3 November 2020

Submission to the Public Administration and Constitutional Affairs Committee inquiry into Data transparency and accountability: Covid 19

Key points:

- COVID-19 has shown that neither the public nor the private sector has strong organisational competence in well governed access to data and its exploitation. There have been significant limitations in data systems that have hampered the UK's response. The government should invest now to establish sophisticated mechanisms for accessing a broad range of external evidence and expertise from the UK and international science community in emergencies. This requires building better data systems to enable rapid and timely access to and analysis of data for policymakers in a way that is ethical, secure and privacy preserving.

- The UK has extraordinary data science capability and there have been large-scale and rapid interventions to build data systems to respond to COVID-19. There is a moment now to build on that rapid digital transformation, through concerted multi-sectoral action in order to establish:
  - **Systematic data collection** – building on the current initiatives to gather comprehensive surveillance data across the whole population
  - **Quality and timely data –** there is a need for up-to-date and accurate data to understand where interventions are needed and whether and where they are working
  - **Connected data –** data which is linked across the various parts of the health system and beyond
  - **Data access –** enabling swifter and easier access to data by researchers helping to respond to the pandemic
  - **Data governance –** ensuring that robust governance arrangements, such as data sharing agreements, are established in advance so that agreements to access and use data can be made rapidly, safely and ethically

- The current crisis has tested the UK public's trust in science and the science community needs to work hard to maintain the trust we enjoy. It is important that we retain our independence and our transparent and open approach to our work. Giving everybody the tools to critically engage with science, including promoting science education, will also support a healthy trust in science and engagement with its findings. There are a number of steps that should be taken to ensure this:
  - **Develop simulations and data visualisations that boost accessibility and understanding:** Decisions are being made based on complex data and analyses. Robust simulation and powerful visualisation has the potential to enhance the understanding of these and improve the ability of policymakers and the public to meaningfully scrutinise decisions.
  - **Recognise uncertainty:** Experts should keep engaging publicly, taking time to explain the evidence and the limitations of this. When they are 'pitted against' other experts, they should take responsibility for explaining why, given the limitations of the data there is no one clear answer, which is why debate continues and it is not possible at this stage to say that one view is right and one is wrong.
  - **Engage with the public honestly, openly and transparently:**
    - Experts should continue to work with organisations like the Science Media Centre (SMC) that can ensure that expert voices are available to the media

quickly to explain concepts and outline the evidence, being clear about its limitations.

- The government should continue to invest in skilled people and organisations that support meaningful public engagement with the challenging and sensitive questions that are raised by the progress and application of science and use this to inform policymaking.
  - o **Longer term, the government should review the education system to ensure that it provides everyone with the tools to critically engage with science, empowering them to make informed choices about scientific and technological development and equipping them to work in an advanced knowledge economy.** The government should establish a royal commission to set out a vision for a secondary education system that will nurture future generations of innovative thinkers and resilient citizens.

Introduction:

1) The Royal Society is the national academy of science for the UK. Its Fellows include many of the world's most distinguished scientists working across a broad range of disciplines in academia, industry, charities and the public sector. The Society draws on the expertise of the Fellowship to provide independent and authoritative advice to UK, European and international decision makers.

2) Expert Fellows and researchers funded by the Society are working alongside researchers around the world to tackle COVID-19. The Royal Society has convened two groups using data science methods to respond to the pandemic. Data Evaluation and Learning for Viral Epidemics (DELVE) is a multi-disciplinary group, supporting a data-driven approach to learning from the different approaches countries are taking to managing the pandemic, analysis to complement the evidence base informing the UK's strategic response. Rapid Assistance in Modelling the Pandemic (RAMP) is an initiative to bring modelling expertise from a diverse range of disciplines to support the pandemic modelling community already working on Coronavirus (COVID-19). RAMP is designed to; provide support for existing research groups; create new models or insights that can be used to inform the work of the Government's scientific advisors, through data science-based approaches; apply knowledge from related epidemiology domains; and triage incoming literature to ensure effective information flows. In addition, the Royal Society has established an internal group, Science for Emergency Tasking – Covid (SET-C) to rapidly analyse and synthesise evidence relevant to the pandemic response.

3) This response draws on experience and insight from across these three initiatives. It also draws on the Royal Society's body of policy advice relating to data and digital technologies, which has highlighted the strategic and societal importance of well-governed access to data.

