COVID-19 Health Data Research

04 August 2020 - Weekly update for SAGE & UKRI/DHSC

Authors:

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
John Aston, Home Office (SAGE sponsor)
John Deanfield, NICOR
Mark Parsons, Scotland National Safe Haven

Members of the HDR UK Public Advisory Board
Melissa Lewis-Brown, Health Data Research UK
Ming Tang, NHS England and Improvement
Nilesh Samani, British Heart Foundation
Rhoswyn Walker, Health Data Research UK
Ronan Lyons, SAIL Databank (UKRI/DHSC sponsor)
Sara Hiom, Cancer Research UK
Tom Denwood, NHS Digital
Alice Turnbull, Health Data Research UK
Ian Young, Health & Social Care Northern Ireland

COVID-19 Health Data Research recommendations – 04 August 2020

Health data research insights on COVID are continuing, with 109 research questions, 131 projects active within the national data Trusted Research Environments (TRE), a further 173 in development, and 157 pre-print publications and 33 published papers. Progress is happening across the 5 recommendations endorsed by SAGE:

#	5 Recommendations endorsed by SAGE on 11 June	Progress on SAGE actions identified on 11 June
1	All **swab & antibody testing programmes** data to be securely linked and used for research. Requires unparalleled cooperation across all four nations between NHS organisations, PHE, data custodians, academic endeavours, and technology partners, whilst building public trust.	**SAGE ACTION:** HDR UK to work with partners to plan and create a serology and testing data research asset that is linkable to other data sources. **PROGRESS:** Decision on funding proposal deferred to a UKRI/NIHR sub-panel, which are meeting w/c 3rd August 2020.
2	Further research, undertaken collaboratively with international partners where appropriate, should address why **BAME groups** have a higher rate of severe COVID-19 outcomes. This will help to target the best interventions and inform the response to future public health crises.	**PROGRESS**: Further insights being generated (see next page). UK Health Data Alliance Board agreed to improve consistency and quality of ethnicity coding and use data to help increase the representativeness of research studies.
3	Enhance data capture on patients and staff in **care homes**, in particular interconnections between settings, to enable research on health, transmission and outcomes. Clarify appropriate use of national Trusted Research Environments for consolidation of relevant care home COVID-19 data.	**PROGRESS**: NHS Digital have released a Data Provision Notice of the adult social care management system coronavirus (COVID-19) status collection. Data quality of initial submissions will be reviewed before confirming how/if the dataset will be available for research. Establishing a broader research response still requires more accessible individual-level data for care home staff, residents and for the domiciliary care (care within own home) sector.
4	Accelerate access to **restricted national datasets**, since lack of availability is holding back crucial research.	**PROGRESS**: *No change to previous report:* priority studies can request access but access via the standard NHS Digital Data Access Request Service (DARS) expected in Aug 20 (Testing data and Community Prescribing) and Sep 20 (CHESS). *Update:* Control of Patient Information (COPI) notice has been extended to March '21.
5	Commission large scale collaborative analyses of the **long-term impacts** of health and social care changes during the COVID-19 lockdown on major diseases. This will require access to linked data from a range of sources (including from COVID-19 laboratory tests, primary and secondary healthcare, death registries, disease-specific audit/registry data). In addition, linkages to cross sectoral data beyond health will be essential to understand the wider impacts of COVID-19.	**SAGE ACTION:** HDR UK to work with ONS and others to accelerate linkage of cross-sectoral datasets. **PROGRESS:** *No change to previous report:* HDR UK and ONS developing national data & connectivity programme in partnership with UK Health Data Research Alliance. *Update:* 13 cardiovascular researchers from 4 institutions actively working via BHF Data Science Centre access to linked datasets in the NHS Digital TRE. This will rise to 16 cardiovascular researchers from 6 institutions once the amended data sharing agreement has been signed (currently in progress).

Priority research questions with new insights generated this week – 04 August 2020

Health data research on COVID-19 continues to grow, now reaching 157 (non peer-reviewed) pre-prints[1] & 33 published papers

[1] The number of pre-prints has increased steadily over the last 3 reports: 127, 139 & 157 on the 7 July, 21 July & 4 August respectively. The previous figure of 161 shared in the 21 July report was an error.

