COVID-19 Health Data Research

10 November 2020 - 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
Clara Fennessy, Health Data Research UK
David Seymour, UK Health Data Research Alliance
Garry Coleman, NHS Digital
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)

Research topics with new insights generated in last 2 weeks

Health data research on COVID-19 continues to grow, now reaching 914 (+9) non peer-reviewed pre-prints & 67 (-) published papers.

Topic	Insights from ongoing studies (links provide further details):
Surveillance & Epidemiology	• Data from the ONS Coronavirus Infection Survey show that risk factors for testing positive varied substantially between the initial and second waves, and a substantial proportion of infections were in asymptomatic individuals, indicating the importance of continued monitoring for SARS-CoV-2 in the community. • COVID Symptom Study app data on symptoms and PCR test results were used to identify 35-80% of regions appearing in the Government's hotspot list regions before they were subject to local government lockdowns. Earlier identification of developing hotspots could save lives. • Daily human mobility data from Facebook for Good, available in near real-time, revealed that communities under the first national lockdown did not reorganise in the same way as they did under locally-targeted interventions.
Immunity	• Prevalence of antibody positivity data in England from REACT2 shows declining prevalence of antibody positivity to SARS-CoV-2 - at the start of the second wave in England, only 4.4% of adults had detectable antibodies. Antibody positivity was greater in those who reported a positive PCR and lower in older people and those with asymptomatic infection. These data suggest the possibility of decreasing population immunity and increasing risk of reinfection as detectable antibodies decline in the population.
Longitudinal health & wellbeing	• Using a large mental healthcare database linked to death registrations (SLaM), it was found that missing data is a barrier to determining whether excess deaths during the pandemic not due to COVID-19 were as a result of increased suicide rates during lockdown. • A USA study set out to determine whether adverse mental health symptoms are either/both consequences and risk factors for COVID-19. Preliminary findings form analysis of anonymised data from 68m electronic health records on the TriNetX Analytics Network, were that survivors of COVID-19 appear to be at increased risk of psychiatric sequelae, and a psychiatric diagnosis might be an independent risk factor for COVID-19.
Transmission & Environment	• Using a method developed as part of COG-UK, linked cases of infection were identified using data collected from two wards at Cambridge University Hospitals, on Covid-19 and non-Covid-19 wards. This method is suitable for the rapid analysis of data from clinical or other potential outbreak settings.
Clinical Trials	• After safety concerns around the REGN-COV2 antibody cocktail, the RECOVERY trial's Data Monitoring Committee reviewed safety/efficacy data and concluded that there was no cogent reason to modify the protocol or intake to the study, and that recruitment of eligible patients to all arms of the RECOVERY trial should continue. Also, to test the hypothesis that aspirin may reduce the risk of blood clots in patients with COVID-19, aspirin has been added as a possible treatment for COVID-19 in the RECOVERY trial. • Building on studies to date on hydroxychloroquine, a further question remains: does routine ongoing use of hydroxychloroquine in people without coronavirus infection protect against new infections or severe COVID-19 outcomes? A study using national primary care data and linked death registrations in the OpenSAFELY platform found no evidence of a difference in COVID-19 mortality among people who received hydroxychloroquine for treatment of rheumatological disease before the COVID-19 outbreak in England. Therefore, completion of randomised trials investigating pre-exposure prophylactic use of hydroxychloroquine for prevention of severe outcomes from COVID-19 are warranted.

Data & Connectivity National Core Study: COVID-19 dataset availability – 10 November 2020

There has been a significant improvement in dataset discoverability via the Health Data Research Innovation Gateway over past 2 weeks including Pillar 1 & 2 testing data with technical metadata