Did the Government have good enough data to make decisions in response to Covid-19, and how quickly was the Government able to gather new data?

4) There has been significant recent investment in surveillance initiatives in PHE (and devolved nation counterparts), the Office for National Statistics (ONS), Health Data Research UK (HDRUK) and the Joint Biosecurity Centre (JBC). However, a better coordinated and responsive surveillance system is required both to understand the nature of the pandemic and to respond in real time. Comprehensive, accessible national data on the proportion of people who have been infected with COVID in different groups sampled is currently limited. Better data is essential to understanding risk, depending on sex, ethnicity, age, occupation, socio-economic factors and comorbidities. Understanding these risks is vital to inform an effective response to protect those most vulnerable and drive down transmission. For example, DELVE's work on nosocomial infections (infections acquired in hospital and health-care settings) highlighted

important gaps in data: " although surveillance systems and large-scale hospital-based studies have recently been set up, there remain gaps in availability of surveillance data on hospital-acquired infections, particularly of healthcare workers (including agency staff) and in nursing homes; important questions remain unanswered, including about the impact on Black, Asian and minority ethnic (BAME) health and social care workers."[1]

5) Disinvestment in both central and local public health departments mean that current systems are now fragmented. A high-quality modern system, with first-class and appropriate expertise, infrastructure and resources is needed for the public health system and NHS to be able to collect and use data to protect the public. In the immediate future, clear, mandated leadership is needed within the Department for Health and Social Care to enable the collection and connection of data from across the health system, building on important work carried out by Health Data Research UK and the Office for National Statistics (ONS) during the pandemic.

Was data for decision making sufficiently joined up across Departments?

6) There is a clear need to connect data across the different parts of the health system. Connected COVID-19 data systems across community, care institutions and the NHS are essential to understand where and how transmission is occurring and, for example, to identify the risk of importing and exporting cases across the health and social care system. Local and national data systems need to work in synergy. This is critical to local public health teams, and local authorities and hospitals, in the rapid identification, investigation and management of outbreaks.

7) Health data also need to be connectable to other forms of socio-economic data to better understand what is happening at the individual and population level in a way that protects personal or sensitive information to ensure individual's privacy. From the point of view of the research community responding to the crisis, a critical weakness in the system was access to important social and behavioural datasets. These are the kinds of data routinely used by major corporations for marketing and business development but inaccessible for scientific research and for decision-making in relation to public policy. Examples include retail sales and loyalty cards, financial transactions, mobile telephone records, smart tickets for transport. These data have the ability to richly inform patterns of mobility, interaction and human contact (hence the likelihood of transmission in business as usual and various lockdown/ pandemic stages) but they were not available to access and exploit at suitable levels of geographical and temporal resolution.

8) There needs to be much more individual-based data available to scientists, public health professionals and political decision-makers regarding the way the density of populations in venues ranging from schools and supermarkets to different outlets in the hospitality sector, showing how people move. Hence, better connecting health and mobility data can have a major impact. There may have been such connections within government, but from the point of view of external researchers collecting and connecting such data was a challenge.

---

[1] Royal Society DELVE initiative (2020) *Scoping Report on Hospital and Health Care Acquisition of COVID-19 and its Control* https://rs-delve.github.io/reports/2020/07/06/nosocomial-scoping-report.html

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?

9) From the point of view of researchers in universities and industry, aiming to help support the COVID-19 response, there was a significant challenge in access data. This was especially the case for researchers with appropriate methodological skills who do not primarily work with health data. The Control of Patient Information (COPI) notice has been helpful, but access to patient data has been difficult to achieve because of the terms imposed by the Data Controllers. Accessing sensitive data has, of course, to be closely managed, following guidelines such as the Five Safes[2] to protect patient privacy. It is however the case that the Trusted Research Environments have found it hard to meet the demand for their services, adding delays to the process. Access can be costly with approval processes for even anonymised data taking up to a year, according to a blog article published by the British Medical Journal which highlighted barriers to data access.[3] Similar calls were made in the Proceedings of the National Academy of Sciences on the urgent need for international social and demographic data. Researchers working with the Royal Society and British Academy have struggled to get data disaggregated by ethnicity, region, age and sex – which is essential to understanding inequalities in excess death rates during the COVID pandemic.

