COVID-19 Health Data Research

7 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
Ian Young, Health & Social Care Northern Ireland

COVID-19 Health Data Research recommendations – 7 July 2020

HDRUK
Health Data Research UK

Health data research insights on COVID are continuing at a pace, with 107 research questions, 112 projects active within the national data Trusted Research Environments (TRE), a further 150 in development, and 127 pre-print publications. 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 on 30 Jun.
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).
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**: Data being consolidated on DHSC Exchange as an immediate practical solution. Opportunity to consider wider research engagement, by enabling data access through an appropriate Trusted Research Environment.
4	Accelerate access to restricted national datasets, since lack of availability is holding back crucial research.	**PROGRESS**: CHESS (COVID-19 Hospitalisation in England Surveillance System) now available from NHS Digital but concerns over data quality and coverage remain. Community prescribing data (England) not yet available. Expiry date of 30 September for current Control of Patient Information (COPI) notices is increasingly highlighted as a risk to research in England and Wales.
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. Proposed health and education dataset linkage for multiple uses recommended for approval by NHS Digital's independent advisory group.

Priority research questions with new insights generated this week – 7 July 2020

Health data research on COVID-19 continues to grow, now reaching 127 pre-print publications

HDR UK — Health Data Research UK

Priority research questions	Insights from ongoing studies (links provide further details):
1. Understanding immunity & susceptibility (R01, 50, 95, 102, 51, 54, 55, 104)	• Researchers have produced the largest and most systematic genetic investigation of human host proteins that interact with SARS-CoV-2 proteins, are related to virus entry, host hyperimmune or procoagulant responses, or are associated with COVID-19 severity. The findings will help to prioritise drug discovery and repurposing efforts. • A study exploring differences in blood samples between mild, severe and critical COVID-19 +ve patients found proteins that could potentially be early biomarkers to monitor disease severity or therapeutic targets.
2. How do we best understand and protect vulnerable groups? (RQ 22, 32, 36, 62, 102)	• Modelling of data from the COVID Symptom Study app can identify potential UK COVID hotspots.. The data has highlighted Dudley and Wolverhampton as areas that could be heading in the same direction as Leicester. • Prevention of obesity and physical activity are at least as important as physical isolation of severely obese individuals during the pandemic. • Early insights from major study on COVID-19 in care homes show that higher infection rates in care home residents are associated with higher rates in staff and vice versa. Higher infection rates among residents and staff are also associated with particular care home practices, such as increased use of bank or agency staff, as well as with geographic region (higher resident infection rates in London and the West Midlands, higher staff infection rates in the North East and Yorkshire and Humber), although regional differences may be affected by different patterns of testing in staff and residents over time. In addition, care homes where staff receive sick pay may have lower levels of infection in residents.
3. Impact on Non-COVID care provision (RQ29, 30, 94)	• Data from hospitals from the south of England to Scotland showed consistent and substantial reductions in total hospital admissions, emergency department attendances and cardiovascular presentations, diagnoses and treatments/procedures. The reductions started in the two weeks before lockdown, with limited recovery of some activities from mid-April 2020 (http://hospitalactivity.com/). These changes signify a major burden of indirect effects of the pandemic, and should be monitored and mitigated urgently.
4. Use of existing treatments (RQ18, RQ98)	• The RECOVERY Trial has shown that there is no clinical benefit from use of lopinavir-ritonavir in hospitalised COVID-19 patients studied in RECOVERY.
5. Health service planning (RQ 22, 94)	• The UK has an adequate supply of hospital beds nationally, but their unequal distribution meant that many trusts operated for a significant period above safe occupancy thresholds, despite substantial capacity in nearby trusts. • The UK started easing lockdown in England when levels of community transmission remained high. Modelling shows that even small changes in R with easing of lockdown can significantly impact expected cases and deaths, even if R remains <=1. Following an elimination strategy rather than one of maintaining R<1 would substantially mitigate the impact of the COVID-19 epidemic in England.

Public Advisory Board Feedback
The PAB emphasized the need for research into overlooked vulnerable sub-populations e.g. those receiving care at home, with disabilities and mental health.

14 COVID-19 taskforce calls with **74** clinical and health data research leaders engaged

1358 academic, industry and NHS participants in COVID-19 Slack channel with 10 sub-channels

107 health data research questions identified – 41 prioritised

127 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 – 7 July 2020

Rate of change across dataset availability, project pipeline and active projects has slowed over the last two weeks. Across England and Wales, researchers are increasingly seeking clarity about the future of the COPI notices.

HDR UK
Health Data Research UK

1. **Primary care data now available for COVID research across all UK nations except NI**

2. **Issues with all testing data except Tier 1 continues to limit research potential.**

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				Data flows agreed but not yet available
Pillar 1 COVID-19 Testing Data (NHS/Public Health)				
Pillar 2 Testing data (UK Gov)	UK level issues with quality, governance and data access reported			
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				

3. **Net reduction of 'In development' reflects change in phase, although continued growth in Wales where linked ONS 2011 Census data is now also available.**

# 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	35 (-9)	32 (-1)	77 (+7)	6 (-)	150 (-3)
Submitted for IG approval	4 (-1)	1 (+1)	0 (-)	0 (-)	5 (-)
Approved but not yet active	3 (+3)	0 (-2)	0 (-1)	0 (-)	3 (-)
Active research taking place	**23 (+3)**	**25 (+1)**	**64 (+3)**	**0 (-)**	**112 (+8)**

4. Rate of approved new projects also reduced for CPRD (+1) and Discover-NOW (+1)

5. Expiry date of 30 September for current Control of patient information (COPI) notices is increasingly highlighted as a risk to research. Concern that a lot of work will come to a halt if permissions are not extended or replaced in a timely fashion.

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