

A National Health Data Research Capability to Support COVID-19 Research Questions

SAGE/UKRI/NIHR Reporting: 5 May 2020

Executive Summary

Health Data Research UK (HDR UK) has teamed up with the UK Health Data Research Alliance, including NHS Digital, NHS national data custodians in Scotland, Wales and Northern Ireland, and national and regional providers of specialist data to:

- Streamline and prioritise health data research expertise and resources to tackle the most important health data research questions relevant to SAGE, policy makers and the NHS
- Enable a scalable approach to linked data across the four nations
- Provide rapid access to secure analytical environments for researchers to answer these questions to improve understanding and treatment of coronavirus.

New health data research insights and progress this week, since the last [report](#) on 28 April, include insights in cardiovascular disease, psychological impacts of lockdown, new research questions on excess deaths in care homes and mental health, and increased accessibility of data in the four nations to support further research.

Update 5 May 2020

A. New Research Insights: Pre-peer Reviewed Findings are emerging across themes with multiple groups using datasets across the four nations. We are encouraging collaboration to help us to reach definitive answers to key research questions. Early outputs include:

- **Care homes (RQ63):**

Researchers at University College London are collaborating with the care home provider Four Seasons Healthcare (FSHC), and are extracting weekly data on COVID-19 cases, hospital admissions, deaths and the use of disease control measures across a national network of 200 care homes with >9000 residents. A similar project is being run by HDR UK North researchers at the Universities of Sheffield and Lancaster using data from a network of 60 care homes across the North East. Activities in Wales integrated into the SAIL COVID-19 eCohort are also exploring relevant analysis in partnership with Public Health Wales as well as within Scottish National Safe Haven. Early insights are expected in the next week.

- **Cardiovascular (RQ29):**

Insight: an approximate halving of the numbers of patients presenting to hospital with acute coronary syndromes in recent weeks compared with the same period last year¹ and an increase in

¹Work led by University of Oxford, Keele, Leeds and Imperial researchers, working with NHS Digital, NICOR and NHS England, as previously reported to SAGE

cardiovascular deaths of about 15% compared with the average over the last five years, peaking at the end of March about a week after lockdown and a month after the first reported case of COVID 19 in Scotland was reported (MedRxiv/2020/086231)². Further insights on the direct and indirect impacts will emerge through analyses of linked nationally collated datasets from all UK nations by the CVD-COVID-UK consortium, coordinated by the BHF Data Science Centre.

- **Psychosocial consequences of lockdown (RQ22, RQ24): [CovidLife](#)**

Insight: 50% of those in employment before COVID-19 are very confident of post-COVID employment, the rest have concerns and there has been a dramatic increase in feeling lonely and socially isolated is an early insight from an online survey developed by CovidLife ([Generation Scotland \(GS\)](#)), focused on the psychological, social and economic impacts of lockdown. By 30 April, there were 10,000 responses, ~80% from Scotland, ~20% from England.

- **Ethnicity (RQ34 & 47):**

Further statistical analysis across a variety of national and regional datasets is a particular focus, given the challenges in determining direct vs confounding contributions within previous observations. One additional emerging insight has been shared by Imperial College London, who have shown that people of black ethnicity may be at increased odds of mortality (Further info [here](#)).

- **Immunology and serology (RQ 01, 03, 50):**

Collaborative efforts have been initiated, supported by coordination from the Wellcome, to capitalise on the UK's diverse population data resources that have significant potential to address urgent serology and immunology questions. These include: (i) a serial ~1year sero-survey of existing National Blood Donor Cohorts³. An online COVID-19 questionnaire will start this week, sample collection is planned for end of May/beginning of June, subject to additional funding support. (ii) Snapshot sero-survey to provide information across age, geography, socio-economic status, ethnicity, incorporating ~40,000 participants consented in existing cohorts⁴ throughout UK, with planning in progress for the use of remote home testing kits and examination of the variation in infection status. Currently seeking ethical approval.

