

HDRUK
Health Data Research UK



British Heart
Foundation

HSC Public Health
Agency
Research and Development



CANCER
RESEARCH
UK

Public Health
Scotland

NHS
Digital

 SAIL DATABANK

UK Health Data
Research Alliance

COVID-19 Health Data Research

02 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
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



COVID-19 Health Data Research recommendations – 02 June 2020

The volume of health data research on direct and indirect effects of COVID continues to grow rapidly, fuelled by weekly improvements to dataset availability. In contrast, in areas without an established approach to making national data available for research (e.g. in social care) progress is slowed as the effort is going into establishing new linkages, collaborations etc.

As this is a weekly report, our recommendations are largely unchanged week on week, changes from last week are underlined below. After next week we will be moving to fortnightly reporting.

1. Ensure data gathered from all swab & antibody testing programmes can be securely linked and used for research. This will require unparalleled cooperation across all four nations between NHS organisations, PHE, data custodians, academic endeavours, and technology partners, whilst maintaining public trust.
2. Commission independent meta-analysis of ethnicity studies in the UK and internationally. Ensure that further research, undertaken collaboratively with international partners where possible, addresses why BAME groups appear to have a higher rate of severe COVID-19 outcomes.
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.
4. Many insights are being generated from the Symptom Study App. Endorse the use of established Trusted Research Environments to accelerate insights from other “risk calculators”, symptom trackers and surveys, and consider integrating with targeted public health messaging and actions.
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.

Health data research on COVID is growing rapidly, with a 40% increase in COVID research questions and 60% increase in pre-print publications over three weeks. However, progress is slower where work on additional data collection and linkage is required

Priority research questions & insights – 02 June 2020

Priority research questions	Insights from ongoing studies (links provide further details):	SAGE Recommendation
<p>1. What impact has COVID-19 having on care home patients? (RQ63 et al)</p>	<ul style="list-style-type: none"> Core insights require the identification of care home residents in routinely collected data. An algorithm has been developed, which uses address matching and core health data to identify the population, with higher sensitivity and specificity than previously published efforts. This tool is supporting work across the UK and is being validated in Scotland, Wales and SW England. Ongoing Projects lead by teams at UCL, QMUL, University of Sheffield, University of Swansea/SAIL, University of Bristol and University of Edinburgh. 	<p>Continue to enhance data capture to understand the interconnectedness of care homes and other healthcare settings (via care givers and care receivers).</p> <p>Ensure an appropriate TRE is available to allow linkage of data to achieve such analysis.</p> <p>Enhance data capture to establish contact structures within a care home and how this impacts likelihood and outcome of local outbreak.</p>
<p>2. How do we best understand and protect vulnerable groups? (RQ 22, 32, 36, 62) & inform an effective phased lockdown release:</p> <ul style="list-style-type: none"> - Risk prediction - Social & mental health - Vulnerable groups 	<ul style="list-style-type: none"> Individuals with cancer have a significantly increased risk of infection compared to the general community. Those being treated with chemotherapy or immunotherapy are particularly at-risk of infection. <p>Data from the COVID Symptom Study app has been used to:</p> <ul style="list-style-type: none"> Predict COVID-19 likelihood without a test, based on symptom-clustering, therefore could play a role in identifying likely infections quickly, so that health authorities can act early. Discover that menopause may affect women's immunity to COVID-19, possibly due to lower oestrogen levels. This is currently being explored further. Predict personal nutritional responses to the food we eat to identify the best foods to eat to support your gut health and immune system during the COVID-19 pandemic. 	<p>Further develop, extend and utilise open “risk calculators”, symptom trackers (e.g. ZOE app) and surveys, to better communicate risk and more targeted public health messaging and actions.</p> <p>Directly endorse the use of trusted research environments, to enable an open approach to health research data access safely and securely. As successfully achieved using the COVID Symptom Study, delivered through the ZOE app, with data made securely accessible through the UKSerP platform.</p>
<p>3. How do we support the restoration of cancer/health care services? (RQ89)</p>	<ul style="list-style-type: none"> Around 2.4 million people in the UK are waiting for cancer screening, further tests or cancer treatment according to a new analysis by Cancer Research UK, which has estimated shortfalls in screening, referrals and treatment based on data for England. 	<p>Commission large scale analyses of the long-term impacts of health and social care changes during the COVID-19 lockdown. This will require access to linked data from a range of sources.</p>

<p>11 COVID-19 weekly taskforce calls with 62 clinical and health data research leaders engaged</p> 	<p>1310 academic, industry and NHS participants in COVID-19 Slack channel with 10 sub-channels</p> 
<p>62 volunteers in HDR UK's COVID-19 Public & Patient Group</p> 	<p>106 health data research questions identified – 34 prioritised</p> 
<p>91 COVID-19 pre-print publications</p> 	



Priority questions where research is proceeding, with no new insights since 27th May 2020

Activity currently focussed on data integration/linkage & the establishment of new data infrastructure to underpin large-scale research:

- How do we best understand and monitor population immunity to COVID-19? (Immunology & seroprevalence R01, 50)
- Why do BAME groups have an increased risk of severe COVID-19 outcomes (RQ34)?

 [Click here](#) for a link to the full prioritised list of questions, status, and prioritisation process

Dataset availability is increasing each week, and we have seen a five-fold increase in active research taking place in three weeks

COVID-19 dataset availability and status of projects using the data – 02 June 2020

1. Continued progress towards UK wide Primary Care data being available.

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	Being loaded	Approval process agreed		
Pillar 1 COVID-19 Testing Data (NHS/Public Health)				
Pillar 2 Testing data (UK Gov)	In internal build	Linkage work underway		
Community Prescribing				
Critical Care				Options under review
Personal Demographic Service				
Secondary Care				
Death registry				
Pillar 3 Testing data - antibody	Data flows being specified across all 4 nations			

2. Growth in active research taking place within SAIL Databank driven by access to COVID-19 symptom study and work directly for Wales Technical Advisory Group

# of COVID-19 Projects by TRE stage (change from previous week)	England (NHS Digital Data Processing Service)	Scotland (National Data Safe Haven)	Wales (SAIL Databank)	Northern Ireland (Honest Broker Service)	Total
In development	39 (+9)	27 (+1)	62 (+6)	3 (-)	131 (+16)
Submitted for IG approval	6 (+1)	3 (+1)	0 (-)	0 (-)	9 (+2)
Approved but not yet active	3 (-1)	0 (-)	3 (-1)	0 (-)	6 (-2)
Active research taking place	8 (+2)	19 (+3)	50 (+32)	0 (-)	77 (+37)

3. Collaborative efforts underway to scale up smaller immunity and testing studies to answer questions at pace

- Collaboration taking shape between Public Health bodies, UKCRC Tissue Directory Co-ordinating Centre and Health Data Research UK to answer immunity and re-infection questions at pace
- Building on COG-UK and the genomics metadataset with potential to bring together small data sets from the many single hospital/ university studies

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

NOTES

- TRE - Trusted Research Environment
- IG - Information Governance
- DPN – Data Provision Notice

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