. These biases have the potential to be massively distorting.

Coronavirus probably began in Wuhan in November 2019. By the end of December, it had infected enough people there for the cluster of abnormal pneumonia cases it caused to be noticed. By early January it was identified, and began to be formally diagnosed. Tests were devised which identified many cases of the new disease in the population. By late January the Chinese authorities enacted drastic measures to isolate and lock down Wuhan and several other cities in Hubei province.

Over the course of February it became apparent the infection had spread everywhere and countries scrambled to start testing for and isolating those

infected.

Initially, testing people with a connection to the perceived source of the outbreak revealed quite a few coronavirus carriers. However, the disease had invisibly established a foothold in other locations and cases were already quite widespread. For some time the number of positive tests in each country was mistaken for the number of cases. Evidence of unseen cases slowly emerged through the detection of community transmission, and it became clear there was a much larger population of infected individuals than previously appreciated.



It is important – while being aware of the rapid spread of the disease – not to overestimate the epidemic’s spread and to overreact

David Harding

In response, governments conducted more tests. And, while the largest-scale testing possible is desirable in assessing the best actions to take, more tests inevitably mean more cases in the short term. There exists a substantial and rapidly growing reservoir of infected individuals waiting to be discovered.

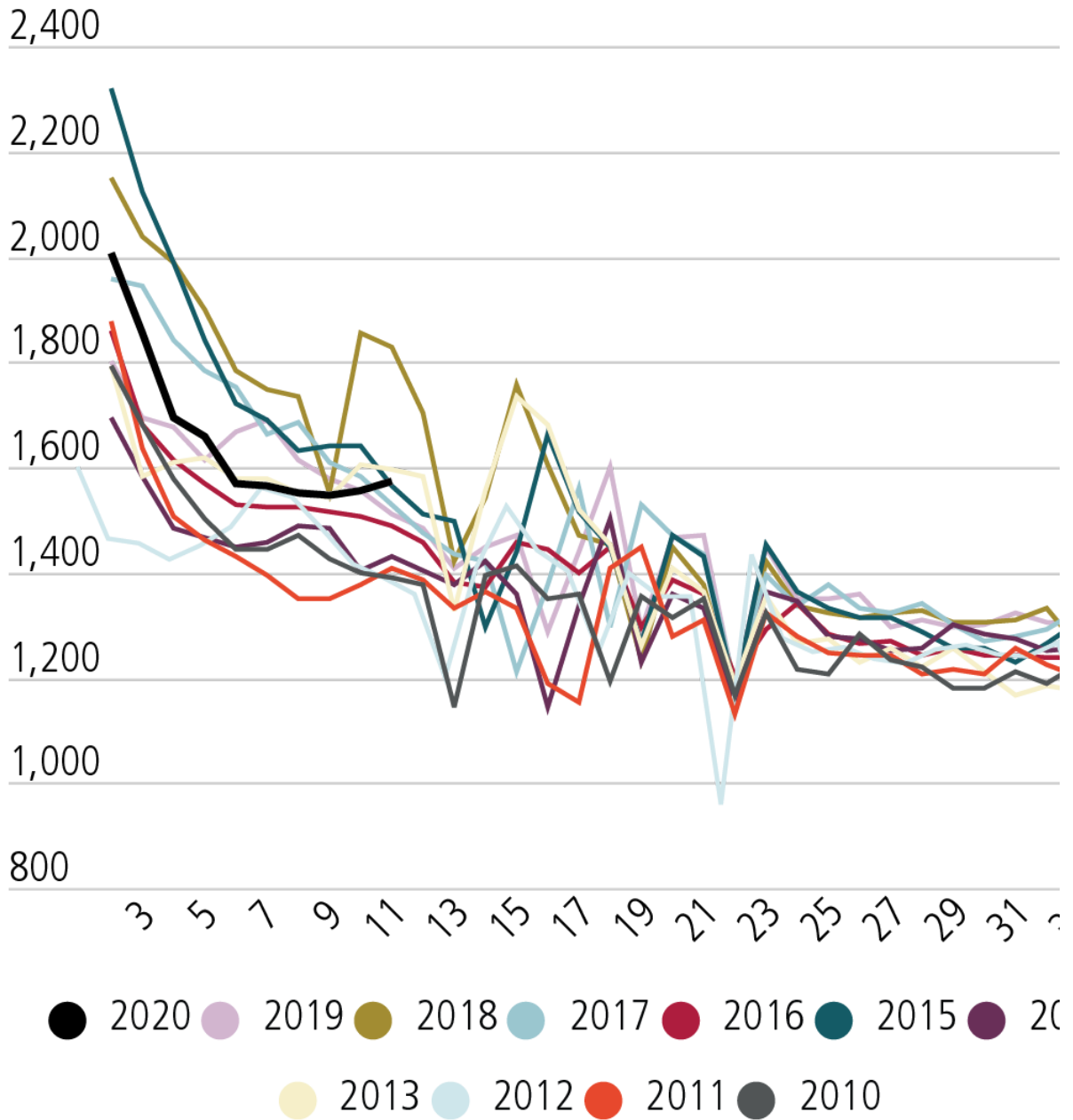
It is important – while being aware of the rapid spread of the disease – not to overestimate the epidemic’s spread and to overreact, spreading fear and panic.

