

HDRUK
Health Data Research UK



HSC Public Health Agency
Research and Development



COVID-19 Health Data Research

24 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
Sophie Morris, 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)



Health data research community highlights this period*



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



1044 COVID-19 pre-print publications



1474 academic, industry and NHS participants in COVID-19 Slack channel with 10 sub-channels



111 health data research questions identified – 42 prioritised



This Week's Patient and Public Voice Feedback:

"It is vital to understand transmission among children in school to ensure they can stay in school for longer. Research about COVID-19 in social care settings must also not be forgotten"

[Click here to read more feedback](#)



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 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 1044 (+130) non peer-reviewed pre-prints & 68 (+1) published papers.

| Topic | Insights from ongoing studies (links provide further details): |
|--|---|
| Surveillance & Epidemiology | <ul style="list-style-type: none"> A study shows that the detection of certain protein levels can accurately predict COVID-19 outcome weeks later, and distinguish between those who will be asymptomatic vs experiencing severe symptoms. Using data from the COVID Symptom Study app, disease incidence, prevalence and effective reproduction number were estimated and geographically granular estimates used to highlight regions with rapidly increasing case numbers, or hotspots. This demonstrates that self-reported data from mobile applications can provide an agile resource to inform policymakers during a fast-moving pandemic. |
| Immunity | <ul style="list-style-type: none"> A study using bespoke online surveys found that risk-mitigating behaviours in people with immune-mediated inflammatory diseases receiving targeted systemic therapies, were responsible, in part, for fewer adverse COVID-19 outcomes compared to patients receiving no systemic treatments. Analysis of data from SARS-CoV-2 PCR-positive results in seropositive and seronegative healthcare workers attending asymptomatic and symptomatic staff testing at Oxford University Hospitals, found that prior SARS-CoV-2 infection that generated antibody responses offered protection from reinfection for most people in the six months following infection. Further work is required to determine the long-term duration and correlates of post-infection immunity. Antibody levels to SARS-CoV-2 wane within months, and faster in younger adults and those without symptoms. Ongoing longitudinal studies are required to track the long-term duration of antibody levels and their association with immunity to SARS-CoV-2 reinfection. |
| Longitudinal health & wellbeing | <ul style="list-style-type: none"> Using routine health data from c10m patients in CPRD, it was found that interactions with primary care dropped dramatically after first national lockdown, and as of July 2020, engagement with primary care had not recovered to pre-lockdown levels (except for unstable angina and acute alcohol-related events). The largest reductions were for diabetic emergencies, depression, and self-harm, which will have substantial ramifications for these vulnerable patients. The Invasive Respiratory Infection Surveillance Initiative reveals significant reductions in invasive bacterial infections during the COVID-19 pandemic, likely owing to introduction of COVID-19 restrictions and associated public information campaigns, leading to a significant global reduction in life-threatening invasive diseases. The Clinical Record Interactive Search tool was used to examine routine health data of people receiving mental healthcare from a NHS Foundation Trust and found no evidence of changes in psychiatric prescribing, as a result of the marked increase in remote consultation since the epidemic began, particularly among younger patients. |
| Transmission & Environment | <ul style="list-style-type: none"> A prospective sero-epidemiological study of healthcare workers at a UK teaching hospital using a SARS-CoV-2 immunoassay, found that risk of infection amongst healthcare workers was found to be influenced by COVID-19 working location, role, age and ethnicity. Analysis of data from the OPENSafely cohort found that for c10m adults living with children, there is no evidence of increased risk of severe COVID-19 outcomes – as compared to adults not living with children. This does not support the hypothesis that close contact with children increases risk of SARS-CoV-2 infection. |
| Clinical Trials | <ul style="list-style-type: none"> A RECOVERY trial participant has become the first person in England to receive convalescent plasma treatment and then donate convalescent plasma. Convalescent plasma is being tested as part of the RECOVERY trial to explore its efficacy in helping people recover than the usual standard of hospital care which all patients receive. |

Data & Connectivity National Core Study: COVID-19 dataset availability – 24 November 2020

Significant data custodian engagement across the sprint 2 priority datasets and continued improvement of sprint 1 metadata, but no new additions yet available for access.

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

| 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 | GPES 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 (Pillar 4) | | | | | |
| COVID-19 Clinical Information Network (CO-CIN) | To be linked to 2011 census | Data for linkage in Scottish Data Safe Haven agreed | Limited metadata | Awaiting decision on data access | |
| 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 | | | | |
| Intensive Care data | | ICNARC – internal use only | SICSAG (updated weekly) | ICNARC – internal use only | |
| Pillar 3 Testing data (NHS labs) | | Internal use only | | | |
| Pillar 3 Testing data (Thrive) | | Internal use only | | | |
| Other Pillar 4 Testing data | VIVALDI, REACT II | | | | |
| ZOE Symptom Study App Data | | | | UK wide (unlinked) Wales (linked) | |
| COG-UK Viral Genome | | | | Finalising governance | |

Status of COVID-19 projects using the data – 24 November 2020

23 additions to active research taking place takes total over 200. Most new projects enabled by Scottish Data Safe Haven plus ONS projects included for first time. Some increase in bottle neck at IG approval and not yet active.

Further information about **Data & Connectivity** can be found [here](#). Including end of **Sprint 1 report**.

[Data Readiness: Lessons from an Emergency](#) recently published by [The Delve Initiative](#)

| # 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 | 13 | 23 (-4) | 42 (+1) | 84 (-) | 6 (-) | 168 (+10) |
| - a/w researcher | 11 | 16(-4) | Not available | 53 (-) | 6 (-) | N/A |
| - a/w data custodian | 2 | 7 (-) | Not available | 31 (-) | 0 (-) | N/A |
| Submitted for Information Governance approval | 5 | 10 (+1) | 9(+2) | 0 (-) | 0 (-) | 24 (+8) |
| Approved but not yet active | 1 | 2 (-1) | 6 (+6) | 2 (-2) | 2 (-) | 13 (+4) |
| Active research taking place | 6 | 68 (+1) | 60 (+14) | 80 (+2) | 0 (-) | 214 (+23) |
| Active Number of Researchers | Not yet available | Not yet available | Not yet available | 178 | Not yet available | |
| Average time from application to active research | Not yet available | Not yet available | Not yet available | | Not yet available | |

Participation in key UK wide studies:

- [PRINCIPLE](#): 2,068 participants, an increase of around 25% in last 2 weeks following establishment of data flow of Pillar 2 COVID positive test results to support recruitment
- [RECOVERY](#): 18,570 participants across 176 active sites
- [CO-CIN \(ISARIC 4C\)](#)
 - 105,144 Tier 0 (case report)
 - 724 Tier 1 (single sample)
 - 1,833 T2 (serial sampling)
- [GENOMICC](#): 5,989 participants across 210 ICUs
- [COVID-19 ZOE symptom study](#): 4,453,193
- [COG-UK](#): 117,375 (viral genomes sequenced)

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.