

Professor Sir David Spiegelhalter FRS OBE
Chair, Winton Centre for Risk and Evidence Communication
Centre for Mathematical Sciences, Cambridge

William Wragg MP
Chair, Public Administration and Constitutional Affairs Committee
House of Commons

Dear Mr Wragg

Thank you for letter following my appearance before the Committee on 24th November.
Your specific requests were:

- *We discussed statistics used by Government to justify decisions where the original data was not produced for that purpose. Could you provide us with some examples of where that has been the case?*
- *Could you provide a few examples of particularly bad presentations of data by Government, such as misused statistics or graphs which are hard to interpret?*

I shall give a bit of background and then answer your specific requests.

1. I am Chair of Winton Centre for Risk and Evidence Communication at the University of Cambridge, Co-chair of the Royal Statistical Society Covid-19 Task-force, and a non-executive director of the UK Statistics Authority, which oversees the Office for National Statistics and the Office for Statistics Regulation. I was President of the Royal Statistical Society for 2017-2018, and Professor for the Public Understanding of Risk at the University of Cambridge from 2007-2018.
2. However, this contribution is being made in a personal capacity, and not as a representative of any of these organisations.
3. There has been some very good data communication by professionals, and I would single out ONS and the PHE dashboard with its API.

4. Philosopher Baroness Onora O'Neill has laid down the basic principles for 'intelligent transparency' around information¹: it should be
 - *accessible* to audiences,
 - *intelligible* to as many as possible,
 - *useable* to answer their concerns,
 - *assessable* by those who wish to check your working.
5. It is clear that Test and Trace, JBC and PHE are using the best available data to make recommendations as to tiering etc, and are also making increasing attempts at making this data available as part of a transparency agenda. For example, it is admirable that the detailed Slide pack for tier decisions announced on November 26th was made [available](https://www.gov.uk/government/publications/slides-to-accompany-coronavirus-press-conference-31-october-2020)² with data-sets. But it is not very accessible: it is a 79-page pdf, and is remarkably difficult to find on the web.
6. **Given the billions of pounds being spent on Test and Trace, it seems remarkable that funds have not been found to provide a better public-face for the excellent analytic work being done behind the scenes, to properly make available the data that is leading to the decisions that affect all our lives and require such sacrifice from the public.**

We discussed statistics used by Government to justify decisions where the original data was not produced for that purpose. Could you provide us with some examples of where that has been the case?

The communication at the October 31st briefing announcing the second lockdown was particularly poor³. They were clearly rushed into it due to a leak the previous day, and the projection data of 'up to 4000 deaths a day' was subsequently widely ridiculed (Figure 1).

¹ <https://royalsociety.org/topics-policy/projects/science-public-enterprise/report/>

² <https://www.gov.uk/government/publications/coronavirus-cases-in-england-26-november-2020>

³ <https://www.gov.uk/government/publications/slides-to-accompany-coronavirus-press-conference-31-october-2020>