The data can lead to false conclusions in a number of ways. Already there have been several published articles and non-peer-reviewed preprints that have received lots of media attention only to be quickly retracted, as Stanford professor John Ioannidis has [pointed out](#).

Taken out of context, headline news about absolute numbers of deaths can be misleading. Each year many old people succumb to respiratory diseases – often pneumonia – which result from a weak immune system encountering cold or flu.

In England and Wales, there are about 1,700 deaths per day in winter compared with 1,300 in summer. As a first approximation, we could attribute these additional 400 deaths per day to winter diseases. During one of the more severe flu epidemics the figure could double to 800.

Daily number of England and Wales deaths by week



Sources

Share

One could anticipate that a very bad winter disease season, over the course of five to 10 weeks, might produce an additional 10,000 to 20,000 deaths in England and Wales, and many hundreds of thousands or millions of deaths worldwide.

During a rapidly expanding epidemic of a new, even relatively mild disease, many deaths in this large cohort group are likely to be found in which the coronavirus will have been a contributing factor. Publishing real-time counts of these numbers and using them to drive policy seems risky.

As Joshua Niferatos and others have [shown](#), it is difficult to estimate the death rate early in an epidemic due to limitations and inconsistencies in testing. Differences in how deaths are recorded can further muddy the picture. As an example, according to Chinese medical convention, most of these deaths will be recorded as being due to an underlying condition and not as a result of coronavirus. In Western countries deaths are more likely to be attributed to the virus, making the death tolls seem ever more alarming.

To accurately inform major public policy decisions, it is critically important that all relevant public data is collected and organised. Estonia could be a model for other countries to emulate. Its government has set up drive-in testing sites and is collecting anonymous data on its citizens' movements. The country is also already conducting random sample testing of workers in vital services as well as those aged 60 or over and has pledged to extend this further.

“ To accurately inform major public policy decisions, it is critically important that all relevant public data is collected and organised

David Harding

There is already evidence that the disease can be mild or even [asymptomatic](#) in many people and thus unbiased estimation of its current prevalence requires fairly large testing of essentially random samples.

This will reveal the true, low – though rapidly increasing – prevalence in the population. The random testing then needs to be repetitive to gain an unbiased view of the trajectory of case numbers. By contrast, the present approach of rapidly increasing the number of tests from a large pool of people with a high prior probability of having the disease will likely, in the short run, exaggerate the growth rate by discovering quickly a lot of infection that is already there.

There is a current among policy-makers and their advisers globally to act aggressively, assuming the disease will rapidly move beyond our ability to control it and become impossible for the hospital system of any nation to handle. There are also great risks, though, associated with overreaction. Those are beyond the scope of this article, but common sense tells us that reasoned analysis and debate must be the best guide to policy-making.

This is no time for wishful or politicised thinking, but for calm analysis. And the raw material of such analysis must be good data.

David Harding is the founder, CEO and co-CIO of Winton Group, an investment management firm. Donations from David established the Winton Professorship of the Public Understanding of Risk and the Winton Centre for Risk and Evidence Communication, both at Cambridge University. He is also the founding donor of the Harding Center for Risk Literacy at the Max Planck Institute for Human Development.

The logo for Risk.net Investing, featuring the text "Risk.net" in a bold, blue, sans-serif font above the word "INVESTING" in a smaller, blue, all-caps, sans-serif font. The background of the logo area is a light green with horizontal lines.

Have you read our latest Investing content?

Focused on the investment management community, you will find content from a new editorial team on the Investing desk page.

Coverage is concentrated across five areas: risk management; regulation; derivatives; markets; and quantitative investing, reporting on key topics for the front to the back office.

[VISIT OUR INVESTING PAGE](#)

Copyright Infopro Digital Limited. All rights reserved.

You may share this content using our article tools. Printing this content is for the sole use of the Authorised User (named subscriber), as outlined in our terms and conditions - <https://www.infopro-insight.com/terms-conditions/insight-subscriptions/>

If you would like to purchase additional rights please email info@risk.net