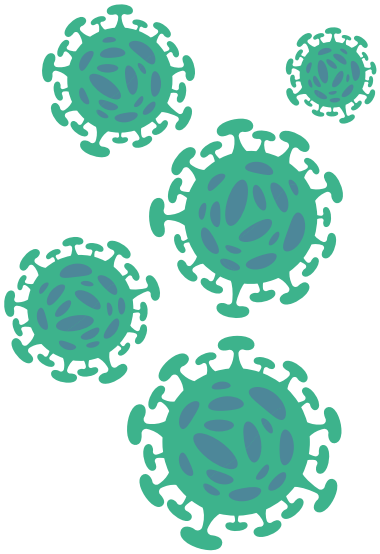
HDRUK
Health Data Research UK

Data & Connectivity National Core Study

What did we achieve in July-September 2021 (Sprints 10-13)

Making COVID-19 datasets available for research

- **92** (+4 from sprint 9) National Core Study priority datasets listed on the HDRUK Innovation Gateway with detailed technical metadata for 87 (+1 from Sprint 9)
- **COVID-19 viral variant genomic data** now available for the four UK nations. **Full viral sequence data for Scottish and English populations** will imminently be available in the Edinburgh Parallel Computing Centre Trusted Research Environment (TRE) and in the Welsh SAIL databank TRE for the Welsh population.
- Progress continues with development of **Outbreak Data Analysis Platform**. This project is linking the key data analysis systems that have supported the UK-wide observational studies of COVID-19: the CLIMB COVID viral sequence analysis system, and the current Outbreak Analysis Platform in Edinburgh. This project will provide vital insights into viral variation on disease severity, vaccine efficacy, and transmission and will be a vital resource for **future pandemic preparedness.**
- **Progress to enable regional, rapid, near real time hospital admissions data to support vaccine safety research continues.** CIPHA, PIONEER data hub, and Discovery (Barts Health NHS Trust) are working towards enabling access to more rapid, granular and timely data (including laboratory data flows) that are not currently available in nationally data feeds. Initial work has resulted in a curated data set across all three regions, and analytical protocol developed which will be implemented across all 3 regions. A clinical validation workshop was held with clinical haematology experts to validate our approach. Plans to scale up across wider English regions and devolved nations and identify use cases beyond COVID underway.
- **The NCS Admin and Health Data Linkage group continues to support the Longitudinal Health and Wellbeing and Transmission and Environment National Core Studies with high priority administrative and health linkage needs.** This cross NCS group has supported an application to link HMRC data to UK Longitudinal Linkage Collaboration study cohort data, identify other UK wide admin data sets and linkages that would support the NCS and aims to raise awareness of the NCS data and connectivity priorities across Government departments.



HDRUK
Health Data Research UK

Data discovery, access, and user experience

- We launched a **£2 million joint research funding call with the Alan Turing Institute** to support research that creates new knowledge and expand the research outputs of the National Core Studies with an advanced data analytics focus. **Nine studies were awarded funding** – they will use large-scale linked data to address priority research questions that will improve understanding of the pandemic and inform the continued policy response. Read the press release here to learn more about each of the studies.
- We have created a brochure for new researchers to explore available data assets: Data available for COVID-19 research across the UK
- Progress continues to **harmonise the data access request process** via the HDRUK Innovation Gateway across the UK network of Triste d
- **Data Use Register:** a register of approved data uses for COVID-19 research TREs will be made available via the Innovation Gateway shortly.
- Progress continues to build an analytical federation across the 5 UK TREs.

HDRUK
Health Data Research UK

Impacts and emerging insights

To date, there have been 1,320 COVID pre-prints using health data and 237 published papers. A selection of key outputs are listed below:

- A self-controlled case series study exploring risk of thrombocytopenia and thromboembolism after covid-19 vaccination and SARS-CoV-2 positive testing found that increased risks of haematological and vascular events were observed for short time intervals after first doses of the Oxford-AstraZeneca and Pfizer-BioNTech vaccines. The risks of most of these events were substantially higher and more prolonged after SARS-CoV-2 infection than after vaccination in the same population.
- An analysis into neurological complications after first dose of COVID-19 vaccines and SARS-CoV-2 infection found that vaccination leads to increased risk of some neurological conditions but confirm the risk of complications is greater following infection with COVID-19.
- Analyses of linked health data from 2.53 million Scottish adults derived from the Early Pandemic Evaluation and Enhanced Surveillance of Covid-19 (EAVE II) database revealed that a first dose of the Oxford–AstraZeneca vaccine, but not Pfizer-BioNTech, is associated with a small increased risk of an autoimmune bleeding disorder (idiopathic thrombocytopenic purpura). The small increase in risk is similar to figures for flu and MMR vaccines, and should be balanced with benefits of vaccination.
- A retrospective cohort study using linked national Census, electronic health records and mortality data for >12 million adults in England found that in people aged 40 and over obesity is a greater risk factor for COVID-19 death in ethnic minorities.
- A cohort study using linked national Census, electronic health records and mortality data for >12 million adults in England found that disparities in rates of COVID-19 vaccination uptake among elderly adults existed within specific ethnic and religious groups, with lowest rates observed in black African and black Caribbean ethnic backgrounds.