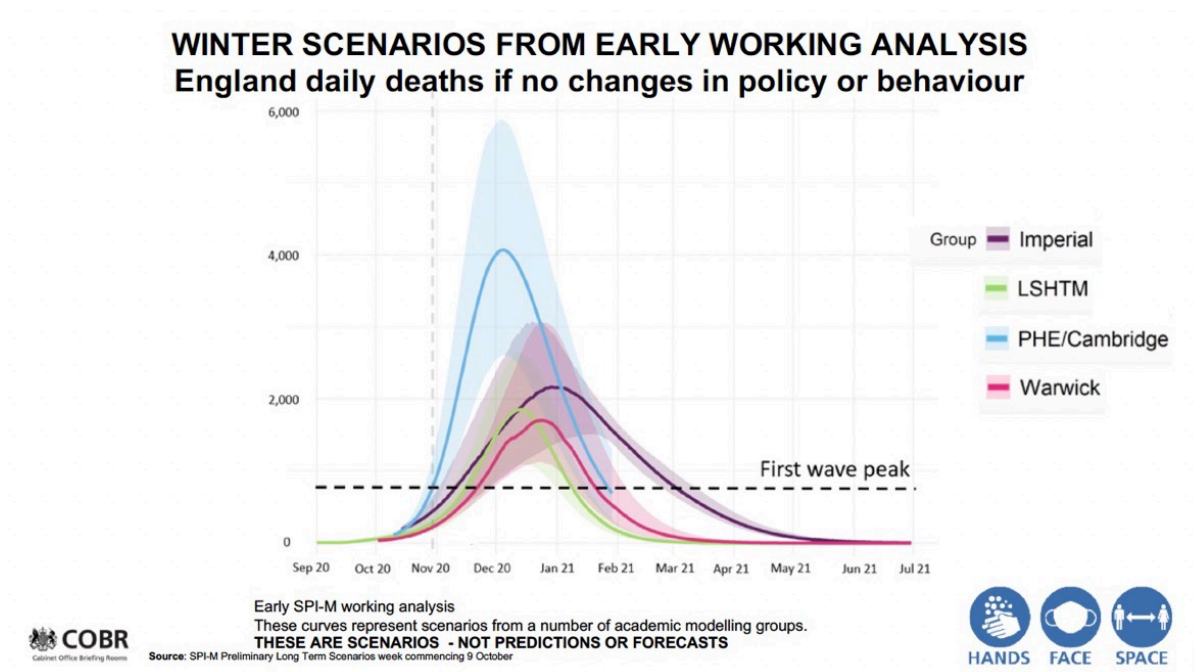
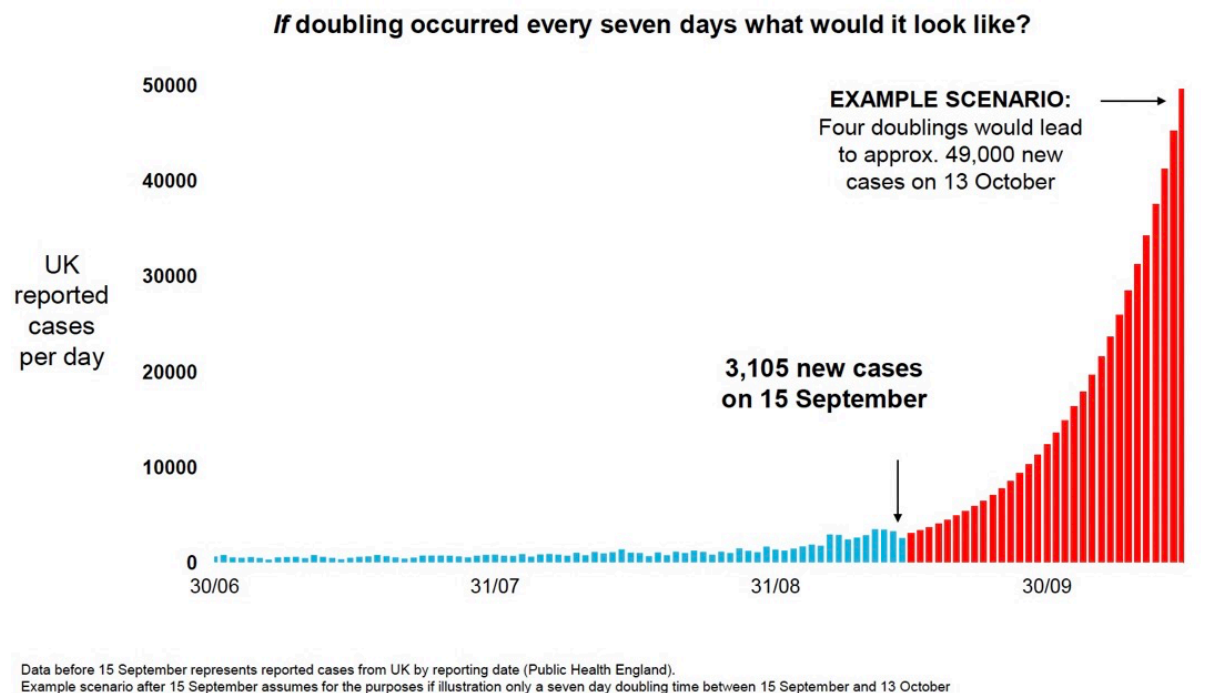


