



COVID-19 Health Data Research

9 February 2021 - Fortnightly update for SAGE, National Core Studies & UKRI/DHSC

Authors:

Alice Turnbull, Health Data Research UK
Andrew Morris, Health Data Research UK
Ben Gordon, Health Data Research UK
Carole Morris, Public Health Scotland
Caroline Cake, Health Data Research UK (lead)
Cathie Sudlow, BHF Data Science Centre
Charlie Davie, DATA-CAN
Sophie Morris, Health Data Research UK
Hollydawn Murray, Health Data Research UK
James Pickett, Health Data Research UK
Lara Edwards, Health Data Research UK
David Seymour, UK Health Data Research Alliance
John Aston, Home Office (SAGE sponsor)

John Deanfield, NICOR
Mark Parsons, Scotland National Safe Haven
Melissa Lewis-Brown, Health Data Research UK
Members of the HDR UK Public Advisory Board & COVID-19 PPIE Group
Ming Tang, NHS England and Improvement
National Core Study Programme Leads
Nilesh Samani, British Heart Foundation
Pete Stokes, Office for National Statistics
Ronan Lyons, SAIL Databank (UKRI/DHSC sponsor)
Garry Coleman, NHS Digital
Ian Young, Health & Social Care Northern Ireland



Health data research community highlights this period*



29 COVID-19 taskforce calls with **183** clinical and health data research leaders engaged



1,168 COVID-19 pre-print publications



751 academic, industry and NHS participants in COVID-19 Slack channel with 10 sub-channels



111 health data research questions identified



This Week's Patient and Public Voice Feedback:

“Use and analysis of vaccine programme data remains of the utmost importance. Urgent research is needed to:

- understand and evaluate initiatives to overcome vaccine hesitancy and low vaccine coverage particularly in ethnic minority populations and communities in deprived areas
- understand the effectiveness of vaccines in clinical extremely vulnerable populations but with granularity given this covers a huge array of conditions”

Click [here](#) to read more feedback



Click [here](#) for a list of regularly updated COVID-19 research questions from the health data community

Research topics with new insights generated in last 2 weeks

Health data research on COVID-19 continues to grow, now reaching 1168 (+5) non-peer-reviewed pre-prints & 98 (+5) published papers.

Topic	Insights from ongoing studies (links provide further details):
Surveillance & Epidemiology	<ul style="list-style-type: none"> Four devices used for COVID-19 antibody tests were compared against each other by running the same c4500 samples through them. All devices displayed a trade-off between specificity and sensitivity – such as reduced sensitivity at lower antibody levels – and this could correspond to 5% false positive rates on the most specific device. A novel algorithm to predict COVID-19 deaths in England (QCovid) developed using linked datasets from general practice, death registry and hospital episode data, has been externally validated in a population-based cohort study. The algorithm, which showed high levels of discrimination between men and women, has potential to support public policy and clinical risk assessments.
Immunity & Vaccines	<ul style="list-style-type: none"> According to data collected using the ZOE app from nearly 40,000 people who received the Pfizer/BioNTech vaccine in December, people who have previously had COVID-19 are almost twice as likely to experience mild systemic after-effects such as fatigue and headache compared to people who have not had COVID-19. Analyses of linked patient-level data on the OpenSafely platform indicate lower vaccination coverage in people over 80 with severe mental illness, dementia, and learning disabilities – as well as lower coverage in people from ethnic backgrounds and people from the most deprived communities. Real-world data from 503,875 individuals in Israel suggest the 1st dose of the Pfizer/BioNTech is 51% effective 13-24 days after immunisation. Recent results from the ONS Opinions and Lifestyle Survey found that 49% of people surveyed from the Black or Black British ethnic group would be very or fairly likely to have the COVID-19 vaccine if offered. This is the lowest proportion amongst ethnic groups surveyed including White (85%), Mixed (80%), and Asian or Asian British (72%).
Longitudinal health & wellbeing	<ul style="list-style-type: none"> A population-based study drawing on data from all 52 English NHS radiotherapy providers identified changes to radiotherapy services during the first wave of COVID-19 including: treatment delays where safe to do so, shorter radiotherapy courses, and increased radiotherapy to compensate for reduction in surgeries. The overall number of patients receiving radiotherapy fell significantly and remained reduced in June pointing to reductions in cancer referrals and diagnoses.
Transmission & Environment	<ul style="list-style-type: none"> Early data from the VirusWatch Household Study found that people living in the most deprived areas of England and Wales were more likely to leave their house for work or school, use public transport and visit an essential shop compared to people living in the least deprived areas – suggesting that differences in daily essential activities may be contributing to the disproportionate infection rates in deprived areas. A representative, household survey of SARS-CoV-2 infection amongst a UK strictly Orthodox Jewish population revealed an infection rate of 64% (5 times national estimates). This study adds to the growing body of evidence demonstrating a disproportionate impact of COVID-19 on minority groups. Using a combination of crowd-sourced app data and existing geo-statistical techniques within a safe data environment, researchers are now able to predict COVID-19 hotspots at fine geographic scales in the UK. This is particularly relevant given that emerging risks such as deprivation may vary over small distances.
Clinical Trials	<ul style="list-style-type: none"> Interim data from the PRINCIPLE trial (Platform Randomised trial of Interventions against COVID-19 In older peoPLE) show that two commonly used antibiotics, azithromycin and doxycycline, have no beneficial effect in patients over the age of 50 in the early stages of COVID-19.

