

HDRUK
Health Data Research UK



HSC Public Health
Agency
Research and Development



Public Health
Scotland

NHS
Digital

SAIL DATABANK

UK Health Data
Research Alliance

COVID-19 Health Data Research

02 September 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 – 02 September 2020

Health data research insights on COVID are continuing, with 110 research questions, 144 (+6) projects active within the national data Trusted Research Environments (TRE), a further 180 (+7) in development, and 666 pre-print publications and 39 published papers.

Incremental progress across all recommendations. Actions are being incorporated into national core study - data and connectivity planning.



#	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: <i>No change to previous report:</i> Iteration ongoing with UKRI/NIHR panel following feedback requesting a more focused resubmission from the collaborative team.
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). Data Officers Group meeting scheduled to enact the UK Health Data Alliance Board commitment 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: Outcome of initial quality reviews on the NHS Digital adult social care management coronavirus status data collection is expected mid-September. The outcome of these reviews will determine how/if the dataset will be available for research. Exploratory analysis by University of Edinburgh of care record data, direct from software provider, has shown variable data completion and quality which limits initial utility but provides a basis to identify priority areas for improvement. Establishing a broader research response still requires more accessible individual-level data for care home staff, residents and domiciliary care (care within own home) sector.
4	Accelerate access to restricted national datasets , since lack of availability is holding back crucial research.	PROGRESS: 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). This recommendation will be integrated into the national core studies data and connectivity work from September.
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: HDR UK continues to work with national core study leads to identify and prioritise datasets and linkages to support priority research questions as part of the data and connectivity study programme, which awaits formal approval. <i>No change to previous report:</i> BHF Data Science Centre: 17 cardiovascular analysts from 6 institutions actively working via access to linked datasets in the NHS Digital TRE, including primary care GPES data.

Priority research questions with new insights generated this week – 02 September 2020

Health data research on COVID-19 continues to grow, now reaching 666 (non peer-reviewed) pre-prints & 39 published papers

Priority research questions

Insights from ongoing studies (links provide further details):

1. Understanding & protecting vulnerable groups? (RQ 22, 32, 36, 62, 102)

- In the UK, important risk-factors such as social deprivation or ethnicity vary over small distances, therefore risk needs to be modelled at fine spatial resolution to avoid aggregation bias. The use of spatial modelling was demonstrated for near-real-time prediction of COVID-19 prevalence at small-area resolution to inform government policy. It was found that [existing geospatial statistical methods - originally developed for global health applications are well-suited to this task and can be used in an anonymised databank environment, thus preserving the privacy of the individuals who contribute their data.](#)
- An international collaboration with the US explored the effects of coronavirus victimization distress and racial bias on mental health among BAME communities in the US. Results suggest [the COVID-19 pandemic has created new pathways to mental health disparities among young adults from BAME communities](#), by reversing formerly protective factors such as employment, and by exacerbating structural and societal inequities linked to race. Findings highlight the necessity of creating mental health services tailored to the specific needs of racial/ethnic minorities during the current and future health crises.
- An [analysis of the number of hospital discharges to all Welsh adult care homes](#) and the subsequent outbreaks of COVID-19 showed that large homes were at considerably greater risk of outbreaks throughout the epidemic, and after adjusting for care home size, a discharge from hospital was not associated with a significant increase in risk.

2. Impact of lockdown release (RQ22)

- A study shows that [reopening secondary schools is likely to have a greater impact on community transmission than reopening primary schools in England](#). Keeping transmission within schools limited is essential for reducing the risk of large outbreaks amongst school-aged children and their household members.
- A study characterised the clinical features of children and young people admitted to hospital with laboratory confirmed COVID-19. It confirmed that [children and young people have less severe acute COVID-19 than adults](#).

