



COVID-19 Data Issues

LOTI has collated feedback on issues London boroughs are experiencing with the COVID-19 data sets they receive from Public Health England the NHS for outbreak prevention and control and for Test and Trace. The issues here relate mainly to restricted data sets, since the public PHE Daily Coronavirus Dashboard has a comprehensive API.

Beyond the specific issues outlined below, the overarching concern is that the complexities in accessing and then bringing together the multiple data sets issued by PHE and NHS are delaying councils' ability to respond to their residents' needs.

For more information about the issues highlighted in this document, please contact Jay Saggar (jay.sagg@loti.london) and Onyeka Onyekwelu (onyeka.onyekwelu@loti.london).

PHE published Line List data

Published to PHE Power BI Dashboard.

Data Set Quality

- **Problem:** The dataset has missing or limited meta-data (e.g. concerning what the data set is for, how it was collected, its accuracy, etc.). This makes it hard to reconcile with other data sources such as LSAT and the publicly available [Daily COVID-19 Dashboard](#).
- **Suggested Solution:** Urgent attention is needed to improve the meta data for these data sets.

Automating Data Access

- **Problem:** Boroughs report that automating the data pull from PowerBI is difficult / not possible. The current system forces the use of spreadsheets increase the risk of error.
- **Suggested Solution:** Enable data to be published in a format which boroughs can pull using an automated process.

Shielding Data

This data was published by MHCLG but has now been handed over to NHS Digital. The data is accessed via NHS SEFT portal.

Automating Data access / Manual Process

- **Problem:** Accessing this data still requires a time-consuming manual process that increases the chances of data transformation errors and that could cause a bottleneck for accessing vital data.

- **Suggested Solution:** Can an API or even a FTP transfer system be set up to pull the data automatically when updates are published? Commitment not to change data format and to publish accompanying guidance to support data teams interpreting the change..

Alignment with legacy shielding data (GDS SPL)

- **Problem:** Councils have built up a more comprehensive picture of shielded individuals through the original shielding period. Since the transfer of publishing responsibility shifted to the NHS, it has become difficult to match current records up with the legacy GDS Shielding Persons List (SPL). Following internal analyses in multiple boroughs, it is evident that the NHS SPL data misses out individuals from the original lists. *NOTE: the London boroughs of Ealing & Hackney have raised this with the NHS data publishers already.*
- **Suggested Solution:** 1) Arrange a discussion between NHS and boroughs to understand the methodology behind NHS SPL so that councils can be confident that those missing from the list have been removed for a valid reason and not a data error. 2) Put in place a mechanism to verify individual cases for the same purpose. 3) Establish a contact point at NHS to raise future issues.

Data Quality

- **Problem:** As with the GDS SPL boroughs are reporting data quality issues including very old contact details/phone numbers, people included on basis of out of data medical information and deceased individuals.
- **Suggested Solution:** Publish further guidance on shielding list methodology and engage with LAs to improve data quality in subsequent publications.

New Risk Stratification Tool

- **Problem:** Boroughs have had limited information on the new tool including how it will identify vulnerability, how it will categories people, how and when it will be deployed.
- **Suggested Solution:** Full borough briefing so they can prepare for its roll out.

Test and Trace

Data passed to LAs via NHS Test and Trace (formerly CTAS).

Data Quality

- **Problem:** Councils report that the data has quality issues (e.g. incomplete addresses, missing or incorrect phone numbers and emails), though some of this is to be expected as the contact data is for those people that the central trace system could not find.

- **Suggested Solution:** Set up a conversation with boroughs to discuss minimum data fields and how these can be effectively collected.

Data Governance

- **Problem:** Councils lack clarity on how they can use Test and Trace data. For example, are they allowed to link it with existing data in the council - such as individuals known to adult social care - to provide targeted support for those asked to quarantine
- **Suggested Solution:** Provide councils with clarity on the range of uses to which this data can be put.

Automating Data Access

- **Manual Process:** Data download requires a manual process.
- **Suggestion:** Can an API or even a FTP transfer system be set up to pull the data automatically when updates are published?

Access Control and Security

- **Problem:** The security requirements for accessing this data set make it difficult for council officers to access the data while working from home (as many have to do). This includes the requirement to have a whitelisted IP address to access the data. Additionally, only a limited number of council officers can download the data, creating a bottleneck in the process.
- **Suggested Solution:** 1. Review security requirements to ensure they are proportional to risk and do not unnecessarily delay council officers' test and trace work. 2. Allow more council officers access to this data set.