Data & Connectivity National Core Study: COVID-19 dataset availability – 9 February

Vaccine data now flowing into the ONS SRS and the Welsh SAIL databank, to be made available to researchers shortly.

Core COVID-19 Datasets available for linkage	Office for National Statistics Secure Research Service	England (NHS Digital Data Processing Service)	Scotland (National Data Safe Haven)	Wales (SAIL Databank)	Northern Ireland (Honest Broker Service)
C-19 vaccine data collection	Currently limited use only	Acting as data processor	Data to be transferred soon	Welsh Immunisation Service (WIS) data now in SAIL	Data flow and IG to be agreed
COG-UK viral genome		Finalising governance	Subset linked to CO-CIN data a/w IG sign off for remainder	Governance agreed, to be linked shortly	Governance agreed, to be linked shortly
Pillar 1 COVID-19 Testing Data					
Pillar 2 Testing data (UK Gov)					Missing results prior to 26 Apr – Data quality issue
Primary Care	GPES linked to census, mortality and hospital data for internal access only	GPES extract – 98% practice coverage, large subset of codes (4bn items) Community Prescribing	Albasoft ESCRO GP Extraction* Prescribing Information System	80%+ coverage of full longitudinal record, with 100% coverage for COVID codes	Enhanced Prescribing Database as proxy
Secondary Care	HES – available for internal access only 100% coverage	100% coverage - HES SUS via DARS extract only	100% coverage	100% coverage	
Personal Demographic Service	100% coverage	100% coverage (via DARS extract only)	100% coverage	100% coverage	
Death registry	Provisional Monthly Extract & Linked Census and death occurrence	100% coverage	100% coverage	100% coverage	
C-19 Infection Survey (CIS)		N/A	Awaiting DEA accreditation	Awaiting decision on data access	Awaiting decision on data access
COVID-19 Clinical Information Network (CO-CIN)	Being linked to 2011 census	English CO-CIN data available in Scottish Nation Data Safe Haven	Limited metadata	Awaiting decision on data access	
Census 2011	Household structure	N/A			N/A
Covid Opinions Survey		N/A	N/A	Awaiting decision on data access	N/A
Business Impact of Covid Survey	c. 5,000 businesses	N/A	N/A	Awaiting decision on data access	N/A
Labour Force Survey	40,000 households, 100,000 individuals	N/A	N/A	Awaiting decision on data access	N/A
Intensive Care data	HES Critical Care, ICNARC	Reviewing IG to share ICNARC. HES Critical care to be onboarded to TRE	SICSAG (updated weekly)	ICNARC COVID weekly, ICNARC quarterly all admissions and critical care routine data (CCDS) monthly	3 datasets being explored
Pillar 3 Testing data (NHS labs)	Available in CIS data	Internal use only			
Pillar 3 Testing data (iELISA)	N/A	Available via DARS extract only			Data to be validated
Other Pillar 4 Testing data	VIVALDI, REACT II				
ZOE Symptom Study App Data				UK wide (unlinked) Wales (linked)	



KEY

1. Custodian engagement

2. Dataset available in secure Trusted Research Environment

3. Linkages established to other priority datasets (within TRE)

4. Datasets available for COVID-19 research via Gateway

Further information about **Data & Connectivity** can be found [here](#). Including the [Sprint 3 report](#).