HDRUK
Health Data Research UK

Priority research questions	Insights from ongoing studies (links provide further details):
1. Understanding & protecting vulnerable groups? (RQ 22, 32, 36, 62, 102)	• An observational study has found that severe COVID-19 is associated with polypharmacy (concurrent use of multiple medications) and in particular with drugs that cause sedation, respiratory depression or dyskinesia, have anticholinergic effects or affect the gastrointestinal system. • Recent research from ALSPAC shows progressively increasing anxiety and decreasing wellbeing in June vs April 2020 in younger populations in particular; other groups more likely to suffer from anxiety are females, individuals with a history of mental health disorders, or history of financial problems.
2. Impact on Non-COVID care provision (RQ29, 30, 94)	• Insights from 20 GP practices in Bristol, North Somerset & South Gloucestershire indicate that practices continue to innovate & cope well with COVID-19 service changes. Ongoing challenges include delivering complex consultations remotely, managing patient expectations as lockdown eases & strategies for restarting physical consultations when national guidance is unclear.
3. Use of existing treatments (RQ18, RQ98)	• There has been a substantial mortality improvement in people admitted to critical care with COVID-19 in England, with markedly lower mortality in people admitted in mid-April and May compared to earlier in the pandemic. Possible causes include the introduction of effective treatments as part of clinical trials and a falling critical care burden. • Preliminary work using ZOE app data suggests estrogen may be protective for COVID-19, including potentially those on hormone replacement therapy and the combined oral contraceptive pill. • If dexamethasone is rolled out in the UK it is estimated that c12,000 lives could be saved by Jan 2021 (and 650,000 globally, various assumptions notwithstanding). • Preliminary findings suggest that those who have been vaccinated (e.g. for polio, geriatric flu etc) in the last 1-5 years are associated with decreased SARS-CoV-2 infection rates, including significantly lower rates in Black people who were previously vaccinated for pneumonia. • There are significant similarities between the biological mechanisms that leads to rheumatoid arthritis (RA) and COVID-19, as well as between the pathways by which inflammation is reduced in both conditions. Treatments for RA could therefore inform treatments for COVID-19. • Models that aim to predict prognosis of COVID-19 patients were evaluated. Admission oxygen saturation of room air was found to be the strongest predictor of in-hospital deterioration; age was the strongest predictor of in-hospital mortality.
4. Understanding immunity, testing reliability & prevalence (R01, 50, 95, 102, 51, 54, 55, 104)	• Significant heterogeneity was found in the production of antibodies to SARS-CoV-2. Those with mild form of the disease produce significantly lower levels of antibodies, highlighting the need to consider future vaccination for COVID-19 patients recovered from asymptomatic or mild illness. • ZOE app data has found that loss of smell was found to show the highest symptom specificity for SARS-CoV-2 antibody response. • Samples for SARS-CoV-2 diagnostic tests can be obtained from the upper or lower respiratory tract. This study suggests that lower respiratory tract infection may be a key factor in the progression to critical illness, therefore virological samples should be obtained from this site where possible, as upper respiratory samples have a significant negative rate.

Patient and Public Voice Feedback
Important for research to explore risks, outcomes and long-term effects within smaller populations as opposed to combining them into one larger population with commonalities.

16 COVID-19 taskforce calls with **81** clinical and health data research leaders engaged

1374 academic, industry and NHS participants in COVID-19 Slack channel with 10 sub-channels

109 health data research questions identified – 42 prioritised

157 COVID-19 pre-print publications

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

COVID-19 dataset availability and status of projects using the data – 04 August 2020

Continued improvements in linked dataset availability in Wales, with no significant change across other national TREs. Increase in projects awaiting IG approval partially related to impact of holiday season. Further projects accessing CPRD linked data as alternative data asset to national TREs.

HDRUK
Health Data Research UK

1. **Linked dataset availability in Wales continue to increase, little or no change reported in other national TREs**

Core COVID-19 Datasets available for linkage	England (NHS Digital Data Processing Service)	Scotland (National Data Safe Haven)	Wales (SAIL Databank)	Northern Ireland (Honest Broker Service)
Primary Care				
Pillar 1 COVID-19 Testing Data (NHS/Public Health)				
Pillar 2 Testing data (UK Gov)	N/C Expected Aug			Missing results prior to 26 Apr
Pillar 3 & 4 Testing data	No confirmed date		Pillar 3 available	Not yet decided to take part
Community Prescribing	N/C Expected Aug			
Critical Care (CHESS, ICNARC, SICSAG)	N/C CHESS expected in DARS Sep			N/C - Under review
Personal Demographic Service				
Secondary Care				
Death registry				

2. **Increase in awaiting IG approval partially linked to holiday season.**

3. **Increase in England 'active research' is due to change in how cardiovascular research is counted.**

# of COVID-19 Projects by TRE stage (change from previous report)	England (NHS Digital Data Processing Service)	Scotland (National Data Safe Haven)	Wales (SAIL Databank)	Northern Ireland (Honest Broker Service)	Total
In development	39 (-3)	37 (-1)	89 (+5)	8 (-)	173 (+1)
- a/w researcher	24 (-)	Not available	66 (+4)	8 (-)	N/A
- a/w data custodian	15 (-3)	Not available	23 (+1)	0 (-)	N/A
Submitted for IG approval	8 (+3)	5 (+4)	0 (-)	0 (-)	13 (+7)
Approved but not yet active	2 (+1)	0 (-)	0 (-1)	0 (-)	2 (+1)
Active research taking place	**37 (+11)**	**27 (+2)**	**66 (-)**	**0 (-)**	**131 (+13)**

KEY

- ■ Data flows specified but not yet agreed
- ■ Data flows agreed but not yet available for linkage
- ■ Fully available

KEY UK WIDE PROJECTS:
- RECOVERY
- CO-CIN (ISARIC 4C)
- COG-UK
- CARDIOVASCULAR CONSORTIUM
- COVID-19 symptom study
- GENOMICC

Datasets available for COVID-19 research via national TREs for Wales Scotland and England

4. **Further progress on dataset availability in SAIL:** Pillar 3 COVID results available; daily flows for Situation Report Data, NHS 111 Call data, Welsh Ambulance Service Dataset; monthly flows for Wales Dispensing DataSet and Healthcare Workers Risk Assessment; next academic year of Education results; and agreement to move Welsh Longitudinal General Practice data from quarterly to monthly.

5. **Three new studies are now active using** CPRD linked data covering risk stratification models to inform management of severe COVID-19, excess mortality and end-of-life morbidity patterns, and indirect acute effects of the COVID-19 pandemic on physical and psychological morbidity.

6. **NHSx announced that Control of patient information (COPI) notices extended to 31 Mar 2021.**

NOTES
N/C – No change
TRE - Trusted Research Environment
IG - Information Governance
DPN – Data Provision Notice
CHESS - COVID-19 Hospitalisations in England Surveillance System
SICSAG - Scottish Intensive Care Audit Steering Group