

HDR UK Covid-19 Strategy

This is a 'live' strategy and will be continually updated

19 March 2020



Introduction

As the national institute for health data science, the Health Data Research UK (HDR UK) community is actively championing the use of health data to address the COVID-19 challenge. Institute members across our sites and hubs are contributing to this global challenge by forging partnerships with government, NHS, industry and academia. Teams are already working on risk factors, genomics, clinical trials, care pathways and surveillance to help, for example, identifying which retrovirals and which clinical practices have the best outcomes for patients with COVID-19 (see annex for further examples of current activity).

Such is the scale and immediacy of the challenge that we have developed a rapid, collaborative and agile strategic response to deploy our national capability through three priorities:

1. Leverage the best of the UK's health data science capability to **address the wider impact of the COVID-19** pandemic supporting vulnerable groups who will be hardest hit
2. **Accelerate safe access** to UK-wide priority data relevant to COVID-19 for research
3. **Co-ordinate and connect** national data science-driven research efforts related to COVID-19

Priority 1: Leverage the best of the UK's health data science capability to address the wider impact of the COVID-19 pandemic supporting vulnerable groups who will be hardest hit

The HDR UK community includes many of the UK's finest health data researchers and innovators who are already supporting the COVID-19 response: these include clinical trials, surveillance, genome sequencing, dynamic modelling, real-time dashboards and close working with virology and infection control teams, and are outlined in the Annex.

BREATHE, the HDR Hub for respiratory health is focusing all resources and efforts on COVID-19. BREATHE will help navigate respiratory health data sets from across the UK, creating a digital resource where trained, approved experts can access data safely and responsibly, on an unprecedented scale. The HDR Hubs for clinical trials, Real World Evidence and other disease areas: NHS DigiTrials, PIONEER, DATA-CAN, Gut Reaction, Discover-NOW, INSIGHT and the BHF Data Science Centre; are also focusing resources and expertise on COVID-19.

In addition, we are refocusing work across all our nationally coordinated priorities to deploy health data science to the COVID-19 challenge, including our work on Understanding the Causes of Diseases, Improving Public Health, Better Care, the Human Phenome and Applied Analytics.

We have identified opportunities to scale up this existing work to deliver UK wide impact and to accelerate insights. These include:

- Augmenting the clinical trials work on existing and retroviral treatments
- Real-time predictive analytics
- Monitoring whether measures we are introducing are working within the UK and internationally – and knowing when to lift restrictions
- UK wide surveillance building on projects in Scotland, North West London and Wales.

We also recognise the importance of sustaining activities in other disease areas, including cardiovascular disease, cancer, diabetes, eye disease, Alzheimer’s disease and inflammatory bowel disease as people living with these diseases may be the most adversely affected by COVID-19.

Whilst some of these potential projects can be resourced through redeployment, funding and access to data (outlined below) are also required.

Deliverables

- Coordinate an inventory of HDR UK COVID-19 research questions and activities and publish this on the [HDR UK website](#), questions such as “At the point of peak need, if patient requirements outstrip ventilation capacity, how should hospitals stratify and prioritise patients for ventilation?”
- Identify, support, accelerate and scale research and innovation projects that address new COVID-19 challenges by:
 - working with funders to shape calls
 - joining up research using longitudinal datasets in, for example, Wales, Scotland, North West London and internationally
 - coordinating different research studies being led by parts of the UK and identifying opportunities to scale through the HDR UK Hubs, Alliance, industry and international partners
- Identify COVID-19 specific deliverables that each HDR Hub and HDR National Priority (including sites) contribute towards.

Priority 2: Accelerate safe access to UK wide priority data relevant to COVID-19 for research

This COVID-19 research endeavour urgently needs access to individual-level, linked data from across the UK. This data must be safely available as near real-time as possible for research. We will work with the 27 members of the **UK Health Data Research Alliance** including national NHS bodies across the UK (NHS

Digital, NHSX, NSS, PHE, NWIS and NHSCIC) to improve the speed of access and utility to data. We will achieve this through the linkage of data collections that are already established to produce a national registry of COVID-19 patients (infected and recovered) for research.

