COVID-19 Health Data Research

8 June 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

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

1,242 COVID-19 pre-print publications, and **158** papers published

771 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:

As we continue to move out of a national lockdown, we must sustain momentum:

- We must not lose sight of the need to treat COVID-19. Further research into this area should be prioritised so that we are better prepared in the future given we do not know how long immunity from vaccines will last.
- With more and more of the population being vaccinated and receiving both doses, there are a number of urgent research questions including
    - how long immunity lasts and if certain populations are at risk of contracting COVID-19 following vaccination
    - the effectiveness of the vaccine on the different and emerging variants – particularly as variants such as the delta variant become more prominent
    - whether it has the same level of effectiveness in different populations e.g., the virus affects different ethnicities in different ways – will the vaccine be the same?
- As people will be encouraged to return to the workplace, further research is urgently needed to understand any trends in different workplace settings and to inform interventions.

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 4 weeks

Health data research outputs on COVID-19 continues to grow, now reaching 1,242 (+20) non-peer-reviewed pre-prints & 158 (+11) published papers.

Topic	Insights from ongoing studies (links provide further details):
Surveillance & Epidemiology	• A large-scale phenome wide association study using electronic health record and genomic data from >450,000 participants of the Million Veterans Programme identified existing conditions (e.g. inflammatory diseases) that share genetic variants with severe COVID-19 – providing new potential pathways to target in the treatment of COVID-19. • COVID Symptom Study data from >250,000 children aged 5-17 confirms that COVID-19 in school-aged children is usually short with mild symptoms – however, some children experience prolonged illness requiring appropriate support. Likewise, data from 4,678 children as part of the VirusWatch study suggest the prevalence of persistent COVID-19 symptoms (>4 weeks) in children is low; teens, girls, and those with long-term health conditions had a higher risk of persistent symptoms. • The latest surveillance data from Public Health England show that the delta variant is now the most dominant COVID-19 strain circulating in the UK, overtaking the alpha variant. There may be immediate implications for the NHS given early evidence of a higher risk of hospitalisation in people with the delta variant. • A time series analysis of weekly mortality data from 29 high income countries estimates that nearly 1 million excess deaths occurred in 2020. In most countries, estimated excess deaths exceeded reported COVID-19 deaths and were much higher in men.
Immunity & Vaccines	• A community-based study using data from >1 million vaccinated UK adults as part of the COVID Symptom Study found older individuals with frailty and those living in deprived areas are at a higher risk of contracting COVID-19 following vaccination – though illness appears less severe. • Analyses of nation-wide linked census and health record data in England have revealed that 1st dose vaccination rates in people >70 years are lower among all minority ethnic groups compared to the White British population, and among individuals who identified as Muslim and Buddhist. This study speaks to the urgent need for informed public health policy and community engagement to promote vaccine uptake in these groups.
Longitudinal health & wellbeing	• Combined mortality and mental health care data from England, Scotland, and Wales showed reductions in new mental health inpatient admission and face-to-face community carefollowing the first national lockdown. The long-term effects of these changes, particularly the shift to virtual services, should be closely monitored. • Analyses of longitudinal data from 708 children as part of a birth cohort study in the UK (Avon Longitudinal Study of Parents and Children) suggest that the pandemic may be associated with increased emotional and behavioural difficulties beyond the age of 2. Continued research and support for carers and children is critical. • Primary care records from 58 million people show the use of diagnostic codes for Long Covid is low and varies between practises, regions, and GP software. Full and consistent use of Long Covid coding is essential to research and planning of services. • An online survey of >4,000 individuals with psoriasis across 86 countries indicates that worsening psoriasis during the pandemic is associated with poor mental health. Concerns about immunosuppressant-related risks of COVID-19 led a subset of individuals to stop immunosuppression – underscoring the importance of access to integrated care and support. • An entire population-based study of Scotland and Wales using person-level datasets from Public Health Scotland and the Secure Anonymised Information Linkage Databank (SAIL) revealed the most substantial reductions in chronic obstructive pulmonary disease (COPD) exacerbations ever across the 2 nations during UK lockdown and no corresponding increases in COPD deaths. Understanding the underlying reasons for this reduction will be advantageous to future care.
Transmission & Environment	• Linking COVID-19 outbreak data with information on workplace settings in England revealed that manufacturers and packers of food had the highest outbreak rates and this was consistent across regions. These findings are important in informing interventions and monitoring programmes.
Clinical Trials	• On-going progress

Data & Connectivity National Core Study: COVID-19 dataset availability – 8 June

Viral genome data now available in Wales. Primary care data now available from ONS as part of the linked data asset (Census-Mortality-HES-GPES)

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	Scottish Vaccination Data	COVID Vaccination Dataset	Data access agreed. Data to be transferred to TRE shortly from Vaccine Management System
COG-UK viral genome	Awaiting data flow from Public Health England	Awaiting data flow from Public Health England (to follow flow into ONS)	COG-UK data available in TRE Subset also linked to CO-CIN data	COG-UK data available in TRE	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	Census-Mortality-HES-GPES linked data asset now available (ONS/NHSD)	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-GPES 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	Preparing data sharing agreement for ICNARC	HES Critical Care (ICNARC available in June)	SICSAG (updated weekly)	ICNARC COVID weekly, ICNARC quarterly all admissions and critical care routine data (CCDS) monthly	ICNARC 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				

HDRUK Health Data Research UK

Office for National Statistics

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**, along with our **latest monthly sprint report**

Data and Connectivity National Core Study webpages and dashboard now live

Status of COVID-19 projects using the data – 8 June

# 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	8 (+3)	24 (-16)	23 (-9)	48 (+2)	5 (-)	108 (-20)
Submitted for Information Governance approval	0 (-2)	12 (+2)	4 (-4)	2 (+1)	0 (-)	18 (-3)
Approved but not yet active	6 (+2)	1 (-3)	21 (+9)	2 (+1)	2 (-1)	32 (+8)
Active research projects taking place	**26 (-2)**	**110 (+10)**	**78 (-)**	**128 (-2)**	**2 (+1)**	**344 (+7)**
Completed projects	3	0	3	25	0	31
Active Number of Researchers	367 (+15)	48 (+3)	154 (-85)	233 (-22)	10 (+8)	812 (-81)
Average time from application to active research	115 days	*Not available*	10 days	7 days	47 days	
Number of disclosure requests in last 4 weeks*	*Not available yet*	156	53	110	38	
Number of items in disclosure requests*	*Not available yet*	263	459	639	2843	

* These are requests by researchers to export safe and anonymous data from the Trusted Research Environments. These requests include items such as graphs or data tables

Participation in key UK wide studies:

- PRINCIPLE: 5,018 participants (+3% in last 2 weeks, with continuing data flow of Pillar 2 COVID +ve test results to support recruitment)
- RECOVERY: 39,967 participants across 181 active sites (+0.5% in last 2 weeks)
- CO-CIN (ISARIC 4C)
  - 209,693 Tier 0 (case report) (+1.2% in last 2 weeks)
  - 2334 Tier 1/2
- GENOMICC: 13,011 participants (+3.3% in last 2 weeks) across 212 ICUs with a total of 5091 intensive care beds
- COVID-19 ZOE symptom study: 4,661,644
- COG-UK: 511,351 viral genomes sequenced (+8% in last 2 weeks)

HDRUK
Health Data Research UK

Office for National Statistics

Data Use 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.