Fig 1. Graph shown at briefing on October 31st

7. This graph (Figure 1) was never intended for public consumption, was based on extreme assumptions, and was demonstrably out-of-date at the time it was used.
8. This reflects a general problem with using reasonable worst-case scenarios (RWCS) for public consumption. This has happened at three important occasions: first in early March when the '500,000' deaths was highlighted, second on September 21st ⁴ when the red-bars showing a projection of cases doubling every week, reaching 49,000 by 13th October: in fact 13,000 new cases were reported that day (figure 2).

⁴ <https://www.gov.uk/government/publications/slides-to-accompany-coronavirus-data-briefing-21-september-2020>



9. I don't want to ascribe motivations to anyone, but if someone were trying to manipulate emotions and wanting to frighten rather than inform, then this is the kind of thing they might do. All these projections were very extreme, and inevitably misinterpreted as predictions. **Scenarios are an extremely valuable way of envisaging possible futures, but worst-case scenarios should not be used selectively in communication. If modelled scenarios are to be shown, there should be a 'fan of possibilities' under different assumptions.**

Could you provide a few examples of particularly bad presentations of data by Government, such as misused statistics or graphs which are hard to interpret?

10. I have publicly complained about the 'number theatre' of briefings. The poor quality of many of the slides shown at past briefings became notorious – too crowded, lots of coloured lines (apparently ignoring guidance on accessibility for those with visual restrictions) with a legend that could not be read, and so on. Things have improved, but there is still some very poor practice which appear more like public relations than genuinely trying to inform the public. For example, the vital and detailed web-page on the government [Winter Plan](https://www.gov.uk/government/publications/covid-19-winter-plan/covid-19-winter-plan)⁵ features the following graph as their Figure 1 (my fig 3).

⁵ <https://www.gov.uk/government/publications/covid-19-winter-plan/covid-19-winter-plan>

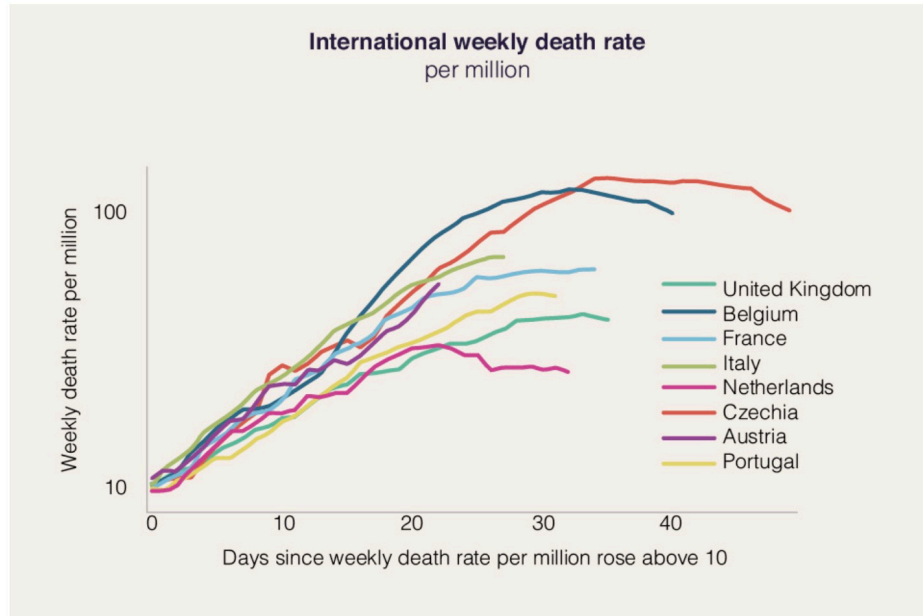


Fig 3. Graph shown on Government Winter Plan website

11. This is truly awful. A tangle of lines going upwards, no idea why these countries were selected (where is Germany?), extremely difficult to match country with line, no proper discussion in the text, and worst of all no idea whether the vertical scale is linear or logarithmic! My colleague Kevin McConway has [reproduced this graph in Excel](#)⁶, and shown that it was a logarithmic scale - if the more appropriate linear scale is used (below) it gives a very different impression, with the UK lying far lower than most of the other countries.

