

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

27 May 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 – 27 May 2020

Insights from health data have further accelerated, fuelled by new data sources made accessible in safe environments. In particular, there has been a big increase in requests for symptom tracker data and additional intensive care data has been made available in Wales and progress with primary care data in England. However, national testing data remains an issue.

Recommendations for SAGE based on current health data research insights:

1. (NEW) Ensure data generated from all swab & antibody testing programmes can be securely linked and used for research. This will require **unparalleled cooperation between NHS organisations, PHE, data custodians, academic endeavours, technology partners, whilst maintaining public trust.**
2. Commission independent meta-analysis on ethnicity analyses within the UK and with international studies and ensure that further research is undertaken to understand why BAME groups appear to have a higher rate of severe COVID-19 outcomes. This should be undertaken collaboratively with international partners as much as possible.
3. Enhance data capture on patients and staff in care homes to enable research on health, transmission and outcomes to the equivalent depth in NHS settings and provide clarity on appropriate use of national Trusted Research Environments for consolidation of relevant care home COVID-19 data.
4. Further develop, extend and utilise open “risk calculators”, symptom trackers (e.g. ZOE app) and surveys, integrated with targeted public health messaging and actions
5. Commission meta-analysis on outcomes across major disease groups to assess the long-term impacts of health and social care changes during the COVID-19 lockdown period, compared with previous 5 years, including research, frontline clinical teams and existing disease registry experts. Create COVID-19 registries for the four nations to provide an ongoing source of data on COVID-19 patients, akin to the national cancer registry. And ensure that government cross-departmental data linkage beyond health is enabled to understand the wider impacts of COVID-19 on all vulnerable populations.

We are seeing rapid growth in the research generated from health data, with a 50% increase in pre-print publications over 2 weeks (from 57 to 86)

Priority research questions, studies & insights – 27 May 2020



Priority research questions	Studies already working on this:	Insights	SAGE Recommendation
<p>1. How do we best understand and monitor population immunity to COVID-19? (Immunology & seroprevalence R01, 50)</p>	<ul style="list-style-type: none"> University Hospital Birmingham NHS Foundation Trust (UHBNFT) 	<ul style="list-style-type: none"> Asymptomatic health care workers were tested for current infection and for COVID specific antibodies in their blood (seroconversion) to determine whether they now have immunity. Asymptomatic infection led to variable seroconversion rates depending on department and role of the worker (e.g. high rates in housekeeping and acute medicine; lowest in intensive care & emergency medicine). Several seroprevalence studies are underway, recruiting participants from existing research cohorts, social and health care workers as well studies directed by Public Health England. However, fragmentation of data collection and availability for research analysis remains a significant challenge. 	<p>Ensure data generated from all swab & antibody testing programmes can be securely linked and used for research. This will require extensive cooperation and coordination between NHS organisations, PHE, data custodians, academia and technology partners, whilst maintaining public trust. Ensuring the right data is collected is essential.</p>
<p>2. Why do BAME groups have an increased risk of severe COVID-19 outcomes (RQ34)?</p>	<ul style="list-style-type: none"> Uni Hospitals Birmingham, Pioneer Hub & DECOVID NHSD & PHE CO-CIN ONS & Discover-NOW Imperial College King's College London 	<ul style="list-style-type: none"> Further research has shown that hospitalised COVID-19 patients of BAME background are on average much younger than white patients and have more co-morbidities such as higher levels of diabetes and high blood pressure. 	<p>Further research, ideally with international partners, to understand why BAME groups appear to have a higher rate of severe COVID-19 outcomes, should be prioritised.</p>
<p>3. What impact has COVID-19 having on care home patients? (RQ63 et al)</p>	<ul style="list-style-type: none"> HDR UK North – University of Sheffield & Lancaster UCL, HDR UK London University of Bristol, HDR UK South-West SAIL Databank University of Edinburgh, HDR UK Scotland 	<ul style="list-style-type: none"> Core insights require the identification of care home residents in routinely collected data. Work is being accelerated to increase the sensitivity of an algorithm, which uses address matching and core health data to identify the population, and to validate the algorithm across geographies. Analysis of care home residents in Wales (using the SAIL Databank) found that deaths more than doubled during lockdown, from 4% (2019) to 9% (2020). Ongoing research to analyse patterns of secondary care referral, patient management – cohort, testing, impact on non-COVID care and outcomes. 	<p>Enhance data capture on patients and staff in care homes with a focus on resident discharge and referral patterns and staff working patterns.</p> <p>Ensure alignment with existing datasets, providing clarity on appropriate TRE for consolidation of relevant care home COVID-19 datasets.</p>
<p>4. 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"> UCL, HDR UK London London School of Hygiene and Tropical Medicine, HDR UK London University of Glasgow, HDR UK Scotland KCL, HDR UK London 	<ul style="list-style-type: none"> Analysis of data from children with suspected COVID-19, including those with underlying health conditions, found no significantly increased risk of either contracting COVID-19 or severe complications, apart from those undergoing chemotherapy. Susceptibility to and transmission of COVID-19 amongst children and adolescents, is 56% lower than adults. There is limited evidence that children and young people play a lesser role in transmission of SARS-CoV-2. An early minimally symptomatic phase is often followed by deterioration in patients with COVID-19 infection. Adding age and a minimal set of common blood parameters to NEWS2 improves the risk stratification of patients likely to develop severe COVID-19 outcomes. The addition of a few common parameters is likely to be much easier to implement in a short time-scale than a novel risk-scoring system. Essential workers in health and social care have a higher risk of severe SARS-CoV-2 infection. This underscores the need for national and organisational policies 	<p>Better communication of individual risk will be vital to the ongoing public health messaging. Recommend that openly available online risk based tools (such as OurRisk.Co.V) are further developed & integrated into targeted public health messaging.</p> <p>Cross-departmental data linkage e.g. Health (NHS Digital) and Dept of Education is essential to understand the impact of COVID-19 on children and other vulnerable populations.</p>

