Written evidence from the Centre for Public Data to the 'Data Transparency and Accountability: Covid 19' inquiry

30 October 2020

Summary

- 1. This is a response from the Centre for Public Data to the Public Administration and Constitutional Affairs Committee's call for evidence on data transparency and accountability during the Covid-19 pandemic¹. The key points of our response are:
 - a. Government lacks data in many areas needed to understand the impact of the pandemic. We focus on its broader impact and find missing data in areas from court delays to NHS waiting lists, and in Government's financial support schemes. This makes it hard for Government, Parliamentarians, researchers and citizens to assess the broader effect of policy interventions.
 - b. Failure by central government to share data has affected the local response. We highlight problems with local access to shielding lists, death projections and case numbers, despite well-established frameworks for health data sharing.
 - c. There have also been failures to share the data that has underpinned decision-making. We highlight data on local lockdowns and ICU statistics.
 - d. Systemic interventions will be needed to address these issues. This will include data-focussed updates to the UK's policymaking processes to make them fit for the 21st century; new data-literate and independent institutions to enforce data quality and accountability; new provisions to support access to third-party data; and where appropriate, legislative changes.

About us

- 2. The Centre for Public Data (CFPD) is a new, non-partisan organisation with a mission to strengthen the UK's public data². We aim to reduce gaps in data that harm civil society and business. We support legislators and policymakers to improve data coverage and quality, via practical interventions in legislation, codes of practice and governance.
- 3. This response focuses on data that was not collected, or not shared, by Government during the pandemic. Our answers are based on desk research, an automated study we have conducted of Parliamentary written questions (PWQs) during the pandemic³,

¹ https://committees.parliament.uk/call-for-evidence/250/data-transparency-and-accountability-covid-19/

² https://www.centreforpublicdata.org/

We will submit this as supplementary evidence.

and a public survey we ran during October 2020⁴. We are happy to discuss any of these issues further: contact@centreforpublicdata.org

Responses to questions

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

- 4. The Committee will undoubtedly receive expert responses about the epidemiological and NHS data available to Government. We focus instead on other data needed for policy responses to the epidemic.
- 5. We have identified important topics where Government appears to lack adequate data via our analysis of PWQs, whose answers typically indicate when data is not available to the Secretary of State. We highlight two areas with evidence of poor-quality data: (i) the social impact of the pandemic and (ii) the use and design of Government's financial support schemes.
- 6. These data issues clearly affect Government's ability to understand the broader costs of policy responses and balance these against the need to save lives.

Data on the social impact of the pandemic

- 7. It is clear that the pandemic has caused significant delays in the courts system. However, a lack of data means it is hard to quantify the extent of disruption. Data appears to have been lacking on e.g.: how many victims of crime have been affected by delays in court proceedings during the Covid-19 outbreak⁵; how many and what proportion of civil cases are currently awaiting a trial date⁶; and how many benefit tribunals have been postponed⁷. These data issues are also discussed in the Justice Select Committee's report on Covid-19's impact on the courts⁸.
- 8. It is concerning that the Secretary of State for Justice does not have the data to understand the full impact of the pandemic on the courts. This reflects long-standing issues with adequate management data from the court system, as discussed by Dr Natalie Byrom in her report on the use of data in the HM Courts & Tribunals Service⁹.
- 9. An area of significant concern early in the pandemic was border control, both the number of overseas arrivals and later whether quarantine measures were being adhered

⁴ https://docs.google.com/forms/d/15EV4taKOw3Fs3vnSyN7FgSPZ71hhn5mLo4BDMZ5mtl8/

⁵ https://www.theyworkforyou.com/wrans/?id=2020-05-01.42029.h

⁶ https://www.theyworkforyou.com/wrans/?id=2020-04-29.41441.h

https://www.theyworkforyou.com/wrans/?id=2020-05-05.43706.h

⁸ https://publications.parliament.uk/pa/cm5801/cmselect/cmjust/519/51905.htm

⁹ https://www.gov.uk/government/news/hmcts-response-and-progress-update-on-dr-natalie-byrom-report

