

**Written evidence from Marie Oldfield (CStat, CSci), Joel Smalley (MBA), Murray
McMonies BEng PgD MSc IEng (DTA 13)**

**Public Administration and Constitutional Affairs Committee
Data Transparency and Accountability: Covid 19**

This is an interdisciplinary paper that addresses the data issues but also the fundamental problems that sit behind the questions being asked here. Therefore, we have dug deeper than simple data questions as data is affected by many issues.

1. Did Government have good enough data to make decisions in response to Coronavirus, and how quickly were Government able to gather new data?

No, as referenced in Lords Select Enquiry of June 2020¹, there were serious issues with data in the UK and modelling had not used modelling or data from abroad within their own modelling, thereby operating in vacuum without considering current and past events. Some UK data could not be obtained at all and other data was not even considered.

“The data management systems we have in place through the NHS and NHS Scotland, which I am more familiar with, are frankly very cumbersome. It has been difficult to extract the right data at the right time for the right person in the right place. There is a lot of difficulty there. That is a historical problem; I have been complaining about it for about 10 years”².

There was sufficient data out of Wuhan and Italy on demographic and comorbidity risk at the end of February to design a suitable heterogeneous response that would have been significantly better on balance than the draconian, homogeneous response. All the data since has not contradicted this. The area in which there was little existing data, was the potential and inevitable impacts of any countermeasures implemented in response to the COVID-19 outbreak. Due to the unprecedented nature of the occurrence, there is little data available to inform what the results of a lockdown and social distancing etc would be. Despite this these measures were implemented. This is a non-evidence based policy. Furthermore, it would also be extremely difficult to predict the economic and social impacts throughout all levels of society, considering the extent of the unprecedented actions, in order to fully generate a risk balance case in the generation of the countermeasures. Although the extent of the countermeasures may have been unprecedented, the impacts of ill-communicated and implemented measures are not. Events such as the nuclear emergencies of Chernobyl 1986, Tokiamura 1999 and Fukushima 2011 are clear evidence of countermeasures implemented across a wide spread of society where the health and social impacts of the countermeasures are significantly greater than the direct impacts of the initiating event. We have seen insufficient analysis in this regard and therefore, a serious lack of a risk balance case.

2. Was data for decision making sufficiently joined up across Departments?

¹ Lords Select Committee on Science and Technology Afternoon Session – Corrected oral evidence: The Science of COVID-19 (London, 2 June 2020)

² Lords Select Committee on Science and Technology Afternoon Session – Corrected oral evidence: The Science of COVID-19 (London, 2 June 2020)

No. Data was fragmented across multifarious reporting organisations - PHE/NHS, ICNARC, ONS but apparently, largely ignored anyway in favour of singular focus on the R-number. Even much later into the event, there is conflicting comprehension of data and decision-making. For example, both Boris Johnson and Dominic Raab made explicit reference to false positive rates (FPR) of PCR tests being above 90% as the reason why travel quarantine is required, whereas Matt Hancock has shown a complete misunderstanding of the FPR and its impact on the data. Homogenous and widespread countermeasures can have a significant impact on the interpretation of the risk and the implementation of the measures at departmental levels and lower. Decisions to implement 'lockdowns' and strict measures in general society can be read across to implement effective closures or substantial reduction in operations for local health organisations, military establishments, etc, where the full risk balance has not been established. "The data management systems we have in place through the NHS and NHS Scotland, which I am more familiar with, are frankly very cumbersome. It has been difficult to extract the right data at the right time for the right person in the right place. There is a lot of difficulty there. That is a historical problem; I have been complaining about it for about 10 years" ³.

3. Was relevant data disseminated to key decision-makers in: Central and Local Government; other public services (like schools); businesses; and interested members of the public?

