COVID-19 Health Data Research

19 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 – 19 May 2020

HDRUK
Health Data Research UK

Insights from health data are now being generated at pace, fuelled by new data sources made accessible in safe environments. In particular, there has been a big increase in the use of the Welsh and Scottish national data and progress with primary care data in England. However, national testing data remains an issue.

Recommendations for SAGE and UKRI/DHSC based on current health data research insights:

1. Commission independent meta-analysis on ethnicity analyses within the UK and with international studies, distinguishing between “population fatality rate” and “infection fatality rate”
2. Support SAGE sub-group to enhance data capture on patients and staff in care homes to enable research on health, transmission (“R” in care homes) and outcomes to the equivalent depth in NHS settings
3. Further develop, extend and utilise open “risk calculators”, symptom trackers (e.g. ZOE app) and surveys, integrated with targeted public health messaging and actions
4. Commission meta-analysis on outcomes across major disease groups compared with previous 5 years (e.g. cancer, cardiovascular disease, diabetes, respiratory, dementia)
5. 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
6. Ensure data flows from national testing programme are available for linkage, within trusted research environments

We are seeing rapid growth in the research generated from health data, with a 32% increase in pre-print publications since last week

Priority research questions, studies & insights – 19 May 2020

HDRUK
Health Data Research UK

Priority research questions	Studies already working on this:	Insights	SAGE Recommendation
1. Why do BAME groups have an increased risk of severe COVID-19 outcomes (RQ34, 37)?	• University Hospitals Birmingham, Pioneer Health Data Research Hub & DECOVID • NHSD & PHE • CO-CIN • ONS & Discover-NOW Health Data Research Hub • Imperial College • King's College London	National Cumulative, but inconsistent insights emerging, potentially due to difference in terminology, in particular 'risk' for population fatality rate vs 'risk' for infection fatality rate (see details of insights from individual studies here).	Initiate meta-analysis of COVID-19 studies concerned with incidence and severity of COVID-19 in different ethic groups and regional variations., e.g. NIHR Complex Reviews Unit (CRSU). Include international comparisons to compare outcomes for different ethnic groups by geography.
2. What impact has COVID-19 having on care home patients? (RQ63)	• HDR UK North – University of Sheffield & Lancaster • UCL • University of Bristol	Expanding and implementing local data collection to increase the specificity and power of research and avoid contradictory insights – exploring the potential value of a shared data resource to build on existing research enabled trusted research environments.	Enhance data capture on patients and staff in care homes to enable research on health outcomes to the equivalent depth in NHS settings. Develop routine care home data specification. Identify care home population in existing electronic medical records.
3. How do we best understand and protect vulnerable populations? (RQ 22, 32, 36, 62) to inform an effective phased lockdown release: - **Risk prediction** - **Social & mental health** - **Vulnerable groups**	• HDR UK London (UCL) & Uni of Cambridge and Oxford/Alan Turing Institute • University of Edinburgh • University of Swansea • University of Bristol • UCL	National • Online, prototype risk calculator (OurRisk.CoV) shows underlying health conditions increase 1-yr risk of death from COVID-19 5-fold (see Lancet study here). • Early results from HAPPEN-at-home study show both positive (e.g. more active, sleeping better) and negative (e.g. more screen time) effects of lockdown on children. • CovidLife study shows that 'lockdown' has raised major concerns about future employment and the economy, particularly in young adults and are accompanied by high levels of anxiety and depression. • Similar results from ALSPAC, which shows a worsening pattern in wellbeing/anxiety in adults of parental age. • Study on homeless people shows the positive impact of temporary accommodation to enable care & self-isolation upon onset of symptoms. International Dataset from Wuhan used to ID patients at risk of developing severe COVID-19, adds to growing evidence that simple blood tests are strong predictors of risk of serious disease.	Open, risk based tools (such as OurRisk.CoV) are further developed & integrated into targeted public health messaging. Government alignment to facilitate 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.
4. What is the impact on treatment and outcomes for non-COVID-19 disease (e.g. cardiovascular & cancer) (RQ29)	• BHF Data Science Centre and UK-wide CVD COVID UK consortium	National 76% reduction in urgent cancer referrals and 60% reduction in chemotherapy attendances compared to pre COVID-19 levels (details here) Substantial declines in presentation of acute coronary syndrome and Percutaneous Coronary Intervention activity to hospital (details here). International Netherlands - drop in cancer diagnoses overall by 26% (60% for skin cancer diagnoses); US - significant changes in cancer services, e.g. drops in cancer screening for cervical, breast and colorectal cancer (83% for PAP smears, 87% for mammograms, 90% for colonoscopies respectively - details here)	Commission meta-analysis on outcomes across major disease groups compared with previous 5 years (e.g. cancer, cardiovascular disease, diabetes, respiratory, dementia) – including research, frontline clinical teams and existing disease registry experts.