- to. The Government lacked e.g. data on the number of arrivals from countries with quarantine restrictions, and the number who registered their contact address¹⁰. This seems to reflect long-standing issues with data quality in border control.
- 10. Remarkably, Government also seems to lack robust data about the effect of the pandemic on NHS services. Data is not held on e.g.: how many operations for cancer have been cancelled since the outbreak of the pandemic¹¹; how many patients had radiotherapy treatments delayed and cancelled¹²; how many total and routine cervical smears were delayed¹³; how many routine appointments and red flag referrals have been cancelled or postponed¹⁴; the effect on waiting lists for child and adolescent mental health services¹⁵.
- 11. This lack of data on NHS services appears to reflect a lack of routine access to timely local management information outside of centralised reporting statistics. While Government created the NHS Covid-19 Data Store to bring together data on the pandemic¹⁶, it is unclear whether it has similar access to broader performance data.

Lack of reporting data about businesses using financial support schemes

- 12. The Government has used the Coronavirus Business Interruption Loans Scheme (CBILS) and related loan and wage schemes to support businesses affected by Covid-19. In most cases, no data has been published on the businesses which have accessed this support. This makes it hard to evaluate their impact, and will have contributed to high levels of fraud. By contrast, the US publishes loan-level data on equivalent programmes¹⁷.
- 13. HMRC has reported to the Public Accounts Committee that up to 10% of CBILS loans may be fraudulent¹⁸. The National Audit Office (NAO) has stated since the names of businesses accessing the Coronavirus Jobs Retention Scheme were not public, employees could not whistleblow on fraudulent furlough claims¹⁹.
- 14. It is surprising that this data was not published in real time, since loan-level data will soon have to be published under EC regulations, so there seems little benefit to delaying publication²⁰. There appears to have been a lack of focus on operational data within HMRC, failure to clarify with the British Business Bank (which operated most loan schemes) what reporting data was required, and confusion within HMRC about its data publication responsibilities.

https://home.treasury.gov/policy-issues/cares-act/assistance-for-small-businesses/sba-paycheck-protection-program-loan-level-data

¹⁰ https://www.theyworkforyou.com/wrans/?id=2020-08-28.82528.h

https://www.theyworkforyou.com/wrans/?id=2020-04-28.HL3613.h

https://www.theyworkforyou.com/wrans/?id=2020-05-11.45139.h

¹³ https://www.theyworkforyou.com/wrans/?id=2020-04-21.38646.h

¹⁴ https://www.theyworkforyou.com/wrans/?id=2020-04-28.41055.h

¹⁵ https://www.theyworkforyou.com/wrans/?id=2020-06-17.60643.h

https://data.england.nhs.uk/covid-19/

https://committees.parliament.uk/event/1837/formal-meeting-oral-evidence-session

¹⁹ https://www.nao.org.uk/report/implementing-employment-support-schemes-in-response-to-the-covid-19-pandemic/

https://questions-statements.parliament.uk/written-questions/detail/2020-10-01/98272

Lack of operational data to support the design of financial support

- 15. The National Audit Office estimates that around 1.1 million people were ineligible for support from the Coronavirus Job Retention Scheme because HMRC had insufficient data to verify their claims²¹. These included people working via short-term contracts, and company directors paid via dividends.
- 16. One survey response we received from a freelancer whose income had vanished stated this movingly: "Tax forms need to glean more detail as to the nature of one's business and the specifics of our jobs... There is no data other than the data we have managed to pull together that articulates the full story for us. This is the nightmare we have found ourselves within".
- 17. We also heard from a senior cultural leader involved in shaping the Cultural Recovery Fund, who highlighted a lack of high-quality data on the extent of the cultural sector and particularly the contribution of freelancers and community organisations. This lack of data on non-traditional workers affected their ability to make a case for the Fund to support the sector.