No and in fact when asked, in a freedom of information request, for evidence on mask wearing DFT actually supplied papers in reference that do not support mask wearing. Additionally Authors have added clarification notes to their submissions to Journals due to the publicity. These articles have been misused by the DFT. Examples of quotes from the DFT referenced papers: "The study found that cloth mask wearers had higher rates of infection than even the standard practice control group of health workers, and the filtration provided by cloth masks was poor compared to surgical masks."⁴. "This study is the first RCT of cloth masks, and the results caution against the use of cloth masks. This is an important finding to inform occupational health and safety. Moisture retention, reuse of cloth masks and poor filtration may result in increased risk of infection"⁵. "Out of 876 participants, only 27 people (3.1%) did not complain of any problems related to face mask wearing. Out of all reported inconveniences, difficulty in breathing appeared to be the most common one (35.9%)⁶. The Government Department responsible for Freedom of Information actually stated that a campaign of disruption was going on with over 100 people asking for the evidence. This is a bold assumption considering the impact that this clearly unsupported policy had on the public.

No - As per Lords Committee⁷, reliance on the R measure was inappropriate and the actual measures relating to risk were not used.

³ Lords Select Committee on Science and Technology Afternoon Session – Corrected oral evidence: The Science of COVID-19 (London, 2 June 2020)

⁴<https://bmjopen.bmj.com/content/5/4/e006577.responses#COVID-19-shortages-of-masks-and-the-use-of-cloth-masks-as-a-last-resort>

⁵ <https://bmjopen.bmj.com/content/5/4/e006577>

⁶ <https://onlinelibrary.wiley.com/doi/full/10.1111/dth.13567>

⁷ Lords Select Committee on Science and Technology Afternoon Session – Corrected oral evidence: The Science of COVID-19 (London, 2 June 2020)

A: "As regards the concentration on R, which appears to me to be an average over London, the Shetlands, individuals in flats, hospitals and care homes—very different groupings all pooled together into the one magic figure of R—are we getting this communication with the public right?"

B: "No, in my opinion (...) The focus on a single R, as you quite rightly summarised, has been a distraction. I have argued consistently for a long time now that to use a single measure as a metric to drive policy would be wrong. I do not think that it is being used a single measure to drive policy, but the impression is out there that this is a particularly critical number. I think it has been a distraction and has kept our eye off the ball, which is, as I keep coming back to, the population over 70. That is where the public health risks lie. R does not capture that". B: "(...) we should probably go back to old fashioned public health and think about it in terms of risk: what is the risk to an individual in this location at that time? Apart from anything else, that is very helpful to the individual concerned, allowing them to make informed choices about how they behave. I am not sure whether the R number helps an individual decide how to behave, but it certainly does not help me".⁸

Due to the misuse of data and interference with data collection as well as using inappropriate metrics, the media then reported the results of this to the public thereby exacerbating the utter misunderstanding of the situation. The policy of coercion and fear⁹ supported this use of misleading statistics and fearful advertising - not based on data, essentially propaganda. Ultimately the data and statistics that have been used here, and in particular the R number, are ineffective risk communication tools, as it poorly represents individual or societal risk factors. Any response should be driven by actual risk, which can be effectively communicated, not by trend.

4. Were key decisions (such as the “lock downs”) underpinned by good data and was data-led decision-making timely, clear and transparently presented to the public?

Absolutely not. In the modelling that was undertaken to inform policy, basic modelling concepts were ignored. The AQUA book and Laidlaw Review were nowhere to be seen, neither was robust analysis. Bad data in = bad data out. “Essentially the exercises that come down have a strict deadline to inform a particular policy decision. There is not a whole lot of room for interpretation and for addressing questions other than that which has been put down from on high and from SAGE”¹⁰. “With regard to age, there is an enormous amount of modelling activity going on around the world.(...) There has been an enormous amount of interest in trying to model alternative approaches to population-wide social distancing, but it has not yet become what you might call the mainstream modelling direction among my colleagues in the scientific community worldwide. (...) We have done lot of work on this in my own group and there is work out there, but as Paul rightly said this is not what has been fed down into SPI-M 5 by SAGE and the

⁸ Lords Select Committee on Science and Technology Afternoon Session – Corrected oral evidence: The Science of COVID-19 (London, 2 June 2020)

⁹https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/887467/25-options-for-increasing-adherence-to-social-distancing-measures-22032020.pdf

¹⁰ Lords Select Committee on Science and Technology Afternoon Session – Corrected oral evidence: The Science of COVID-19 (London, 2 June 2020)

Cabinet Office”¹¹. Driven by SAGE and not two-way dialogue this turned any output into untrustworthy output.