<p>10 COVID-19 weekly taskforce calls with 67 clinical and health data research leaders engaged</p> 	<p>1306 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>99 health data research questions identified – 32 prioritised</p> 
<p>86 COVID-19 pre-print publications</p> 	



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

Active research has doubled in the last 2 weeks. Priorities to scale-up data use are unchanged: 1) National testing data 2) continue moving projects from development to IG scrutiny

COVID-19 dataset availability and status of projects using the data – 27 May 2020

1. Increasing research interest in the national antibody testing programme highlights need to include future data flows as a core dataset available for linkage.

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	First extract 29/5			
Pillar 1 COVID-19 Testing Data (NHS/Public Health)				
Pillar 2 Testing data (UK Gov)				
Community Prescribing				
Critical Care			Now available	Options under review
Personal Demographic Service				
Secondary Care				
Death registry				
Pillar 3 Testing data - antibody	Data flows being specified across all 4 nations			

2. There has been a 29% increase in active research taking place via the national TREs. The proportion of projects 'in development' is virtually unchanged at 68% of total.

# 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	30 (-)	26 (+1)	56 (+18)	3 (-)	115 (+19)
Submitted for IG approval	5 (-1)	2 (-)	0 (-)	0 (-)	7 (-1)
Approved but not yet active	4 (+2)	0 (-2)	4 (+3)	0 (-)	8 (+3)
Active research taking place	6 (+3)	16 (+4)	18 (+2)	0 (-)	40 (+9)

3. The Cardiovascular Consortium is working closely with the national data custodians in preparation for demonstrating the benefits of research at UK-wide scale

- Working with NHS Digital to help to shape the specification for the interim GP data flows for England and prepare for the initial data curation activity once the first extract 'lands' on 28 May.
- Sourcing hospital level data to undertake analysis of COVID-19 on cardiovascular activity to provide early insights similar to work undertaken by DATA-CAN, the health data research hub for Cancer.

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**