COVID-19 Health Data Research

6 April 2021 - Fortnightly update for SAGE, National Core Studies & UKRI/DHSC

Authors:

Alice Turnbull, Health Data Research UK
Andrew Morris, Health Data Research UK
David Seymour, UK Health Data Research Alliance
Caroline Cake, Health Data Research UK (lead)
Sinduja Manohar, Health Data Research UK
Susheel Varma, Health Data Research UK
Cathie Sudlow, BHF Data Science Centre
Sophie Morris, Health Data Research UK
Ashleigh Smith, Health Data Research UK
Hollydawn Murray, Health Data Research UK
James Pickett, Health Data Research UK
Lara Edwards, Health Data Research UK
Melissa Lewis-Brown, Health Data Research UK

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

Health data research community highlights this period

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

1,204 COVID-19 pre-print publications

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

111 health data research questions identified

Patient and Public Voice Feedback:

"It's great to see the funded research projects progressing and 'leaving the starting blocks' as more data are being made available. The three highlighted in this report will help to provide the much needed inclusion criteria of what symptoms and clinical observations can be classified as Long Covid, as well as focusing on vaccine uptake, efficacy and safety and further understanding the differences of C-19 impact in men and women."

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 outputs on COVID-19 continues to grow, now reaching 1,204 (+11) non-peer-reviewed pre-prints & 131 (+6) published papers.

HDRUK
Health Data Research UK

Topic	Insights from ongoing studies (links provide further details):
Surveillance & Epidemiology	• A population-based cohort study involving >15,000 participants (COVIDENCE UK) found that people from Asian/Asian British ethnicities and people with raised body mass index have an increased risk of developing COVID-19 – whilst people with allergic diseases such as eczema and hayfever, have a decreased risk. • Using data from nearly 7,000 participants in UK Biobank, researchers have shown that older *biological* age is associated with a greater risk of COVID-19 hospitalisation and death. Biological age may be older or younger than a person's chronological age, and was predicted using telomere length in this instance. • A model,developed using health records from 2,815 COVID-19 inpatients from Wuhan, predicts patient outcomes at the point of COVID-19 hospitalisation, helped clinicians identify appropriate treatments at admission, with applications for long COVID. The model will benefit from being tested in larger datasets and different cohorts. • A study using laboratory and clinical data from hospitals in London and Oslo to validate published COVID-19 prediction models for hospitalised patients, found varying performance between the 2 sites – underscoring the need for local/regional/national recalibration of models and validation across multiple locations to improve accuracy.
Immunity & Vaccines	• Analyses of antibody tests linked to national SARS-CoV-2 testing programme data & vaccination records from >10,000 residents of 310 long term care facilities in England as part of the VIVALDI study, indicate that a 1st dose of the Pfizer-BioNTech or Oxford-AstraZeneca vaccine substantially reduces the risk of COVID-19 infection in older adults. Closely related, an observational study of >14,000 vaccinated older care home residents in Wales, combining electronic health records and administrative data from the SAIL databank, found a small number of residents were infected with COVID-19 following vaccination; these infections were associated with frailty and usually occurred within 28 days of vaccination. These findings are relevant for policymakers prioritising 2nd vaccine doses or revisiting the control measures in long term care facilities. • Recent research combining Canadian demographic, epidemiological, economic and Prospective Urban Rural Epidemiology (PURE) cohort data, estimates that those with obesity, diabetes or hypertension are at higher risk of death compared to individuals without these risk factors, therefore prioritising individuals with obesity, diabetes, or hypertension may be an efficient way to prevent deaths. • Survey response data from >20,000 adults as part of the VirusWatch household study indicate that 86% of participants who were reluctant or intending to refuse a COVID-19 vaccine in December were planning on accepting (or already had accepted) a vaccine in February – and these findings were consistent across ethnic and social groups. This shift in attitude highlights the need to offer vaccines repeatedly as people change their minds over time.
Longitudinal health & wellbeing	• A social media survey of >2,500 people, co-produced with patients, found that long COVID affects people's daily life, including their mental health, ability to work and do domestic chores. For the majority, symptoms fluctuate or relapse, and common triggers include physical activity, stress, sleep disturbance and cognitive activity. • Patient reported outcomes from >1,000 patients discharged from hospital following treatment for COVID-19 as part of the UK multicentre PHOSP-COVID study, revealed most survivors are not fully recovered five months after discharge; ~1 in 5 people developed a new disability, stopped working or changed jobs due to their health, and/or experienced symptoms of anxiety or depression. • Analyses of data from a smaller group of participants using International Severe Acute Respiratory and emerging Infections Consortium (ISARIC) WHO Clinical Characterisation Protocol (CCP-UK), indicate that COVID-19 survivors experience long-term symptoms, new disability, increased breathlessness, and reduced quality of life. • Individuals discharged from hospital after COVID-19 had increased rates of multiorgan dysfunction compared with the expected risk in the general population. The increase in risk was not confined to the elderly and was not uniform across ethnicities. The diagnosis, treatment, and prevention of post-covid syndrome requires integrated rather than organ or disease specific approaches, and urgent research is needed to establish the risk factors.
Transmission & Environment	• Ongoing progress.