Question 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?

- 18. No. We have found a number of areas where relevant data was not, and in some cases is still not, disseminated.
- 19. This is not the fault of the UK's health data sharing framework. Before the crisis, there was a robust pre-existing legal and governance framework for the sharing of health data. For example, the Data Protection Act 2018 and Health Service (Control of Patient Information) Regulations 2002 contain provisions to share personal data for public health purposes²². The Health and Social Care (Safety and Quality) Act 2015 requires the use of consistent identifiers for patients across health-related services²³. Additional data sharing guidance was put in place early in the pandemic. On 1 April, DHSC formally notified health authorities to share confidential patient information²⁴. The Information Commissioner's Office and the National Data Guardian for Health and Adult Social Care stated that data protection law should not stand in the way of sharing data to help with public health needs²⁵.

https://www.legislation.gov.uk/uksi/2002/1438/regulation/3/made

https://www.legislation.gov.uk/ukpga/2015/28/contents/enacted

https://www.gov.uk/government/speeches/data-sharing-during-this-public-health-emergency

²¹ https://www.nao.org.uk/report/implementing-employment-support-schemes-in-response-to-the-covid-19-pandemic/

https://www.legislation.gov.uk/ukpga/2018/12/contents/enacted

²³ https://www.legislation.gov.uk/ukpga/2015/28/contents

https://www.gov.uk/government/publications/coronavirus-covid-19-notification-of-data-controllers-to-share-information

https://ico.org.uk/about-the-ico/news-and-events/news-and-blogs/2020/03/data-protection-and-coronavirus/

20. Despite this robust framework, data was not shared with local authorities effectively, for reasons that have not been publicly explained.

Shielding lists to help local authorities deliver support to vulnerable people

- 21. On 21 March 2020, Government announced that patients who were clinically vulnerable to Coronavirus should shield until further notice²⁶. Local Resilience Forum (LRF) members have informed us that the Government was slow in securely sharing its list of shielded patients with local authorities, so that support could be delivered to vulnerable residents before national services were in place. One LRF member told us: "This information was requested regularly, sometimes on a daily basis, and it was difficult to establish why it wasn't being shared".
- 22. It is unclear why this happened, given the guidance above, and given that local authorities had existing processes for managing lists of vulnerable residents.

Excess death projections to help local authorities plan mortuary capacity

23. During March 2020, LRF members worked to ensure local mortuary capacity. A survey responder told us that their LRF requested modelling data to aid planning (the Government Reasonable Worst Case Scenarios²⁷), but were refused: "Government had the modelling, presumably through SAGE, but they told us there was no ministerial approval to share it". This affected their planning: "We were given 'hints' about what the numbers might have been, such as 'it is likely to be a bit worse than the normal worse winter week, but you should build some contingency in' without any sense of what that contingency level should be… We would… possibly not have built in so much capacity, if we had anticipated numbers to work with". It is unclear why this data required ministerial approval to share with LRFs.

Local case data to support the first local lockdown

- 24. On 29 June 2020, DHSC announced local outbreak control restrictions in Leicester City and parts of Leicestershire County, to be reviewed every 2 weeks²⁸. This was the first occasion since the onset of the pandemic that local measures had been imposed.
- 25. The local authority expressed concern about its lack of access to detailed data on cases, allowing it to understand and target its lockdown response. It was not until late June that Public Health England (PHE) agreed to share data with Leicestershire County

²⁶

https://web.archive.org/web/20200322111238/https://www.gov.uk/government/publications/guidance-on-shielding-and-protecting-extremely-vulnerable-persons-from-covid-19/guidance-on-shielding-and-protecting-extremely-vulnerable-persons-from-covid-19