Our research community has identified an initial set of priority datasets for this endeavour:

- COVID-19 testing & isolates data
- Primary care data
- NHS 111
- Prescribing data
- Secondary care data
- Critical Care data from Hospital Electronic Health Records
- Genomic data (including international sequence submissions)

We will work with national data custodians and regulators to identify the action required to enable the “priority” datasets to be linked and made available and updated daily (as far as possible) for research. For deeper electronic health record data, we will work with our Better Care Partnership network of digitally mature NHS sites.

We will also support efforts already underway including with UK BioBank, the COVID-19 Genomics UK (COG-UK) sequencing consortium based at Sanger Institute, UKCRC Tissue Directory and Coordination Centre and PHE’s efforts in the discovery and access of COVID-19-relevant samples and linked data, and the NIHR Health Informatics Collaborative (HIC) theme on critical care.

Whilst the potential public benefit of accelerating access is immense, it is vital that privacy is protected. We will work with data custodians to agree proportionate governance requirements which reflect the level of emergency and uses the “Five Safes” approach to protecting privacy. We will seek input from HDR UK’s existing national network of patients and members of the public, including members of the HDR UK Public Advisory Board and HDR Hub Public & Patient Advisory Groups, to ensure we are working to our value of transparency and that our work earns public trust. We will accelerate our Innovation Gateway development to build an appropriate environment to facilitate the research, supported by data use agreement templates.

Deliverables:

- Launch the **HDR UK COVID-19 workbench** as part of the **Innovation Gateway** that builds upon existing capability to improve availability, accessibility, interoperability, linkage and findability of data that will move research in COVID-19 farther and faster. This will comprise of integrated access to COVID-19 datasets, community developed tools deployed in a Trusted Research Environment.
- Engage all 27 Alliance members in adding COVID-19 relevant datasets for research onto the workbench.

- Engage other data custodians, including industry partners, in making COVID-19 relevant datasets accessible for research through the Gateway, and in line with the [UK Health Data Research Alliance Principles for Participation](#)
- Launch a coordinated effort across the HDR UK Public Advisory Board, HDR UK Hub and Site Public & Patient Advisory Groups, existing patient and public groups (e.g., useMYdata) and national partners such as Understanding Patient Data to involve patients, vulnerable populations and wider public in shaping the approach to data access.
- Put all COVID-19 related **metadata** from existing Alliance members, HDR UK studies (see Annex), and international sources on the HDR UK COVID-19 Workbench.

Priority 3: Co-ordinate and connect national data science-driven research efforts related to COVID-19

The HDR UK One Institute UK-wide model and existing collaborations with over 200 NHS, academia, industry, patient and public groups, enables HDR UK to connect and co-ordinate national COVID-19 data-driven research projects.

Deliverables

- Lead the **HDR UK COVID-19 Taskforce** with a weekly COVID-19 taskforce summit to support/co-ordinate the research effort.
- Provide a single, open SLACK channel ([#covid-19challenge](#)) as a point of information for all HDR UK investigators, collaborators and studies on COVID-19. This channel already has over 150 investigators (on day 3) and is growing fast.
- Leverage HDR UK's GitHub (<https://github.com/hdruk/covid-19>), linked to other endeavours, to provide unrestricted access to open source tools and pipelines and resources relevant to COVID-19 data science research.
- Provide a 'shop window' of opportunities for industry partners wishing to contribute to the challenge.
- Publish a weekly communications update on the HDR UK website on progress in partnership with Government, other researchers and national/international partners e.g. International Severe Acute Respiratory and Emerging Infection Consortium (ISARIC) reporting on real-time intelligence on COVID-19 HDR UK research expertise and opportunities to support the COVID-19 data science research effort with relevant industry partners.
- Collect and collate all relevant pre-print and open-access papers related to COVID-19 research in a central location to allow the HDR UK research community to 'stand on each other's shoulders'.

Principles of the HDR UK COVID-19 Taskforce

The HDR UK community has subscribed to the following ways of working that are vital to rapid progress on the COVID-19 pandemic:

1. Guided by the HDR UK values: Transparency, Optimism, Respect, Courage and Humility
2. [UK Health Data Research Alliance Principles for Participation](#)
3. [HDR UK Attribution Policy](#)
4. [HDR UK Development Principles](#)

We are heartened by the immediate response from our whole community in rising to this global health challenge and will work together in an agile way over the coming weeks and months to bring the best of HDR UK.