Data & Connectivity National Core Study: COVID-19 dataset availability – 6 April

Vaccine data now available for Wales and England and will soon be available for Scotland. Viral genome data also soon to be available for Wales.

HDR UK Health Data Research UK

Office for National Statistics

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	To be made available shortly. Accepting applications now	Vaccines Events & Adverse Reactions	Available in TRE, available via Gateway shortly.	COVID Vaccination Dataset	Vaccine Management System Operational – Data to be transferred to Data warehouse
COG-UK viral genome	Awaiting data flow from Public Health England	Awaiting data flow from Public Health England	Subset linked to CO-CIN data Remaining data being transferred, available shortly.	Data flow started from Public Health Wales	Governance agreed, automation of data flow to PHA in progress
Pillar 1 COVID-19 Testing Data	To be linked to Test and Trace data	COVID-19 Second Generation Surveillance System (SGSS)	Electronic Communication of Surveillance in Scotland (ECOSS)	COVID-19 Test Results	COVID antigen testing - Pillar 1
Pillar 2 Testing data (UK Gov)	To be linked to Test and Trace data	COVID-19 UK Non-hospital Antigen Testing Results (Pillar 2)	Electronic Communication of Surveillance in Scotland (ECOSS)	COVID-19 Test Results	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	Census-Mortality-HES linked data asset now available (ONS/NHSD)	100% coverage – HES. SUS via DARS extract only, available in TRE soon	100% coverage	100% coverage	Admissions & Discharges
Personal Demographic Service	Internal use only	100% coverage (via DARS extract only)	100% coverage	100% coverage	
Death registry	Provisional Monthly Extract & Linked Census and death occurrence	100% coverage Civil Registrations - Deaths	100% coverage	100% coverage	
C-19 Infection Survey (CIS)	Linked to Test and Trace data	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	Data for English CO-CIN participants available in Scottish Nation Data Safe Haven	Limited metadata. Includes English linked data, and COGUK/CO-CIN data asset	Awaiting decision on data access	Discussions ongoing to collect data in NI
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	HES Critical Care Reviewing IG to share ICNARC.	SICSAG (updated weekly)	ICNARC COVID weekly, ICNARC quarterly all admissions and critical care routine data (CCDS) monthly	Critical Care Minimum dataset to be acquired
Pillar 3 Testing data (NHS labs)	Captured within Test and Trace data	N/A		COVID-19 Test Results	
Pillar 3 Testing data (iELISA)	N/A	COVID-19 UK Non-hospital Antibody Testing Results (Pillar 3)			Data to be validated
Other Pillar 4 Testing data	VIVALDI, REACT II				
ZOE Symptom Study App Data	Finalising data sharing agreement			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 recently published **Sprint 5 report**.

Data and Connectivity National Core Study webpages and **dashboard** now live

Status of COVID-19 projects using the data – 6 April

17 new research projects now active, taking the total number of active research projects over 320. 10 of 12 rapid funded project now with at least one data request approved (increase of 6 projects since last report).

# 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	6 (-)	40 (+2)	35 (-1)	16 (-4)	3 (-1)	100 (-3)
Submitted for Information Governance approval	1 (-5)	10 (-2)	10 (+3)	1 (+1)	0 (-1)	22 (+2)
Approved but not yet active	1 (-5)	4 (-)	8 (-)	1 (-1)	3 (+1)	17 (-5)
Active research taking place	22 (+6)	100 (+3)	75 (+4)	129 (+4)	1 (-)	327 (+17)
Completed projects	Coming soon	Coming soon	0	0	Coming soon	
Active Number of Researchers	312 (+39)	56 (+4)	232 (+32)	254 (+5)	1	885 (+80)
Average time from application to active research	102 days	Not yet available	10 days	3 days	Not yet available	

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.

Participation in key UK wide studies:
- PRINCIPLE: 4,717 participants (+2% in last 2 weeks, with continuing data flow of Pillar 2 COVID +ve test results to support recruitment)
- RECOVERY: 39,546 participants across 181 active sites (+0.5% in last 2 weeks, and +1 active site)
- CO-CIN (ISARIC 4C)
    - 200,287 Tier 0 (case report) (+3% in last 2 weeks)
    - 645 Tier 1 (single sample)
    - 1,665 Tier 2 (serial sampling)
- GENOMICC: 12,208 participants (+2% in last 2 weeks) across 212 ICUs with a total of 5091 intensive care beds
- COVID-19 ZOE symptom study: 4,641,484
- COG-UK: 413,687 viral genomes sequenced (+15% in last 2 weeks)

Data & Connectivity National Core Study projects now underway. Examples include:

- **Can phenotypes developed from enhanced remote primary care assessment of COVID-19 be used to identify a cohort of community cases, and enable comparison of recovered and long COVID** led by Professor Greenhalgh, University of Oxford
- **How is COVID-19 impacting women and men's working lives in the UK?** led by Professor Tracey Warren, University of Nottingham
- **Uptake and comparative safety of new COVID-19 vaccines by age, sex, region, ethnicity, comorbidities, medication, deprivation, risk level and evidence of prior COVID infection** led by Professor Hippisley-Cox, University of Oxford