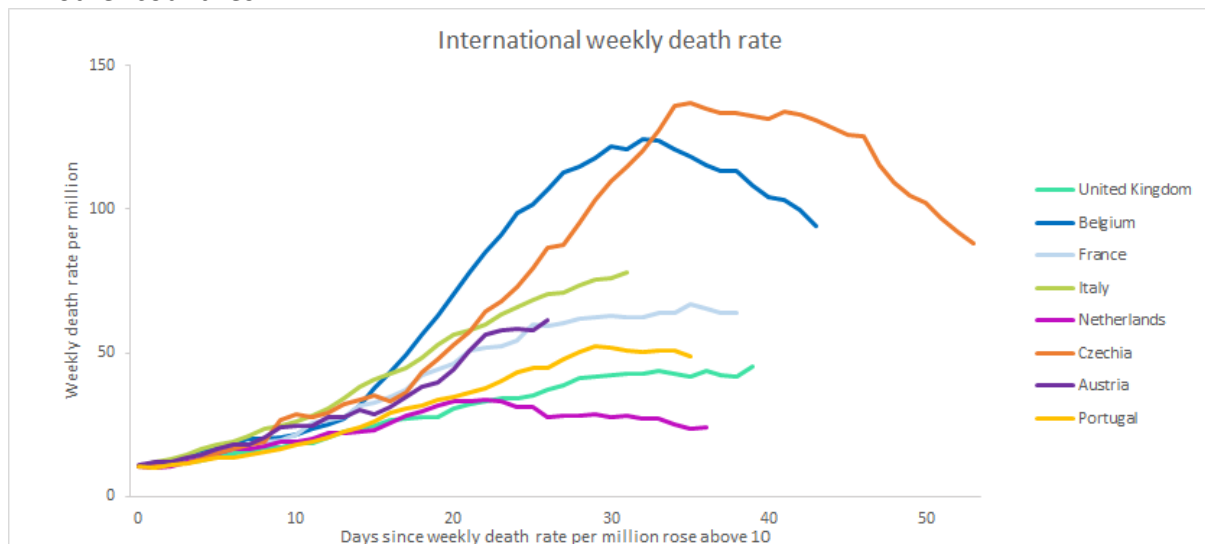


Fig 4. Redrawing of Figure 3 on a linear scale, giving a very different impression.

12. If there had been a spelling or grammatical error in the web-page, there would have been a hasty and embarrassed change. But this appalling and misleading graph, which would fail any test given to a secondary school student, seems to be considered as just

⁶ <https://twitter.com/kjm2/status/1331197478750785548?s=20>

something colourful to give a vague impression of things getting worse in lots of places. There are so many skilled analysts and data-visualisers in government, and yet this piece of embarrassingly amateur Excel is used in a vital document.

13. While the above example is perhaps not of vital importance, it illustrates an attitude in which data is used as a rhetorical addition to emphasise an argument, rather than genuinely trying to inform the public.
14. I would like to finish with an appeal for trustworthy communication. Baroness Onora O'Neill has emphasised the need for organisations to demonstrate trustworthiness, and this is the first pillar of the official [Code of Practice for Statistics](https://code.statisticsauthority.gov.uk/)⁷. Our team has written a recent [Commentary in Nature](https://www.nature.com/articles/d41586-020-03189-1)⁸, in which we argue for five principles for trustworthy evidence communication.
 - Inform, not persuade
 - Offer balance but not false balance
 - Disclose uncertainties
 - State evidence quality
 - Pre-empt misinformation**I appeal to government to take these issues of trustworthy communication seriously when communicating data to the public.**

⁷ <https://code.statisticsauthority.gov.uk/>

⁸ https://www.nature.com/articles/d41586-020-03189-1?utm_source=facebook&utm_medium=social&utm_content=organic&utm_campaign=NGMT_USG_JC01_GL_Nature