New [Trusted Research Environment collection](#) added to Gateway collections.

Status of COVID-19 projects using the data – 9 February 2021

39 additions to active research takes total over 290. Most new projects enabled by the SAIL Databank in Wales. 15 data access requests covering 63 data assets now submitted via the Gateway for the 12 rapid funded projects.

# of COVID-19 Projects by stage (change from previous report)	Office for National Statistics Secure Research Service	England (NHS Digital Data Processing Service)	Scotland (National Data Safe Haven)	Wales (SAIL Databank)	Northern Ireland (Honest Broker Service)	Total
In development	11 (-6)	32 (+1)	40 (+4)	90 (-)	6 (-)	179 (-1)
- a/w researcher	11 (+1)	25(+6)	Not available	55 (-1)	6 (-)	N/A
- a/w data custodian	0 (-3)	7(-5)	Not available	35 (+1)	0 (-)	N/A
Submitted for Information Governance approval	4(+1)	7 (-1)	5(+1)	0 (-1)	0 (-)	16 (-)
Approved but not yet active	2 (-4)	3 (+2)	9 (-)	3 (+1)	1(-)	18 (-1)
Active research taking place	15 (+4)	87 (+3)	69 (+2)	121 (+30)	1(-)	293 (+39)
Active Number of Researchers	Not yet available	Not yet available	Not yet available	205 (+3)	Not yet available	
Average time from application to active research	Not yet available	Not yet available	Not yet available	3 days	Not yet available	

Participation in key UK wide studies:

- [PRINCIPLE](#): 4,053 participants (+8% in last 2 weeks, with continuing data flow of Pillar 2 COVID +ve test results to support recruitment)
- [RECOVERY](#): 35,672 participants across 178 active sites (+11% in last 2 weeks, and +1 active site)
- [CO-CIN \(ISARIC 4C\)](#)
 - 176,413 Tier 0 (case report) (+9% in last 2 weeks)
 - 650 Tier 1 (single sample)
 - 1,658 Tier 2 (serial sampling)
- [GENOMICCC](#): 10358 participants (+12% in last 2 weeks) across 212 ICUs with a total of 5091 intensive care beds
- [COVID-19 ZOE symptom study](#): 4,582,772
- [COG-UK](#): 250,673 viral genomes sequenced (+17% in last 2 weeks)

Data & Connectivity National Core Study projects now underway

- 15 data access requests covering 63 data assets from 12 studies now submitted via [HDRUK's Innovation Gateway](#). Examples include:
- [Is exposure to airborne fine and ultrafine particulate matter a determining factor in COVID-19 infection and outcome within the UK?](#) led by Professor Kevin Wyche, University of Brighton
- [Enhancing the Utilisation of COVID-19 Testing in Schools Studies: The Joint Analysis of the ONS COVID-19 School Infection Survey and COVID-19 Mapping and Mitigation in Schools \(CoMMInS\) Study](#) led by Dr Rachel Denholm, University of Bristol
- [What are the relative contributions of different exposures and settings to COVID-19 community transmission? Analysis of community cohort studies linked to national testing data](#) led by Professor Andrew Hayward, UCL

Data Access Registers

For more information on the active projects:

- [ONS Secure Research Service](#): List of accredited researchers and research projects under the Research Strand of the Digital Economy Act)
- [NHS Digital](#): Register of approved data releases (includes all access)
- [Scotland](#): Public Benefit and Privacy Panel approvals
- [SAIL Databank](#): COVID-19 projects listed on gateway
- [NI Honest Broker Service](#): Projects currently being carried out.