<u>-covid-19</u>
²⁷ It appears that national-level information was eventually published by SAGE in October 2020:

https://www.gov.uk/government/publications/spi-m-o-covid-19-reasonable-worst-case-planning-scenario-21-may-2020

https://www.gov.uk/government/speeches/local-action-to-tackle-coronavirus

- Council at postcode level for 'planning and service purposes'²⁹. This postcode-level data revealed that factories were not behind the local surge, as had been speculated.
- 26. Dame Mary Ney's rapid stocktake of lessons learned from Leicester reports that initial decision-making suffered from a lack of clear data sharing arrangements between central and local government: "There was much concern expressed by the councils at their lack of access to detailed local data sufficient to assist them in targeting interventions and in providing explanations to the public as part of the community engagement plan".
- 27. It has not been explained adequately why the need for these additional data sharing agreements on case-level data was only discovered in June, given the prior work on data sharing in March, April and May.

Sharing of Pillar 2 test data to enable local contact tracing

- 28. On 17 July 2020, the British Medical Journal (BMJ) reported that local public health teams were finding that community (Pillar 2) test data often lacked postcode and occupational data³⁰. As a result, local teams could not use the data to examine possible connections between cases. It is unclear why this was the case, given that the law required testing labs to collect this data³¹.
- 29. The BMJ also reported that local teams were being denied access to identifiable information, preventing local contact tracing. Again, it is unclear why access was denied, given law requiring the use of consistent patient identifiers across health authorities³², and guidelines permitting the sharing of patient information for direct care, as above.

Other data

30. A researcher who responded to our survey reported that they had been unable to access data on death numbers in individual care homes. Although care homes must submit statutory notifications of deaths to the Care Quality Commission³³, an FOI request for these numbers was refused on the basis that the data would prejudice commercial interests. The researcher noted that the US, Canada, and Australia all freely publish similar data on individual care homes.

²⁹ https://www.leicestershire.gov.uk/news/covid-19-further-data-released-by-county-council

³⁰ https://www.bmj.com/content/370/bmj.m2883

³¹ In England, Section 5 of the Health Protection (Notification) Regulations 2010, testing laboratories have a duty to notify PHE of positive tests for notifiable diseases. Covid-19 became a notifiable disease on 16 March 2020. https://www.legislation.gov.uk/uksi/2010/659/regulation/5

https://www.legislation.gov.uk/ukpga/2015/28/contents

³³ https://www.legislation.gov.uk/uksi/2009/3112/regulation/16/made

Question 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?

31. We present some examples where this has not been the case

The evidence framework for local lockdowns

- 32. On 10 May, the Prime Minister announced a new organisation, the Joint Biosecurity Centre (JBC), to provide "evidence-based analysis to inform local and national decision-making in response to outbreaks", stating "the JBC will follow a transparent approach to the assessments and analysis it undertakes"³⁴.
- 33. To date, the JBC has not published any of its evidence, minutes or decision-making processes. It remains unclear how decisions to move local areas between tiers of lockdown are made.
- 34. In the House of Lords, the Minister responsible suggested that criteria for moving between local tiers may not be entirely data-driven: "We have not published strict criteria for each exit strategy for moving from one tier to another. It is part art and part science, in any case."
- 35. Unusually, the Secondary Legislation Scrutiny Committee has urged the Government to publish its supporting data around lockdown decisions, stating³⁵: "For public trust to be maintained, the Government have to be much more transparent in explaining the basis of their decision-making, including setting out how they balance the competing health, social and economic interests and the data to support the decision".

The evidence around ICU capacity

- 36. The availability of intensive care unit (ICU) beds in local areas is a topic of increasing public and media interest. However, official statistics on the topic have been paused, and no alternative official data has been supplied.
- 37. Each month since 2010, NHS England (NHSE) has published data at NHS Trust level on the numbers of available and occupied ICU beds, and counts of urgent and routine cancelled operations³⁶.