10) European Union nations, which are also subject to GDPR, have had public, unrestricted access to crucial Covid-19 data.[4] [5] At the click of a button, a UK scientist can see the daily flow of people from origin to destination between more than 3000 districts in Spain – a country that complies with the same data protection rules as the UK – but this data is not readily available for the UK. Valuable UK-based research was carried out over these datasets in the absence of UK data and the opportunity to build valuable knowledge was missed as a result of this absence of data.

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?

11) DELVE work *on Economic Aspects of the COVID-19 Crisis in the UK* highlighted the importance of bringing together epidemiological and economic data to better understand the impact of non-pharmaceutical interventions. [6] As argued in this report:

    a. Designing sound policy – and effectively tailoring it to ensure maximum impact – requires access to high-quality data. Those developing and monitoring the implementation of economic interventions would benefit from access to a range of different data types, including data relating to transmission of COVID-19, use of public

---

[2] ONS (2017) *The 'five safes' – data privacy at ONS* https://blog.ons.gov.uk/2017/01/27/the-five-safes-data-privacy-at-ons/

[3] The BMJ opinion (2020) *Reducing barriers to data access for research in the public interest—lessons from covid-19* https://blogs.bmj.com/bmj/2020/07/06/reducing-barriers-to-data-access-for-research-in-the-public-interest-lessons-from-covid-19/

[4] COVID-19 Belgium Epidemiological Situation. Municipality Monitoring. https://datastudio.google.com/embed/reporting/c14a5cfc-cab7-4812-848c-0369173148ab/page/giyUB, accessed 15 October 2020

[5] Robert Koch-Institut (Germany) infections binned into medium sized geographical regions https://www.arcgis.com/home/item.html?id=f10774f1c63e40168479a1feb6c7ca74, accessed 15 October 2020

[6] https://rs-delve.github.io/reports/2020/08/14/economic-aspects-of-the-covid19-crisis-in-the-uk.html

transport, consumer spending and financial transactions. Many useful indicators of the economy's response to COVID-19 could be found in data typically held by private financial institutions and are generally considered to have commercial value. While some progress has been made in making such data available, the UK lags behind its European counterparts in its ability to mobilise this data to develop policy responses, and action is needed by government to ensure fine-grained financial data is made available for such analysis.

12) The same DELVE report recommended that "UK government and supporting bodies continue pursuing one of the most important and urgent goals of obtaining more and more representative and fine-grained transactions data from UK financial institutions for the purpose of monitoring the impact of COVID-19 and associated policy interventions. Such data are valuable assets in the current situation, and the UK should be the vanguard in efforts to harness its power for effective and timely policymaking."

13) The granularity and nature of data at a local level is also an important concern. The current lockdowns are based on crude data with respect to localities, as one of the key issues is movement of people between places and those movement patterns need to be taken account of. Consideration of the density of populations and how these vary over periods of hours and weeks can lead to a better understanding of the 'temporal' aspects of transmission: in other words, do infections spread faster during the evenings when people are out at pubs and restaurants; in work where there is sustained activity but social distancing? Temporal restrictions could then become a greater tool in the armoury of social distancing, building on current curfews. Much more could be done with better data and better understanding.

14) Both data, and modelling and simulation using that data, are important for decision making. Modelling and simulation capabilities demonstrated by the Royal Society RAMP initiative in partnership with the Alan Turing Institute shows the potential to inform lockdown decisions at a local scale, to project the evolution of cases through time, and most importantly to evaluate the impact of alternative lockdown scenarios on future R-numbers, incidence of cases and the disease burden.

Was data shared across the devolved administrations to enable mutually beneficial decision making?

15) We do not have a detailed overview of the extent of data sharing and mutual decision making across the devolved administrations, however it is important to note that this is an area that may benefit from greater coordination – for example for a long period, the devolved administrations had different criteria for what constitutes death from COVID-19, creating challenges in extracting a consistent measure.

Is there sufficient understanding of statistics by the public to understand the detailed information that is being regularly published during the pandemic? Equally, is there enough understanding among journalists and parliamentarians to enable reasonable presentation and interpretation of data for the public; and informed questions to be asked of the Government? What could be done to improve understanding and who could take responsibility for this?

16) The current crisis has tested public in trust in science internationally, including in the UK. Science and scientists have been under the spotlight and at the heart of political decision-taking like never before.

17) We have learned a lot amazingly fast. This is a new virus about which we knew nothing just a few short months ago. However, there is a lot that we still do not know.

