The NCRM wayfinder guide to Covid-19 secondary data resources

While the Covid-19 pandemic has led to many changes in primary social research practices, adaptations made by providers and producers of data have also altered the way in which secondary analysis of existing data can be conducted. In particular, the push for faster, easier, and increased access to Covid-19 datasets, and the addition of new survey modules and questions, has led to a large number of resources becoming available to those wishing to use secondary analysis to study the pandemic. This guide provides readers with a brief overview of these resources, where they can be found, and important points to consider when accessing and using them.

Why use secondary data on Covid-19?

Secondary analysis allows researchers to access and use data that they may not have the time, ethical clearance, or resources to collect themselves. This is particularly relevant during the Covid-19 pandemic, where social distancing measures have put limits on fieldwork and primary data collection. Existing data allows researchers to continue to conduct research even under Covid-19 restrictions, while adaptation to existing surveys allow established research projects to continue, while also taking into account the potential effects of Covid-19.

Additionally, many of the topics of interest during this time, such as health and income, are considered sensitive and the approval to conduct primary research on such topics can often be lengthy. In this instance, large scale surveys with pre-existing respondents enable a rapid researcher response. In the case of Understanding Society (which is featured as a case study throughout this guide), data owners were able to ask and collect data on these subjects and make this available to researchers quickly.

What resources are available and where can they be found?

Resources for secondary analysis of Covid-19 data are available in a variety of forms from a variety of different providers.

Survey data

Survey data is one of the most commonly used sources when conducting secondary analysis and many new surveys have been commissioned specifically to study Covid-19. They cover a broad range of topics from infection rates to time use, to social impacts. The ONS Covid Infection Survey for example, was created to investigate the prevalence of COVID-19 infection in the community, while the Evidence for Equality National Survey (EVENS), is designed to collect data on how the pandemic is affecting ethnic and religious minority groups in England, Scotland and Wales.

New surveys are not the only source of survey data on Covid-19. Many data producers have also made use of established samples to adapt pre-existing surveys. Long running surveys such as the Opinions and Lifestyle Survey and the English Longitudinal Study of Ageing (ELSA) have commissioned new modules of questions or sub-studies to assess the impact of the pandemic.

The data from Understanding Society's COVID-19 study was also designed to be linked with answers respondents have given in previous (and future) waves of the annual Understanding Society survey, with the annual 2019 interview data included in the Covid-19 dataset. While this survey data can be used by itself, linking such as this, allows researchers to compare with pre-pandemic responses and explore the longer-term impact of the pandemic (see publications).

Administrative and Government data

Administrative data is another useful resource. ADR (Administrative Data Research) UK facilitates and supports the use of administrative data in the response to Covid-19. Their 'Local Data Spaces' programme used

administrative data to create reports on the impact of Covid-19 for all local authorities in England. These reports are free to access and download.

Governmental sources such as the Office for National Statistics have also produced publicly available Covid-19 data, as well as regular statistical bulletins summarising these. The data are available to download in aggregated forms and they cover topics such as rates of hospitalisation and death, vaccination hesitancy, rates of self-isolation, travel, homeworking, and social impacts.

Visualisations

Throughout the pandemic data visualisation has been utilised to aid comprehension of the spread and impact of Covid-19. It has helped complex data to be understood by wider audiences and it provides a useful starting point for researchers wishing to explore patterns and trends.

The UK government's interactive COVID-19 map, for example shows cases and vaccination data by local authority and can be used to look at the changing pattern over time. Simple summaries of the data can be created, and images of the map are available to download. The BBC also provide up-to-date visualisations of cases, deaths and vaccinations by country, and Worldometer keeps a constant running total of cases, deaths and recoveries worldwide.

Data pages, lists, and hubs

Central lists and data hubs are another resource for exploring Covid-19 data. CLOSER's COVID-19 Longitudinal Research Hub brings together longitudinal survey data, research reports and expert opinions together in one central location. The UK Data Service's COVID-19 theme page provides a list of key Covid-19 survey datasets that can be accessed through their catalogue. The theme page for Understanding Society COVID-19 longitudinal data provides information on the study and links to the main parent study. This enables comparisons to earlier answers and allows researchers to explore the longer-term impact of Covid-19. The UN COVID-19 Data Hub makes Covid-19 response data from many individual nations available as geospatial data, designed to be easily downloaded in multiple formats. Central lists and hubs are an excellent starting point for researchers interested in exploring certain data types, or groups of resources.