³⁴ https://www.gov.uk/government/groups/joint-biosecurity-centre

https://committees.parliament.uk/committee/255/secondary-legislation-scrutiny-committee/news/120231/government-should-be-more-open-on-how-it-decides-on-regional-covid-tiers/

³⁶ https://www.england.nhs.uk/statistics/statistical-work-areas/critical-care-capacity/ https://www.england.nhs.uk/statistics/statistical-work-areas/cancelled-elective-operations/

- 38. However, in March 2020, NHSE suspended the publication of these datasets, stating: "Due to the coronavirus illness and the need to release capacity across the NHS to support the response, we are pausing the collection and publication of these and some of our official statistics"³⁷. Publication has not resumed since.
- 39. NHSE states that the suspension is in line with Office for Statistics Regulation (OSR) guidance for statistics producers during Covid-19, though it is unclear that this is so³⁸³⁹.
- 40. On 7 September the Secretary of State for Health and Social Care was asked when publication of these statistics would resume, and stated that no date was decided⁴⁰. This lack of an expected date contradicts the OSR's guidance.
- 41. In October, there was a high-profile public dispute over whether Greater Manchester should enter Tier 3 lockdown. This included public disagreement about the use of selective statistics by Government, and the comparability of ICU occupation numbers with previous years⁴¹.
- 42. Members of the public have not been not provided with data or analysis to help them draw their own conclusions. Members of the press have requested official data on trust-level ICU occupation, but to no avail⁴².

Test and trace performance metrics

- 43. NHS Test and Trace (NHSTT) does not report various types of basic performance data, including the availability of tests by region⁴³. Data is also heavily lagged, typically by around two weeks.
- 44. This makes it difficult for the public to evaluate the case for local lockdowns, since case numbers in each region may be affected by the availability of tests. Researchers from University College London have highlighted this and other concerns, including timeliness: "[we] emphasise the need for real-time information to be prioritised to inform and support the necessary responses, including regional or local lockdowns".
- 45. It is unclear whether NHSTT does not capture this basic operational metric internally, or whether it merely chooses not to report it.

³⁷ https://www.england.nhs.uk/statistics/covid-19-and-the-production-of-statistics/

https://www.england.nhs.uk/statistics/covid-19-and-the-production-of-statistics/

https://osr.statisticsauthority.gov.uk/wp-content/uploads/2020/07/Regulatory-guidance_changing-methods_Coronavirus.pdf

40 https://guestions-statements.parliament.uk/written-guestions/detail/2020-08-28/81520

⁴¹ https://www.greatermanchester-ca.gov.uk/news/statement-from-andy-burnham-and-sir-richard-leese/

https://www.manchestereveningnews.co.uk/news/greater-manchester-news/secrecy-spin-surrounding-greater-manchesters-19131905

⁴³ https://www.gov.uk/government/collections/nhs-test-and-trace-statistics-england-weekly-reports

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

- 46. We welcome the transfer of responsibility for government use of data from DCMS to GDS⁴⁴. However, the problems we have highlighted here show that systemic reforms are needed.
- 47. The problems span weaknesses in statistical reporting, operational data collection, access to third-party data, and performance reporting. We suggest reforms in each area to support future decision making.

OSR's role and gaps in statistical data

- 48. The Office for Statistics Regulation (OSR) is a statutory independent regulator to ensure that statistics are produced and disseminated in the public good⁴⁵. This includes relevant statistics being collected to meet the needs of the public. A previous report by the Committee has suggested that to meet its remit, OSR needs to fill data gaps and engage with a wider range of users⁴⁶.
- 49. We welcome OSR's new guidance that when Ministers refer to management information publicly, it should be published⁴⁷. However, OSR remains primarily reactive rather than proactive in its interventions and its approach to data gaps. For example, to our knowledge, OSR has not intervened on the issue around ICU statistics that we identify above.
- 50. The OSR's current strategy does not contain details of how it plans to engage a broader range of stakeholders, or of how it will change its practices to address the problems identified in the Committee's previous inquiry⁴⁸.
- 51. To become more proactive in identifying issues, OSR will need more resources and potentially more powers. There is a corresponding need for continuing expert Parliamentary scrutiny to ensure those resources are well-used, and that the OSR carries out its independent regulatory role.

