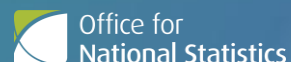


**HDRUK**  
Health Data Research UK



**HSC** Public Health Agency  
Research and Development



# COVID-19 Health Data Research

26 January 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  
Lara Edwards, Health Data Research UK  
David Seymour, UK Health Data Research Alliance  
Ian Young, Health & Social Care Northern Ireland  
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



## Health data research community highlights this period\*



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



**1,163** 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:

The access, use and analysis of data from the vaccine programme roll-out continues to be of great public interest and concern and needs to be a major priority area of work. Meaningful patient and public involvement in this area of research needs to be embedded to build public trust and ensure research questions of importance to the public are addressed to help build and sustain vaccine uptake.

Click [here](#) and to read more feedback



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

\*Following on from the establishment of the National Core Studies Programme we will be tracking the SAGE-endorsed recommendations from November 10, 2020 through this report going forward.

# Research topics with new insights generated in last 2 weeks

Health data research on COVID-19 continues to grow, now reaching 1163 (+10) non peer-reviewed pre-prints & 93 (+6) published papers.

Topic	Insights from ongoing studies (links provide further details):
<b>Surveillance &amp; Epidemiology</b>	<ul style="list-style-type: none"> <li>• Early serological data from over 12,000 healthcare workers as part of Public Health England’s SIREN (SARS-CoV-2 Immunity &amp; Reinfection Evaluation) study, indicate that <a href="#">antibodies from past COVID-19 provide some immunity lasting for at least 5 months</a> though a small number of people with antibodies may still carry and transmit the virus.</li> <li>• A retrospective cohort study leveraging data from 114 hospital trusts in England has found that <a href="#">increased bed occupancy is associated with mortality in Intensive Care Unit (ICU) patients</a>. This work highlights the need to keep ICU capacity low by controlling COVID-19 incidence and prevalence.</li> <li>• Using hospital electronic health records in UK, Norway &amp; China, researchers found <a href="#">the National Early Warning Score system, currently used by most UK NHS Trusts to predict severe COVID-19 outcomes in patients, is moderately successful in the short term (3 days), but only poor-to-moderate success in the medium-term (14 days)</a>.</li> <li>• The latest analyses from the community survey, REal-time Assessment of Community Transmission study-1 (REACT-1), indicate an <a href="#">increased prevalence of SARS-CoV-2 associated with large household size, living in a deprived neighbourhood, and Black and Asian ethnicity</a>. Of note, keyworkers (including health and social care workers) were more likely to test positive for SARS-CoV-2 compared to other workers.</li> </ul>
<b>Immunity &amp; Vaccines</b>	<ul style="list-style-type: none"> <li>• Analyses using an <a href="#">openly accessible tool integrating SARS-CoV-2 genome sequence, immunological, and mutation data</a> indicate that transmembrane proteins (Orf3a) may be a suitable target for diagnostic tests. This online tool is likely to prove useful in monitoring for mutation as vaccination programmes introduce new selection pressures on the virus.</li> <li>• Initial data showing a <a href="#">small drop in effectiveness of the Pfizer BioNTech vaccine (BNT162b2) against mutated forms of SARS-CoV-2 is limited to only 23 participants</a> (and <a href="#">similar small samples from Moderna</a>) - reinforcing the need for large scale, systematic data collection as vaccine programmes are rolled-out across the UK and globally.</li> </ul>
<b>Longitudinal health &amp; wellbeing</b>	<ul style="list-style-type: none"> <li>• Analysis of primary care data on the OpenSAFELY platform, suggests <a href="#">adaptive changes in primary care during the UK’s first COVID-19 wave</a>: low priority tests decreased, such as those for cholesterol, whilst some high priority tests remained stable. Ongoing monitoring of these data is essential to support continuity of high priority services.</li> <li>• Using population-based datasets spanning England, researchers have shown a <a href="#">sustained reduction in the number of people referred, diagnosed, and treated for colorectal cancer with 3,500 fewer patients beginning treatment between April and October 2020</a>. These findings add to the body of evidence on the detrimental effect of the pandemic on cancer detection and management and the growing backlog of undetected cancers within the NHS.</li> <li>• An observational retrospective matched cohort study involving using care records of nearly 50,000 patients in England has revealed that <a href="#">individuals discharged from hospital following acute COVID-19 face increased rates of mortality, readmission and multi-organ dysfunction</a>.</li> </ul>
<b>Transmission &amp; Environment</b>	<ul style="list-style-type: none"> <li>• New estimates from a random sample including 185,342 UK households show that <a href="#">the new UK SARS-CoV-2 variant leads to higher infection rates but is not adapted to any particular age group, and likely less transmissible than previous estimates</a>.</li> </ul>
<b>Clinical Trials</b>	<ul style="list-style-type: none"> <li>• In response to the current urgency to alleviate slow recruitment to the PRINCIPLE trial (Platform Randomised trial of INterventions against COVID-19 In older peoPLE), <a href="#">a method using near-real time NHS Test and Trace data to recruit participants complemented by unique access to the Summary Care Records (SCRs) to ensure safe prescribing</a>, has been developed and implemented as part of this key national therapeutic trial. The approach, which increased recruitment, should be considered by others where slow recruitment is a barrier to trial delivery.</li> </ul>

