

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

23 June 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
Ian Young, Health & Social Care Northern Ireland



COVID-19 Health Data Research recommendations – 23 June 2020

Health data research insights on COVID are continuing at a pace, with 106 research questions, 100 projects active within the national data Trusted Research Environments (TRE), a further 150 in development, and 115 pre-print publications. Progress is happening across the 5 recommendations endorsed by SAGE

#	5 Recommendations endorsed by SAGE on 11 June	SAGE Action and Progress Update
1	All swab & antibody testing programmes data to be securely linked and used for research. Including directly supporting collaborative programmes that build on 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 define a plan for the creation of serology and testing data research asset that is linkable to other data sources, and to report back to SAGE in due course PROGRESS: Collaborative proposal in preparation for funding submission to UKRI/NIHR by 30th June.
2	Ensure that further research, undertaken collaboratively with international partners where appropriate, addresses why BAME groups appear to have a higher rate of severe COVID-19 outcomes. Including understanding whether BAME groups are more likely to contract COVID-19 and/or have an increased risk of severe outcomes once infected. This will help to target the best interventions and inform the response to future public health crisis.	PROGRESS: Further insights being generated on likelihood of contracting and on outcomes once infected by ethnicity (See slide 3)
3	Enhance data capture on patients and staff in care homes, in particular interconnections between settings, to enable in-depth research on health, transmission and outcomes. Provide clarity on appropriate use of national Trusted Research Environments for consolidation of relevant care home COVID-19 data.	PROGRESS: Ongoing work in Scotland & Wales to validate algorithms to identify care home residents plus potential to link to routine medical records for detailed characterisation of care home populations. Research community support for DHSC care home working group to address data linkage challenges and issues with attribution to individual care homes.
4	Accelerate access to currently restricted national datasets, including CHES (COVID-19 Hospitalisation in England Surveillance System) – this important data will not be fully available to researchers via the NHS Digital Data Access Request Service until the 22nd June, holding back research.	PROGRESS: CHES available as standalone dataset via Public Health England. Available for request as linked dataset from NHS Digital who are working through data quality and coverage concerns.
5	Commission large scale analyses of the long-term impacts of health and social care changes during the COVID-19 lockdown on major diseases, involving researchers, frontline clinical teams and disease registry experts. 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 on all vulnerable populations.	SAGE ACTION: HDR UK to work with ONS and other partners to accelerate linkage of cross-sectoral datasets, and to report back to SAGE on progress in due course PROGRESS: HDR UK, ONS, JBC and PHE identifying opportunities to support data and analysis for JBC

Priority research questions with new insights generated this week – 23 June 2020

Health data research on COVID-19 continues to grow, now reaching 115 pre-print publications



Priority research questions Insights from ongoing studies (links provide further details):