The Lords Select Committee for Science and Technology¹² was called in order to question modellers on the robustness of their analysis that has been used to underpin government policy in relation to COVID-19. It was discussed that there was significant disagreement on modelling aims but also a lack of ability to challenge the modelling requests and that the modellers did not use all the data available to them despite having access to it. The data used was focused on social distancing but not aspects such as care homes or non-lockdown measures. Shielding, for example, was neither examined nor used as a parameter. Neither data nor modelling from across Europe was examined. NHS data was extremely difficult for modellers to even obtain. The R measure was deemed an inappropriate average by modellers and Committee members and it was stated that risk perception would have been a more relevant measure to use, ”the impression is out there that this is a particularly critical number. I think it has been a distraction and has kept our eye off the ball, which is, as I keep coming back to, the population over 70. That is where the public health risks lie. R does not capture that”¹³. The right data were not used, neither were the right modelling parameters modelled despite modellers expressing up their chain of command that more relevant parameters should be modelled. “Where we really needed the granularity for this was not in schools. There is very little going on in schools; there has never been an outbreak in a school worldwide that we know of. Care homes are a different matter, but they were not in the models”¹⁴. If the right data cannot be obtained or the modellers are prevented from using the right data or asking relevant questions of it then we cannot possibly generate the right data or results to drive policy. Here we can see the Government Policy was not reflective of the modelling and indeed the non-evidence based policy started to change so quickly that modellers could do little to inform the Government on the right action to take. The mixed messaging that was generated from this activity did little to provide transparent messaging to the public and the non-evidence based policy has defied logic on multiple occasions.

The model premises that were used to determine key decisions have never been tested against the empirical data and presented, as one would expect from a basic data science perspective. On the contrary, the vast amount of external empirical analysis of the decision-making model unanimously demonstrate that the measures have done very little to affect transmission as purported, let alone save any lives.

A significant amount of the data that has appeared to drive the measures implemented are extremely pessimistic in nature. For example, the number of COVID related deaths has been defined by Public Health England is defined in two categories¹⁵:

¹¹ Lords Select Committee on Science and Technology Afternoon Session – Corrected oral evidence: The Science of COVID-19 (London, 2 June 2020)

¹² Lords Select Committee on Science and Technology Afternoon Session – Corrected oral evidence: The Science of COVID-19 (London, 2 June 2020)

¹³ Lords Select Committee on Science and Technology Afternoon Session – Corrected oral evidence: The Science of COVID-19 (London, 2 June 2020)

¹⁴ Lords Select Committee on Science and Technology Afternoon Session – Corrected oral evidence: The Science of COVID-19 (London, 2 June 2020)

¹⁵https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/916035/RA_Technical_Summary_-_PHE_Data_Series_COVID_19_Deaths_20200812.pdf

1. A death in a person with a laboratory-confirmed positive COVID-19 test and either died within 60 days of the first specimen date or died more than 60 days after the first specimen date, only if COVID-19 is mentioned on the death certificate.
2. A death in a person with a laboratory-confirmed positive COVID-19 test and died within (equal to or less than) 28 days of the first positive specimen date.

The result of such broad data is that the risk may be over-exaggerated such that it cannot be fully measured against the opposing risks in implemented countermeasures. It may also mask the trends in deaths resulting from other conditions. It appears that the data has been oversimplified in the attempt to adopt a 'better safe than sorry' attitude. This will ultimately generate an unproportional response and heightened risk perceptions and prevent the ability to properly balance the risk against those introduced by any measures taken. Equally, as highlighted by Labib ¹⁶ there should be greater expansion in the measurement of COVID-19 related deaths, particularly to include the introduction deaths in-directly caused by COVID-19.

5. Was data shared across the devolved administrations and local authorities to enable mutually beneficial decision making?

The fact that local authorities are arguing with central government about the appropriateness of measures and the inconsistency of policies across the devolved administrations would suggest that this is not the case. If the data was consistent, robust and readily shared, one would expect a much more uniform policy response. The virus does not respect any geographical boundaries, especially ones in a homogeneous geography like the UK. How is it feasible to have such varied policy responses?

6. Is the public able to comprehend the data published during the pandemic. Is there sufficient understanding among journalists and parliamentarians to enable them to present and interpret data accurately, and ask informed questions of Government? What could be done to improve understanding and who could take responsibility for this?