The equivalent role on the governance and reporting of operational data

52. The OSR's remit is statistics. There is currently no equivalent institution to ensure that government data collection and sharing supports the public good.

https://osr.statisticsauthority.gov.uk/news/covid-19-production-and-use-of-management-information-by-government-and-other-official-bodies/

⁴⁴ https://questions-statements.parliament.uk/written-statements/detail/2020-07-22/HCWS417

⁴⁵ https://www.statisticsauthority.gov.uk/ https://www.legislation.gov.uk/ukpga/2007/18/contents

https://publications.parliament.uk/pa/cm201719/cmselect/cmpubadm/1820/182002.htm

https://uksa.statisticsauthority.gov.uk/about-the-authority/strategy-and-business-plan

- 53. Some of the problems we highlight such as failures to share details of shielding lists, or to report data from the financial support schemes could have been addressed by early data-focussed interventions.
- 54. With relation to reporting performance data, the Government Digital Service's (GDS) service standard requires new digital services to publish key performance indicators⁴⁹. However, in practice, GDS has struggled to enforce this.
- 55. The recent National Data Strategy states that the ONS and Cabinet Office will work together to create a Data Quality Framework for high-quality data, supported by a new unit at the ONS, the Government Data Quality Hub (DQH)⁵⁰. The remit for the DQH is not yet public. However, there is a need for an institution to support the collection, governance, and sharing of data to support the public good (as well as its technical quality) along the lines of the OSR's remit for statistics.
- 56. Such an institution would need deep digital as well as data expertise, statutory independence, and a proactive approach. It would also need to make careful review of the need for permanent reforms to data practices, versus measures more appropriate for civil contingencies.

Controls over data generated by third parties

- 57. A number of the problems identified here relate to data generated by commercial contractors, including Pillar 2 testing and NHSTT data. We recommend updating Cabinet Office guidance to ensure that Government retains full control of analytic data when commissioning services from contractors⁵¹.
- 58. We also recommend that the Freedom of Information Act should be extended to contractors providing public services, as recommended by the Campaign for Freedom of Information and many others⁵².
- 59. We also recommend considering new general legislation that reporting data on all government spending (grants, loans, tax waivers and contracts) above a *de minimis* should be automatically published.

Embedding data in the policymaking process

60. Some of the problems we have identified were caused by a failure to capture basic performance data, such as NHS Test and Trace's failure to report regional test availability.

10

⁴⁹ https://www.gov.uk/service-manual/service-standard/point-10-define-success-publish-performance-data

⁵⁰ https://www.gov.uk/government/publications/uk-national-data-strategy/national-data-strategy

⁵¹ The Mayor of London's Office of Technology and Innovation has published sample wording for contracts to ensure Government can access analytic data: https://loti.london/resources/tender-wording-data-access-api

⁵² https://www.cfoi.org.uk/campaigns/extending-foi-to-contractors/

- 61. We recommend that the Government's formal policymaking guidance for civil servants should be updated for the digital age, to ensure that basic metrics are always defined when a new service is initiated.
- 62. These should include basic metrics on how well a service is meeting the needs of users and meeting its broader goals, broken down by protected characteristics to ensure compliance with the Equalities Act 2010. These metrics can then be captured from the outset.
- 63. Government should, as a matter of urgency, strengthen the guidance on capturing and reporting metrics in the Cabinet Office's formal pre-legislative guidance for the Civil Service; the Government Digital Service's guidance on collecting data; and the Treasury's Magenta Book and other spending-related guidance⁵³.

Parliament's role

64. Finally, the Committee may wish to explore how Parliament could take a data-focussed approach to examining draft legislation: for example, ensuring that Acts introducing significant new powers or policy changes (such as the Coronavirus Act) contain formal reporting duties.

_

⁵³ https://www.gov.uk/guidance/make-better-use-of-data