18) At the frontiers of science, there is always uncertainty, and to pretend otherwise would be foolish. What science does is to try to gather evidence to reduce the uncertainty, but this happens only gradually as data are gathered and hypotheses tested and discarded until some idea of the truth emerges. But even those "truths" can fall by the wayside in the face of new and contradictory evidence. The entire process is based on honesty, openness and transparency, in which the evidence is published for all to see and argue about. This can, however, be a very confusing process for those observing from the outside, who may be unfamiliar with the process of science and the uncertainties that it operates within and therefore may be left unsure who, or what, to believe. Research suggests that there is a correlation between poor numerical literacy and susceptibility to mis-information about COVID-19[7]. It is important to be open and honest with the public about what we know and the ongoing uncertainties.

19) It is right for government to take an evidence-based approach. However when talking about this, government ministers need to be clear that they are making the best decisions now based on the available evidence, but are also prepared to change tack later when further evidence comes to light. They must also make clear that science advice is only one of the things under consideration when they are making hard choices about how to respond to the pandemic. Also under consideration are economics, ways of implementation, broad consequences to society, and the need to take the public with them. This clarity, that it is not simply a case of 'following the science' is important for parliamentarians, the media and the public who wish to scrutinise government decisions. Alongside access to the scientific evidence, it will enable them to direct their questions more effectively.

20) During the pandemic to date, we have observed a rapid increase in the familiarity of the UK government, parliamentarians, the media and the public with complex statistical terms, such as the 'R number'. Discussions of the latest scientific evidence are hosted on mainstream media. As noted above, this represents valuable transparency of the scientific process in action, but can also be confusing. While there may be an increase in scientific literacy, there may also be simply an increase in familiarity with scientific terms with little understanding of their meaning. There are a number of steps that can be taken to ensure that this higher profile is helpful:

- **Develop simulations and data visualisations that boost accessibility and understanding:** Decisions are being made based on complex data and analyses. Robust simulation and powerful visualisation have the potential to enhance the understanding of these and improve the ability of policymakers and the public to meaningfully scrutinise decisions. Limited visualisation can impact on understanding for example, there was an assumption amongst many at the start of the pandemic that denser places had more than proportionate numbers of infections and this was rarely tested. London was often singled

---

[7] Roozenbeek et al (2020). Susceptibility to misinformation about COVID-19 around the world
https://royalsocietypublishing.org/doi/10.1098/rsos.201199

out, but in fact if the data on infections had been normalised on a per capita or household basis the picture might have been different.

- **Recognise uncertainty:** Experts should keep engaging publicly, taking time to explain the evidence and the limitations of this. When they are 'pitted against' other experts, they should take responsibility for explaining why, given the limitations of the data there is no one clear answer, which is why debate continues and it is not possible at this stage to say that one view is right and one is wrong.
- **Engage with the public honestly, openly and transparently:**
    - Experts should continue to work with organisations like the Science Media Centre (SMC) that can ensure that expert voices are available to the media quickly to explain concepts and outline the evidence, being clear about its limitations.
    - The government should continue to invest in skilled people and organisations that support meaningful public engagement with the challenging and sensitive questions that are raised by the progress and application of science and use this to inform policymaking.
- **Longer term, the government should review the education system to ensure that it provides everyone with the tools to critically engage with science, empowering them to make informed choices about scientific and technological development and equipping them to work in an advanced knowledge economy.** The government should establish a royal commission to set out a vision for a secondary education system that will nurture future generations of innovative thinkers and resilient citizens.

How will the change in responsibility for government data impact future decision making?

21) The use of data, not just in and by government, but in partnership with scientists and academic research teams, should be a key part of long-term strategy. This needs to be driven across government and the Cabinet Office is well-placed to do so. Such partnerships would support not just the immediate response, but research to better understand the impact of COVID-19 and similar or related shocks in the future, and potentially build public trust given relatively high levels of confidence in science and universities. Such work could potentially be more actively promoted through partnerships between government, academics and business organisations possibly with more emphasis on the development of Data Trusts, or data sharing agreements, following on from the Hall-Pesenti review into AI. The National Data Strategy led by DCMS, and the Cabinet Office drive to improve government use of data, offer the opportunity to embed such valuable partnerships and must be closely coordinated. The Royal Society will be responding to the consultation on the National Data Strategy, expanding on the points made here.

For further information, please contact public.affairs@royalsociety.org