There is clearly not a marriage between measures that could be used to help public understanding and the use of them. Scientists, at the time, and subsequently the Lords Select Committee, recommended alternative methods of expressing complex data to the public. This was unfortunately not acted on by the Government. Therefore, an R measure, which was deemed unfit for purpose, has been relied upon erroneously to drive policy decisions. No other reasonable information except out of context debatable death rates, and spurious "case" data determined by faulty use of the PCR test, has been widely distributed.

This is despite teams of modellers requesting that more relevant parameters be examined and better methodology for communication be used. The media have, as a consequence, been presenting a picture to the public that is not only illogical but highly misleading in terms of measures and statistics. To follow this path along with a SAGE endorsed campaign of coercion is highly unethical¹⁷. Scientists working with the Government on COVID and Government

¹⁶ Labib, P. L., & Aroori, S. (2020). Expanding the definition of covid-19 deaths will show the true effect of the pandemic. *bmj*, 369.

Departments are responsible for public understanding of statistics and data as is the RSS, who offered help with this aspect of COVID, but it is entirely the Government's responsibility to listen and take action. It is embarrassing for the general public to be able to pull the modelling and statistics apart that the Government is disseminating to them and discredit this work with basic analysis.

And for the CSO to present fatuous charts like the one on cases doubling every week in September with the message “this is not a prediction but a big if” is disingenuous at the very least. The reality of the situation was and is much less dire than presented.

A Caveat to the SPI-B Report, and many SAGE reports (although it has been noticed this has been removed in later versions) is that much of the evidence that has been drawn on is very recent and has not been subject to peer review. This report has been put together rapidly and been subject to limited scrutiny and review. This is counter to the Government policy of evidence based decision making and robust analysis. The lack of interdisciplinary work within this field has led to a tunnel vision approach of one type of person with one type of aim. The countermeasures and their impact were not based on evidence nor subject to scrutiny.

SPI-B produced APEASE (Acceptability, Practicability, Effectiveness, Affordability, Spill-over effects, Equity)¹⁸, An evaluation grid for options to rapidly increase general social distancing.

The media’s purpose here is not to hold government to account, but to act as a conduit for the ‘behavioural’ message (coercion and psychological messaging).

They recommend that the media is used:

- to increase sense of personal threat
- to increase sense of responsibility to others
- to promote positive messaging around actions

On 17 April this year it was announced that the Government and the newspaper industry have formed a three-month advertising partnership called All in, all together to help “keep the public safe and the nation united” throughout the Covid-19 ‘pandemic’¹⁹. News media now have a commercial interest in repeating what they are told to repeat and not critiquing or holding the Government Accountable.

7. Does the Government have a good enough understanding of data security, and do the public have confidence in the Government’s data handling?

As a professional I have serious doubts about the data handling. Firstly, in terms of security and GDPR the initial collections of data for tracing were not secure, and it is debatable if they still are. Consent should have been integral to this. Secondly, the data around covid, when collected had clearly been of poor quality. The changing guidelines and interference on data collection that

¹⁷https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/887467/25-options-for-increasing-adherence-to-social-distancing-measures-22032020.pdf

¹⁸https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/887467/25-options-for-increasing-adherence-to-social-distancing-measures-22032020.pdf

¹⁹ <http://www.newsmediauk.org/Latest/government-partners-with-newspaper-industry-on-covid-19-ad-campaign>

had been recognised as reasonable for many decades caused data to cease to be robust. This interference has now had a serious effect on national statistics that mean the data can no longer be relied upon.

8. How will the change in responsibility for Government data impact future decision making?

The most important factors are to have robust data available to everyone in a timely manner and then for the Government to actively seek out analysis and insights from this analysis from outside of its own self-determined echo chamber. Any change in responsibility is irrelevant as long as that is respected. Instead, it seems that the data used for analysis is, due to interference, faulty and conflated, and the Government seems determined to censor dissent and counter-argument rather than question their own narrative. We have made a request to the ONS for a timeseries of raw data extracted from death certificates which is not forthcoming as it is “provisional”. It doesn’t matter who is responsible for data, it is the quality and availability of crude data to the public that is important, not data that has been filtered and manipulated by Government bodies before being made available.

October 2020