Annex: HDR UK activities aligned with COVID-19 and Taskforce leads

Live and continually being updated

	Lead	HDR UK affiliation	Activity (use-case)	Status
1	Richard Dobson and Angus Roberts	HDR UK London (King's)	<ul style="list-style-type: none"> Replaced one of the exemplar projects of their HDR UK National Text Analytics Project (Natural Language Processing of unstructured patient records) with one specifically addressing COVID-19. A number of data scientists working on a number of key questions identified by infection control and virology departments. 	Active
<p>Lay summary: The notes that doctors type up into the electronic health record of a patient who has tested positive for COVID-19 are critically important. They may include language that describes symptoms and complications in great detail. Within these notes may lie nuances that could be vital to understanding the development of the disease, the way it was transmitted in the first place, and the most effective treatment with least side effects. Richard Dobson, Angus Roberts and team are developing ways in which we can search the medical notes of all known COVID-19 patients rapidly, to glean information and insights that may help to understand more about this disease. They use a technique called Natural Language Processing. Developing these methods are part of a pre-existing national HDR UK project, however the team have switched their disease focus and are now applying the techniques they are developing to COVID-19.</p>				
2	Rosalind Eggo	HDR UK London (LSHTM); HDR UK Fellow	Quantifying the impact of respiratory viral triggers of chronic disease exacerbations using dynamic transmission models.	Active
<p>Lay summary: Roz is a HDR UK fellow who has been working on developing statistical models to help understand how viruses that affect our breathing spread throughout communities. When COVID-19 emerged, Roz shifted her focus to study this virus specifically. She and her team have been using latest data from China and beyond to measure how COVID-19 can exacerbate other pre-existing conditions.</p>				

3	Ewan Harrison	HDR UK Cambridge (Sanger); HDR UK Fellow	Seconded to lead the COVID-19 Genomics UK (COG-UK) sequencing consortium.	Active
<p>Lay summary: Ewan Harrison is a HDR UK fellow. Prior to COVID-19 his fellowship focussed on how microbes that live in our nose move from person to person, the end game being to understand how to prevent this, and stops these kinds of infection from running rife in our communities. He does this by studying and gleaning insight from the DNA sequence of these microbes. The approaches he and his team have developed can also be applied to COVID-19, which can also live in our noses (as well as in our mouths and faces – hence the advice to wash hands regularly, avoid touching our faces and catch coughs/sneezes in a tissue and bin it straight away). Since the emergence of COVID-19, Ewan has been seconded to lead the COVID-19 Genomics UK sequencing consortium, working to understand the DNA of the virus to look for “chinks in its armour” that we can exploit to control this pathogen.</p>				
4	Axel Heitmueller	Imperial College Health Partners and the Discover-NOW Hub	Working on a COVID-19 dashboard using the hub data to provide near real time information at a population health level. They are also designing a potential trial of Point of Care Testing equipment.	Active
<p>Lay summary: Axel runs HDR UK’s Discover-NOW hub, the Health Data Research Hub for Real World Evidence which brings together health data from the NHS in London and Manchester, with appropriate controls and consent, to support research to find new ways to diagnose and treat disease and make care for patients better and safer. Since the emergence of COVID-19, Axel and his team have turned their attention to:</p> <ul style="list-style-type: none"> • Providing the most up to date clinical information on people who have tested positive for the infection in Northwest London. Typically this kind of information becomes available weeks, months or even longer after cases of infectious diseases are first reported. The COVID-19 situation, is changing rapidly day-to-day; researchers, clinicians and commissioners need information in real-time, or as close to that as possible, in order to plan and improve services. Axel’s team are developing a dashboard using cases from the London area (the part of the UK most heavily burdened by COVID-19) to 				

show a near real-time picture of the spread of the disease. This could help us to flag hot spots where control measures should be increased, for example.

- Identifying patients at risk of complications and to follow care pathways. The Hub also plans to run predictive modelling platform using DataRobot machine learning to enable rapid analysis and modelling including identifying the patient variables that can be used for future high risk/vulnerable patient case-finding criteria.
- Supporting the NWL Gold Command using Hub data and resources.
- Securing daily pathology data into the Hub data set that includes COVID-19 test results (The Discover-NOW Hub already has linked primary and secondary care data).
- Exploring linkage of NHS 111 data which is being used as a front line service and likely to contain data on those referred to self-isolate but not tested and those referred to hospital