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)
Pillar 1 COVID-19 Testing Data	To be linked				
Pillar 2 Testing data (UK Gov)	To be linked				Missing results prior to 26 Apr – Data quality issue
Primary Care	Onboarding underway	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	100% coverage	100% coverage	
Personal Demographic Service	100% coverage	100% coverage	100% coverage	100% coverage	
Death registry	Provisional Monthly Extract & Linked Census and death occurrence	100% coverage	100% coverage	100% coverage	
C-19 Infection Survey					
COVID-19 Clinical Information Network (CO-CIN)		Data for linkage in Scottish Data Safe Haven agreed	Limited metadata		
Census 2011	Household structure				
Covid Opinions Survey					
Business Impact of Covid Survey	c. 5,000 businesses				
Labour Force Survey	40,000 households, 100,000 individuals				
ZOE Symptom Study App Data				UK wide (unlinked) Wales (linked)	

NB: Physical access only for research - not currently possible due to COVID-19 infection control measures

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**

Data & Connectivity Sprint 2 linkage priorities include:

- Viral genome data from COVID-19 Genomics UK (COG-UK) consortium
- Serology (multipe sources)
- Intensive Care Data
- Occupational data

Status of COVID-19 projects using the data – 10 November 2020

Eight additions to active research taking place takes total over 190 with most new projects enabled by NHS Digital data. 'In development' remains relatively flat (158), whilst some growth in bottle neck at IG approval and not yet active (increase of 10 to 25).

# 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	*to be added*	27 (-3)	41 (+7)	84 (-3)	6 (-)	158 (+1)
- a/w researcher	*for*	*20 (+2)*	*Not available*	*53 (-8)*	*6 (-)*	*N/A*
- a/w data custodian	*future*	*7 (-5)*	*Not available*	*31 (+5)*	*0 (-)*	*N/A*
Submitted for Information Governance approval	*reports*	9 (+3)	7 (+2)	0 (-)	0 (-)	16 (+5)
Approved but not yet active		3 (+1)	0 (-)	4 (+4)	2 (-)	9 (+5)
Active research taking place		**67 (+7)**	**46 (-)**	**78 (+1)**	**0 (-)**	**191 (+8)**
Active Number of Researchers	*to*	*be*	*added*	*for*	*future*	*reports*
Average time from application to active research	*to*	*be*	*added*	*for*	*future*	*reports*

Key UK wide studies
RECOVERY
PRINCIPLE
CO-CIN (ISARIC 4C)
COG-UK
CARDIOVASCULAR CONSORTIUM
PHOSP-COVID
COVID-19 symptom study
GENOMICC

Seven new active research enabled by **NHS Digital** includes daily access to positive COVID-19 test results within 24 hours to increase recruitment to PRINCIPLE (Platform Randomised trial of INterventions against COVID-19 In older people). Rapid turnaround enabled by work across Data & Connectivity and engaged stakeholders including ICO, DHSC and HDR UK Public and Patient Panel

SAIL continues to support access to ZOE Symptom Study data.
Recent data provision to:

- NICE
- NHS Tower Hamlets CCG
- Bradford Teaching Hospital Foundation Trust
- Liverpool University
- Joint Biosecurity Centre.

Scotland National Data Safe Haven hosts UK-wide ISARIC dataset now with approved linkages for England and Scotland. The Tier Zero (Case report forms) represent a cohort of approaching 100,000 hospitalised patients (95,794 patients). The source data for the COVID-19 Clinical Information Network (CO-CIN)) is available to other researchers.

ONS has made available via its **Secure Research Service** a Bespoke linked dataset that contains death registrations linked to the 2011 Census. The dataset that runs from 2011 to most recent available includes approaching 100 variables.

Further information about **Data & Connectivity** can be found **here**. Including end of **Sprint 1 report**.

Health data research community highlights this period*

24 COVID-19 taskforce calls with **167** clinical and health data research leaders engaged

1460 academic, industry and NHS participants in COVID-19 Slack channel with 10 sub-channels

914 COVID-19 pre-print publications

111 health data research questions identified – 42 prioritised

Patient and Public Voice Feedback

Vital to ensure research uses data across all four nations of the UK to inform more appropriate and representative research with priority areas of focus around domiciliary (at home) care, mental health and psychological conditions, and "Long Covid".

Click here for a link to the full prioritised list of questions, status, and prioritisation process

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