

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

21 July 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 – 21 July 2020

Health data research insights on COVID are continuing, with 109 research questions, 118 projects active within the national data Trusted Research Environments (TRE), a further 172 in development, and 161 pre-print publications and 30 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: Funding proposal submitted to UKRI/NIHR, outcome expected w/c 20 Jul.
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). HDR UK Cambridge members leading the 50,000 participant Bangladesh Longitudinal Investigation of Emerging Vascular Events (BELIEVE) cohort study, are collecting data on COVID-19 symptoms, which will be linked with baseline data to understand the risk factors for COVID-19 infection, severe illness and death in South Asian populations.
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: Provider-level data is being consolidated on DHSC Exchange to support immediate policy response. Engagement with software providers suggests more detailed data may be accessible from a subset of care homes. To support a broader collaborative research response gaps need to be addressed in individual-level data for staff, residents and for the domiciliary care sector (care within own home).
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) not expected until Aug 20 (Testing data and Community Prescribing), Sep 20 (CHESS). Control of Patient Information (COPI) notice to expected be extended for 6 months.
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 and ONS developing national data & connectivity programme in partnership with UK Health Data Research Alliance partners. Cardiovascular researchers now actively working via BHF Data Science Centre access to linked datasets in the NHS Digital TRE

Priority research questions with new insights generated this week – 21 July 2020



Health data research on COVID-19 continues to grow, now reaching 161 (non peer-reviewed) pre-prints & 30 published papers

Priority research questions	Insights from ongoing studies (links provide further details):
1. How do we best understand and protect vulnerable groups? (RQ 22, 32, 36, 62, 102)	<p>BAME groups:</p> <ul style="list-style-type: none"> Excess mortality in past and present users of a large and diverse mental health service in London (SLaM), found higher rates of death in BAME communities, & in particular Black African/Caribbean patients. Black and Mixed ethnicity are independently associated with greater admission risk with COVID-19 and may be risk factors for development of severe disease. Comorbidities and socioeconomic factors only partly account for this. <p>Care homes:</p> <ul style="list-style-type: none"> Analysis in NHS Lothian shows strong association between care home size and likelihood of outbreak. with increased excess deaths (COVID and non) for those providers experiencing an outbreak. Analysis in 179 long term care facilities shows many residents report COVID symptoms without laboratory confirmed infection. Lower staffing ratios & higher occupancy rates are associated with increase infection risk in residents & staff.
2. Impact on Non-COVID care provision (RQ29, 30, 94)	<ul style="list-style-type: none"> Presentation, diagnosis and treatment of cardiovascular disease started to decline 1-2 weeks before lockdown and fell by 31-88% after lockdown with the greatest reductions in cardiovascular surgical procedures. During lockdown, urgent two-week wait GP referrals in England for suspected cancer dropped by up to 84 per cent – raising fears that undiagnosed cancers could be progressing from early-stage tumours to advanced, incurable disease. Furthermore, substantial increases in the number of avoidable cancer deaths in England are to be expected as a result of diagnostic delays due to the COVID-19 pandemic in the UK.
3. Use of existing treatments (RQ18, RQ98)	<ul style="list-style-type: none"> The RECOVERY Trial has published preliminary evidence demonstrating that in patients hospitalized with COVID-19, hydroxychloroquine was associated with an increased length of hospital stay and increased risk of progressing to invasive mechanical ventilation or death and not associated with reductions in 28 day mortality.
4. Understanding immunity, testing reliability & prevalence (R01, 50, 95, 102, 51, 54, 55, 104)	<ul style="list-style-type: none"> An international study shows that despite considerable under-reporting of COVID-19 in many locations, consistent with emerging serological data, the proportion of each country's population infected is generally low. RT-PCR misses detection of people with SARS-CoV-2 infection: sampling soon after symptoms occur minimises false negative diagnoses; beyond ten days post-symptom onset, lower RT or faecal testing may be preferred sampling sites. Significant decreases in incidence of COVID-19 and R were observed on the Isle of Wight immediately after the launch of the Test, Track and Isolate pilot programme. This sub-epidemic had the 3rd highest R number before this intervention and 10th lowest after.
5. Symptoms & susceptibility (RQ35, RQ68)	<ul style="list-style-type: none"> ZOE app data has found that skin rash is an early symptom of COVID-19 and in some cases the only clinical sign. Using health data and satellite-derived data from USA, England and Italy, an observational study found that ambient UVA exposure is associated with lower COVID-19 specific mortality. And that this effect is independent of vitamin D. An observational study has suggested that COVID-19 could be seasonal, declining with rising temperatures and humidity. Low relative humidity of indoor air might significantly contribute to severity of the disease.

Public Advisory Board Feedback

Need to explore **why** there are additional risks or adverse effects in vulnerable populations to help us better protect and ensure policies are put in place to support.



15 COVID-19 taskforce calls with **79** clinical and health data research leaders engaged



1366 academic, industry and NHS participants in COVID-19 Slack channel with 10 sub-channels



109 health data research questions identified – 41 prioritised



161 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 – 21 July 2020

Positive progress on a number of elements: availability of testing data; large growth (22) of in development projects; Cardiovascular consortium now active in NHS Digital TRE; cross-sectoral linkage in Wales; NI data access; and extension of COPI notices expected.

1. Pillar 1 & 2 data now available across all four nations but not yet available to request via DARS (NHS Digital)

2. Some progress on Pillar 3 data but mixed picture across UK.

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)	Expected Aug			Missing results prior to 26 Apr
Pillar 3 & 4 Testing data	Expected Aug			Not yet decided to take part
Community Prescribing	Delayed – Now Expected Aug			
Critical Care (CHESS, ICNARC, SICSAG)	CHESS expected in DARS Sep			N/C - Under review
Personal Demographic Service				
Secondary Care				
Death registry				

3. Strong growth of 'In development' and where breakdown is available it highlights that many potential projects have actions with researchers to progress.

# 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 (+7)	38 (+6)	84 (+7)	8 (+2)	172 (+22)
- a/w researcher	24	Not available	62	8	N/A
- a/w data custodian	18	Not available	22	0	N/A
Submitted for IG approval	5 (+1)	1 (-)	0 (-)	0 (-)	6 (+1)
Approved but not yet active	1 (-2)	0 (-)	0 (-1)	0 (-)	1 (-2)
Active research taking place	26 (+3)	26 (+1)	66 (+2)	0 (-)	118 (+6)

4. Cardiovascular Consortium now active in NHS Digital TRE, the first project to be fully operational in the environment. Congratulations and thanks to all involved across the consortium and NHS Digital.

5. New cross-sectoral datasets now available in SAIL: Healthcare Worker Risk Assessment - anonymised list of all NHS staff in Wales; Substance Misuse Dataset now updated and accessible; School Workforce Annual Census list of all teachers in Wales accessible.

6. NI planning to re-open the Honest Broker Service physical Safe Haven early Aug with restricted availability to facilitate social distancing. Pilot progressing using the the SeRP platform in partnership with HDR-UK Wales/NI. Link to [Business Intelligence dashboard](#).

7. Current Control of patient information (COPI) notices expected to be extended.

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