# Data & Connectivity National Core Study: COVID-19 dataset availability – 26 January 2021

ONS/NHS Digital joint health data asset launching imminently, significant work to understand and establish C-19 vaccine data flows.

**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**.

First industry dataset accessible via Innovation Gateway ([GSK Salford Lung Study](#) dataset). New [Trusted Research Environment collection](#) added to Gateway collections.

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	Data flow established – agreeing process for providing access	Consulting on joint controllership		Expected this week	
COG-UK viral genome		Finalising governance	Awaiting IG sign off	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 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 (Pillar 4)		N/A		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
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 – internal use only. HES critical care available via DARS	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 <b>New!</b>	Internal use only			
Pillar 3 Testing data (iELISA)	Available in CIS data <b>New!</b>				Data to be validated
Other Pillar 4 Testing data	VIVALDI, REACT II				
ZOE Symptom Study App Data				UK wide (unlinked) Wales (linked)	



# Status of COVID-19 projects using the data – 26 January 2021

11 additions to active research taking place takes total over 250. Largest number of new projects enabled by the Office for National Statistics. Rapid Call projects starting to complete data access requests via Gateway.

# 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	17 (-5)	31 (-1)	36 (+6)	90 (-)	6 (-)	176 (-)
- a/w researcher	10 (-)	19(-5)	Not available	56 (-)	6 (-)	N/A
- a/w data custodian	3 (-5)	12(+4)	Not available	34 (-)	0 (-)	N/A
Submitted for Information Governance approval	3(-6)	8 (+5)	4(+1)	1 (+1)	0 (-)	18 (+1)
Approved but not yet active	6 (-1)	1 (-3)	9 (+2)	2 (-)	1(-)	19 (-4)
<b>Active research taking place</b>	<b>11 (+5)</b>	<b>84 (+2)</b>	<b>67 (+2)</b>	<b>91 (+2)</b>	<b>1(-)</b>	<b>253 (+11)</b>
Active Number of Researchers	Not yet available	Not yet available	Not yet available	202 (+22)	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](#): 3,743 participants (+16% in last 2 weeks, with continuing data flow of Pillar 2 COVID +ve test results to support recruitment)
- [RECOVERY](#): 32,212 participants across 177 active sites (+18% in last 2 weeks)
- [CO-CIN \(ISARIC 4C\)](#)
  - 161,586 Tier 0 (case report)
  - 657 Tier 1 (single sample)
  - 1,672 Tier 2 (serial sampling)
- [GENOMICC](#): 9,246 participants across 211 ICUs with a total of 5091 intensive care beds
- [COVID-19 ZOE symptom study](#): 4,562,628
- [COG-UK](#): 214,039 viral genomes sequenced

## Data & Connectivity National Core Study projects now underway

Six studies now undergoing Data Access Requests (DARs) process via [Innovation Gateway](#). Examples include:

- [Economic scarring from the COVID-19 induced crisis: monitoring inequality in economic and education outcomes](#): led by Professor Stephen Machin, LSE
- [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 Julia Hippisley-Cox, University of Oxford
- [How can NCS healthcare data be connected with wastewater surveillance of COVID-19 in a privacy-preserving fashion to inform epidemiological models and democratise data access?](#) Led by Professor Till Hoffman, Imperial College London

## 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.