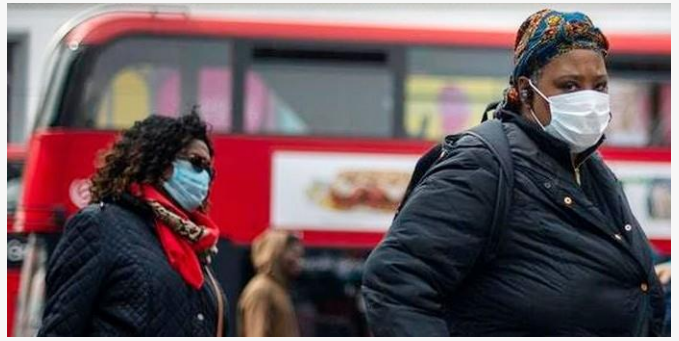
9 COVID-19 weekly taskforce calls with **67** clinical and health data research leaders engaged

1296 academic, industry and NHS participants in COVID-19 Slack channel with 10 sub-channels

62 volunteers in HDR UK's COVID-19 Public & Patient Group

72 health data research questions identified – 28 prioritised

75 COVID-19 pre-print publications



Click here for a link to the full prioritised list of questions, status, and prioritisation process

Priorities to scale-up data use: 1) national testing data 2) continue moving projects from development to IG scrutiny

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

HDRUK Health Data Research UK

1. The scaling up of the national testing programme provides a potentially powerful dataset but the expected data flows are not yet 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	DPN Issued 14/5			
UK Gov Testing data				
Community Prescribing				
Critical Care				Options under review
Personal Demographic Service				
COVID-19 Surveillance Testing				
Secondary Care				
Death registry				

KEY

- Data flows specified but not yet agreed
- Data flows agreed but not yet available for linkage
- Fully available

2. Expedited approval and fulfilment processes appear to be working well once requests are submitted for approval with less than 10% of projects at these stages of the pipeline

# 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 (-)	25 (+1)	38 (+7)	3 (+3)	96 (+11)
Submitted for IG approval	**6 (+4)**	**2 (-)**	**0 (-)**	**0 (-)**	**8 (+4)**
Approved but not yet active	**2 (-)**	**2 (-2)**	**1 (-2)**	**0 (-)**	**5 (-4)**
Active research taking place	3 (+1)	12 (+9)	16 (+2)	0 (-)	31 (+12)

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

3. COVID-19 symptom study app developed by King's College London, Twins UK and health science start-up ZOE is being widely used to support research with projects representing most of the active research projects reported by SAIL Databank which hosts the research data.

- Launched under two months ago it has 3.6m UK users who are reporting regularly. Highlighted loss of sense of smell or taste as a strong predictor of coronavirus as very early insight (1 Apr);
- Data securely transferred to SAIL Databank and listed on Health Data Research Gateway to enable access for non-commercial researchers beyond the original collaborators. Being accessed at local (e.g., councils and NHS trusts) and national level (e.g., Welsh Government, Food Standards Agency, MoD, DfE).

Datasets available for COVID-19 research via national TRES for England, Scotland and Wales.