HDRUK
Health Data Research UK

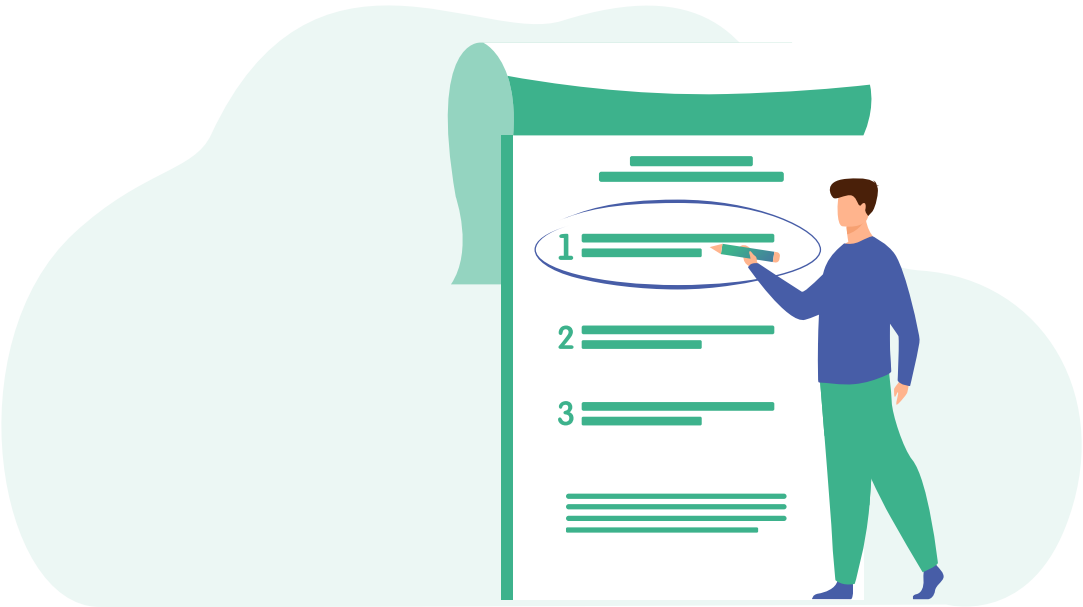
Communicating and engaging with partners, patients, and the public

- HDRUK hosted a webinar 'Harnessing the benefit of COVID-19 data for public benefit' to showcase available data to new researchers.
- Work continues with patient and public representatives to co-design a template for data custodians to **develop lay summaries of data assets** for the Gateway.
- Planning public involvement and engagement work across the Mersey and Cheshire, NE London and West Midlands regions to input to our work to **enable regional, rapid and more granular hospital admissions data flows** to support vaccine safety research.
- Scoping work underway to plan public involvement and engagement activity to support the **NCS Admin and Health Data Linkage Group**.
- Developed **best practice guidelines to include PPIE in funding calls**. Document is now finalised and available here.

HDRUK
Health Data Research UK

Emerging Priorities and next steps

- Continue to progress further linkage and approved researcher access to **high value COVID data assets, such as viral genomic data, intensive care data, and data from the REACT study, and clinical trial data** to:
  - Enable a deeper understanding of the impact of viral and host genomics on susceptibility and immunity to COVID-19 disease, severity, and outcomes.
  - Enable further analyses which will help inform future research and pandemic policy
- Developing **standardised protocols across several regions** to enable **rapid, acute admissions data flows and linkages** to support vaccine research and surveillance.
- Developing a roadmap to enable **wider linkage of priority health and administrative data** sets and conducting **joint public involvement and engagement** work with ADRUK in this area.
- Progressing the implementation of the **cohort discovery tool** across our four nation delivery partners.
- Conducting more in-depth user surveys of researchers working in Trusted Research Environments to **further inform TRE improvements and enhancements**.
- Progressing further **harmonisation of the Data Access Request process** across the four nations



Glossary

HDRUK - Health Data Research UK

ADRUK - Administrative Data Research UK

COG-UK - Covid Genomic Consortium UK

NCS - National Core Study

TRE - Trusted Research Environment

ONS SRS - Office for National Statistics Secure Research Service

SAIL - Secure Anonymised Information Linkage

PPIE - Public and Patient Involvement and Engagement