3. Understanding immunity, testing reliability & prevalence (R01, 50, 95, 102, 51, 54, 55, 104)

- An interactive tool was developed to assess the impact of different combinations of test sensitivity (e.g. false positives), specificity and infection prevalence. It found that, [to avoid multiple unnecessary restrictions on whole populations, and in particular individuals, from widespread population testing for SARS-CoV-2, the crucial roles of extremely high test specificity and of confirmatory testing must be fully appreciated and incorporated into policy decisions.](#)

4. Understanding contributing factors/increased risk (RQ102, 72)

- Poor vitamin D status has previously been associated with increased susceptibility to viral and bacterial respiratory infections. Therefore a study looked at the association between vitamin D and COVID-19 risk and outcomes, to explore potential causal effects. They found [suggestive evidence of association between vitamin D and the risk or severity of COVID-19](#), although further studies are required.

5. Changes in healthcare practice (RQ77, 62)

- Existing evidence suggests success rates for CPR in COVID-19 patients is low and the risk to healthcare professionals from this aerosol-generating procedure complicates the benefit/harm balance of CPR. [The pandemic seems to have changed practice, in that a higher proportion of patients have had "do not attempt CPR" status documented](#). Understanding patient preferences around life-sustaining treatment vs comfort care is part of holistic practice and supports shared decision making. It is unclear whether these attitudinal changes will be sustained after COVID-19 admissions decrease.

Patient and Public Voice Feedback



Priority focus should be ensuring safe and secure access to key datasets, particularly those related to swab and antibody testing.

18 COVID-19



taskforce calls with 87 clinical and health data research leaders engaged

1398 academic,



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

110

health data research questions identified – 42 prioritised



666

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 – 02 September 2020

Limited improvements to dataset availability but increase in both active research projects and ‘in development’ projects. New active projects include COVID-19 research related susceptibility and resilience, antibody study, and cardiovascular.

1. Community Prescribing Data now available in England (NHS Digital)

2. Introduction of Pillar 2 Testing data for England not yet complete within NHS Digital TRE.

3. Six further research projects now active across the national TREs.

4. Seven more ‘in development’ projects compared to net increase of one over previous 4 weeks.

5. NHS Digital new active research includes data release to: National Institute for Health Research (NIHR) Bioresource Consented research cohorts to identify susceptibility & resilience factors in cohorts (HES & Mortality on IBD patients) plus IPSOS MORI/Imperial REACT II Antibody Study

6. Latest approved projects in Scotland National Data Safe Haven include: Cardiovascular Consortium led by British Heart Foundation Data Science Centre, COVID-19 Infection and Subsequent Thromboembolic Events research led by University of Glasgow and University of Stirling’s Ambulance call-outs for psychiatric emergencies during the COVID-19 pandemic. Details of all approved projects available [here](#).

7. Highly requested COVID-19 Symptom Study (ZOE) App available from SAIL Databank is feature of recent pre-print [Real-time spatial health surveillance: mapping the UK COVID-19 epidemic](#).

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)	Delayed: Was expected Aug			Missing results prior to 26 Apr
Pillar 3 & 4 Testing data	No confirmed date		Pillar 3 available	Pillar 3 expected Sep
Community Prescribing	Now available			
Critical Care (CHESS, ICNARC, SICSAG)	N/C CHESS expected in DARS Sep			N/C - Under review
Personal Demographic Service				
Secondary Care	SUS only - not HES			
Death registry				

# 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	42 (+5)	36 (+1)	96 (+1)	6 (-)	180 (+7)
- a/w researcher	31 (-)	Not available	73 (+1)	6 (-)	N/A
- a/w data custodian	11 (+5)	Not available	23 (-)	0 (-)	N/A
Submitted for IG approval	6 (+1)	5 (-1)	0 (-)	0 (-)	11 (-)
Approved but not yet active	2 (-)	0 (-)	0 (-)	2 (-)	4 (-)
Active research taking place	41 (+3)	37 (+3)	66 (-)	0 (-)	144 (+6)



Health Data Research UK

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](#)

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
 HES – Hospital Episode Statistics
 SUS – Secondary Uses Service