<p>1. Understanding immunity & testing reliability (R01, 50, 95, 102, 51, 54, 55, 104)</p>	<ul style="list-style-type: none"> SARS-CoV-2 antibody responses become detectable after the first week of illness. Dual (nucleic acid & antibody) point of care SARS-CoV-2 testing can significantly improve diagnostic sensitivity, whilst maintaining high specificity.
<p>2. Why do BAME groups have an increased risk of severe COVID-19 outcomes (RQ34, 68)?</p>	<ul style="list-style-type: none"> The ISARIC CCP-UK study has shown that ethnic minorities with COVID-19 were more likely to be admitted to critical care, despite similar disease severity on admission, similar duration of symptoms, and being younger with fewer comorbidities. South Asians are at greater risk of dying, due at least in part to a higher prevalence of pre-existing diabetes. Studies using linked UK Biobank data have demonstrated that being overweight is more strongly linked to COVID-19-related deaths in younger people and non-white ethnicities and that multimorbidity, especially cardiometabolic multimorbidity, and polypharmacy are associated with a higher risk of developing COVID-19, particularly in those of non-white ethnicity. ZOE Symptom Tracker app data found the risk for a positive COVID-19 test was increased across racial minorities, not completely explained by other risk factors, comorbidities, and sociodemographic characteristics.
<p>3. How do we best understand and protect vulnerable groups? (RQ 22, 32, 36, 62, 102) - Risk prediction - Social & mental health</p>	<ul style="list-style-type: none"> Research using longitudinal research cohorts (ALSPAC and Generation Scotland) has shown increases in anxiety and lower wellbeing since COVID-19, particularly in young people. Zoe app data was used to identify six distinct symptom presentations, using time-series data that has the potential as a clinical prediction tool. Analysis of ONS data from the early phases (Dec'19-Mar '20) of the pandemic has shown that paradoxically lower than average mortality rates were observed.
<p>4. Impact on Non-COVID care provision (RQ29, 30, 94)</p>	<ul style="list-style-type: none"> Supply and demand for cardiovascular disease services have dramatically reduced, with potential for substantial, but avoidable, excess mortality during and after COVID-19. A study looking at the impact on provision of mental healthcare found significant reductions in caseloads and total contacts for home treatment teams March to May 2020, although they are now back on the rise (Stewart).
<p>5. Use of existing treatments (RQ18, RQ98)</p>	<ul style="list-style-type: none"> The RECOVERY Trial has shown that low-cost dexamethasone reduces death by up to one third in hospitalised patients with severe respiratory complications of COVID-19. The OpenSAFELY study found that inhaled corticosteroid use in people with asthma did not protect against COVID-19 related deaths.

Newly submitted & prioritised research questions (RQ104, 98, 87, 68) include:

- Identifying the proportion of the population not susceptible to COVID-19 i.e. groups who have tested positive but have not displayed symptoms.
- Linked to priority 5 above - the impact of starting inhaled corticosteroids early in the course of COVID-19 illness.
- The short term and long-term impact of hospitalisation on severe COVID-19 survivors.
- Linked to priority2 above - the extent the difference in mortality by ethnicity is driven by urban/rural environments and social deprivation.

Public Advisory Board Feedback

Vital to make full use of datasets available for research across the UK to identify differences and share learning within and between communities.



13 COVID-19 taskforce calls with **67** clinical and health data research leaders engaged



1330 academic, industry and NHS participants in COVID-19 Slack channel with 10 sub-channels



106 health data research questions identified – 40 prioritised



115 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 – 23 June 2020

There are now over 100 active projects accessing data from the national TREs, with a further 150 developing their applications.

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

1. England approaching primary care data coverage of over 50 million population through [GPES collection](#).

2. Establishing high quality Testing data flows remains priority.

3. Over 100 requests now approved and active, but ‘In development’ remains the largest category.

4. A range of data assets, complementary to the NHS Digital capability within England, are also being accessed.

- Fourteen (+3 since last report) approved research projects using [CPRD](#) resources, latest focussed on risk factors for different population groups.
- Twenty (+2) projects approved to access North West London COVID-19 data repository via [Discover-NOW](#), including system resilience research.
- [NHS COVID-19 Data Store](#) can be accessed via the single triage system for NHS data in England, along with the NHS Digital capability. Details of requests yet to be published.

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	94% coverage, 54% loaded	N/C - Approval process agreed		
Pillar 1 COVID-19 Testing Data (NHS/Public Health)				
Pillar 2 Testing data (UK Gov)	N/C - Data completeness & quality issues			
Pillar 3 & 4 Testing data	N/C - Data flows being specified across all 4 nations			
Community Prescribing	N/C - Expected July			
Critical Care (CHESS, ICNARC, SICSAG)	CHESS data quality issues			N/C - Under review
Personal Demographic Service				
Secondary Care				
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	44 (-)	33 (+3)	70 (+10)	6 (+3)	153 (+16)
Submitted for IG approval	5 (-)	0 (-5)	0 (-)	0 (-)	5 (-5)
Approved but not yet active	0 (-)	2 (-2)	1 (-)	0 (-)	3 (-)
Active research taking place	20 (+10)	24 (+5)	60 (+3)	0 (-)	104 (+18)