5	Harry Hemingway	Human Phenome National Priority	Insert COVID-19 related data on the Innovation Gateway	Active
6	Martin Landray	HDR UK Oxford; NHS DigiTrials Hub	Designed a study in ten days of different potential interventions for treating coronavirus.	Due to start end Mar 20

Lay summary: Martin leads HDR UK’s national priority in clinical trials. He leads a national team that is developing ways to easily recruit large numbers of the right patients to clinical trials and be able to conduct long-term follow-up with minimum effort, so that we can improve and speed up the development of new medical treatments. The team have been working on trials to help patients with various conditions. They are now focussing on recruiting as many patients as possible who have tested positive for COVID-19 across the UK. This is a lot harder than it sounds because the systems are not set up to be able to rapidly identify all UK COVID-19 patients in all four nations rapidly, so even that first step is a challenge. There are three potential treatments that might effectively treat those with the disease: opinavir-Ritonavir (commonly used to treat HIV); inhaled Interferon (usually given by injection to treat multiple sclerosis, hepatitis C, and some blood disorders); or dexathasone (a type of steroid, which is used in a range of conditions typically to reduce inflammation). For various reasons

any of these may also double-up as a treatment for COVID-19. The team developed this trial in record time (10 days) and is due to kick-off at the end of March 2020.				
7	Andrew Morris	HDR UK Central	Working in partnership with Gates Ventures about how we might support analytics at scale.	Under review
8	Liz Sapey	HDR UK Midlands; PIONEER (the Health data Hub in Acute Care)	Part of a team that has submitted two COVID-19 related bids which will require the hub's data and services.	Active
9	Aziz Sheikh	HDR UK Scotland (Edinburgh); BREATHE Hub	Will enhance an approach successfully used during the 2009 H1N1 pandemic to estimate epidemiological characteristics of COVID-19 in Scotland and monitor the effectiveness of any antiviral or other interventions deployed.	Active
<p>Lay summary: There is a lot to learn from previous pandemics. In 2009, there was a flu pandemic, dubbed swine flu (H1N1). At this time Aziz Sheikh, who is now a part of HDR UK Scotland and leads the BREATHE Hub, developed an approach that was successful in predicting the patterns, causes and risk factors associated with that virus, in Scotland. The approach also accurately and rapidly measured the effectiveness of medicines and behaviour changes. These were instrumental in bringing the 2009 swine flu under control. Aziz has just received funding to enhance this approach for use in the current pandemic.</p>				
10	Cathie Sudlow	BHF Hub; HDR UK Scotland	Tbc	Tbc

11	Helen Parkinson	HDR Cambridge (EMBL-EBI)	<p>Raw biomolecular data from sequencing, structural analysis and drug screens. Host and Viral Genome Sequence hosting (where research is consented).</p> <p>Helen and her team are working on acquiring raw biomolecular data from sequencing, structural analysis and drug screens etc. This is a national discussion with European countries and builds on Data Hubs. They are able to accommodate host and viral genome sequence where research data are consented. Helen and her team are already involved in a direct collaboration with Public Health England and seeking other ways to extend this work for the benefit of UK healthcare. Phil Quinlan has offered to 'donate' some of his teams to help analyse once ready.</p>	TBC
12	Charlie Davie	DATA-CAN Hub	<p>The DATA-CAN hub, led by Charlie Davie, aims to transform how cancer clinical data across the UK can be used to improve patient care. Charlie and his team are identifying use cases relating to COVID-19 and the information governance requirements to use their datasets for this. One suggestion is tracking particular cohorts e.g. immunosuppressed children in our paediatric cancer cohorts. This is an under-studied area internationally, however the need is great.</p>	
13	tbc	HDR UK Central	<p>Exploring whether it may be appropriate for HDR UK to apply for a combined S251 application to allow data use without consent in the public interest.</p> <p>[suggested by Charlie Davie, DATA-CAN Hub]</p>	
14	tbc	HDR UK Central	<p>COVID-19 models that underpin government policy to be made open access.</p> <p>[suggested during group discussion]</p>	
15	Phil Quinlan	HDR UK Midlands	<p>Phil leads the national UKCRC Tissue Directory and Coordination Centre. Phil is reaching out across their network of biobanks to find those that are able to support access to COVID-19 samples. He has already identified a biobank based at the Royal Free Hospital that should now be in a position to deliver samples</p>	

			for analysis. He has encouraged HDR UK members and beyond to get in touch should they wish to access these samples.	
--	--	--	---	--