Using these resources

There are a few considerations that researchers must make when using secondary data on Covid-19.

Changing methodologies

The Covid-19 pandemic has necessitated a change of methodologies to allow data collection to continue in this new 'socially-distant' environment. Questionnaires previously administered face-to-face have been adapted into web or telephone surveys to allow for the continuation of data collection. The Understanding Society survey, for example, rapidly transitioned to a protocol without face-to-face interviews, giving all participants the option to complete their interview online or by telephone. However, this can come at a cost, such as reduced sample size or a reduced number of questions. Researchers must ensure they consult and read any accompanying documentation or methodological guides thoroughly to ensure that they consider the impact of these changes on their analyses.

Access

While much of the data made available during the pandemic is open access, others (such as many surveys and data that deal with sensitive topics) have additional access requirements. Data providers may require sign in with an institutional login or have an agreement stating the data will not be shared onwards that must be signed. Some datasets may be restricted for commercial use, or use by undergraduate students, and others can only be accessed in secure settings due to increased disclosure risk. Special license data such as the Understanding Society Covid-19 school codes and census and local authority geographies require an application. Although greater restrictions apply, if approved, analysis can be conducted at a more granular level. It is therefore important researchers check the access requirements for any data they plan to use.

Covid-19 has also altered the data access landscape. Secure datasets (i.e., those datasets with higher possible disclosure risks) could historically only be accessed in secure rooms on designated sites. In line with increased homeworking and restrictions on travel, some secure access services have made it possible for researchers to access secure data from home through 'virtual secure rooms'. In the UK, the new UKRI-ESRC SafePod Network (SPN) is set to become the first

service in the world to provide a nationwide network of standardised safe settings or secure environments (known as SafePods) to enable approved researchers across the UK to use sensitive datasets. Researchers will also be able to use the SafePods to securely access government, study and survey datasets from Administrative Data Research UK (ADR UK) projects, ONS Secure Research Service and SAIL Databank.

Case Study – Understanding Society - the journey and benefits of linking COVID-19 data with longitudinal data

In response to the coronavirus outbreak, Understanding Society, with funding from the Economic and Social Research Council and the Health Foundation initiated a new COVID-19 survey. The research community's rapid response initiative resulted in the Covid-19 survey being announced, along with an invitation (on 23 Apr 2020, one month after the first lockdown) to researchers to contribute 'in the moment' questions for future months. The Covid-19 survey is an integral part of the main Understanding Society survey: also known as the UK Household Longitudinal Study, it is the largest longitudinal household panel survey in the UK. All 42,000 adult sample members who had responded in the last two waves of the main survey were invited to participate in the first wave in April 2020 (released in May 2020).

The survey includes questions on core content, repeated in each wave, plus rotating and additional questions on the impact of the coronavirus on the welfare of UK individuals, families, and wider communities. Blood samples were also collected from respondents, in March 2021, and tested for Covid antibodies. Currently data from eight monthly waves have been released (April 2020 – March 2021). The full content for the Covid-19 study is available in the long-term content plan.

The timely 'in the moment' findings from this study have been used to inform the Government, with evidence to SAGE on vaccine hesitancy and analysis from the Treasury on the impact on working households' income. Researchers have been able to gain quick access to these data via the UK Data Service. Research findings have provided researchers with vital evidence on the changing impact of the pandemic, covering many different topics. Topic based briefing notes on the Covid-19 theme page show how the data can be used.

The data has also provided an entry point for those new to longitudinal data, introducing both new and existing researchers to the benefits of longitudinal and secondary data use.

Useful links

UK Data Service

Secondary Data Analysis – UK Data Service

Covid-19 | Understanding Society

COVID-19 — UK Data Service

COVID-19: Social surveys are now more important than ever – UK Data Service Blog

What is the UK Data Service SecureLab? — UK Data Service

How Understanding Society: The UK Household Longitudinal Study adapted to the COVID-19 pandemic
DOI: https://doi.org/10.18148/srm/2020.v14i2.7746

This guide was produced in 2021 by Alle Bloom (UK Data Service, University of Manchester) and Annette Pasotti (Understanding Society, University of Essex) as part of a series produced for the Changing Research Methods for Covid-19 Research Project.

National Centre for Research Methods
Social Sciences
University of Southampton
Southampton, SO17 1BJ
United Kingdom.

Web http://www.ncrm.ac.uk **Tel** +44 23 8059 4539
Email info@ncrm.ac.uk? **Twitter** @NCRMUK