B. Prioritised Research Questions:

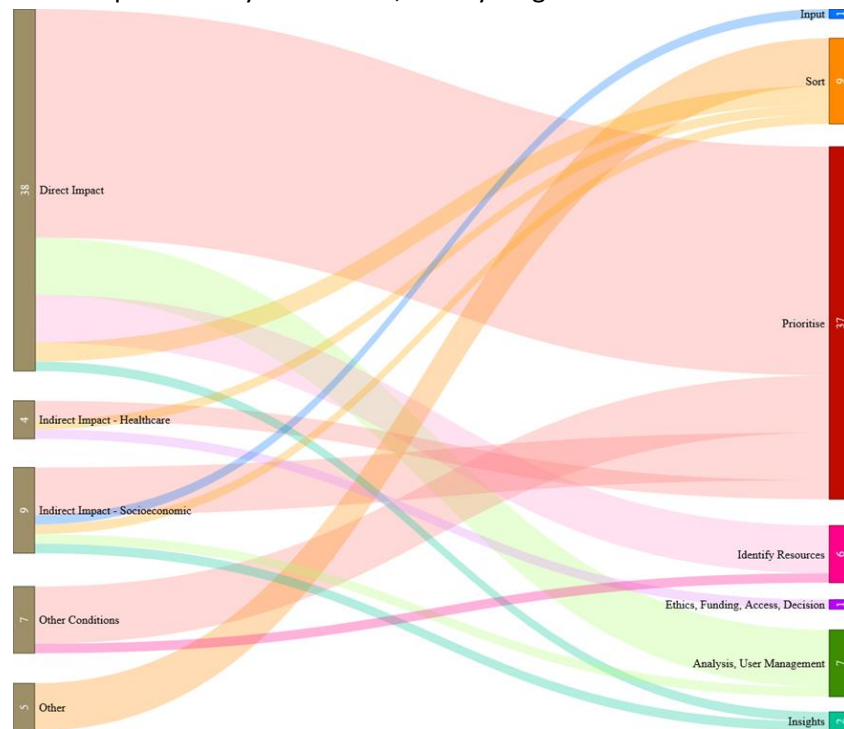
- The Research Funnel¹ has received **64 questions** to date There is an increasing spread of questions across the SAGE priority areas, and more questions are into analysis and insight stage, as demonstrated by the increasing number of insights shared in this paper.
- The three newly prioritised questions this week cover (i) vulnerable patient groups (ii) excess deaths in care homes and (iii) mental health.
- 8 previously prioritised questions have not progressed from question to specific data requests. This is surprising as the prior expectation was that the constraint would be access to data rather than submission of the data access. It also suggests that the enhanced health data assets (see section C) are not yet being as fully utilised by researchers as they could be to support the COVID-19 response.

² Work led by University of Edinburgh researchers who examined cause-specific mortality data from Scotland.

³ INTERVAL, COMPARE, and STRIDES

⁴ ALSPAC, Born in Bradford, Generation Scotland, EXCEED, and Twins UK

The following figure shows the questions by SAGE area², and by stage in the funnel:



¹ Health Data Research Question funnel – process description provided in Appendix 1 [21 April SAGE paper](#)

² 51 questions have been scored by the Urgent COVID Review Group to date and a further 3 have been prioritised by NIHR

C. Linked NHS data in English, Scottish, Welsh & Northern Irish Trusted Research Environments (TRE)

Linkage and access improvements to key data assets to support the COVID-19 response are summarised in Appendix 2. Provision of COVID-19 linked datasets with national coverage in Scotland and Wales is in place, utilising existing infrastructure. Permissions are being sought to increase further the datasets available and frequency of updates.

Specific developments in Scotland TRE:

- Since the start of April, the Scottish National Safe Haven team⁵ has been working to develop a COVID-19 Research Database⁶, bringing together datasets from across the Scottish Public Sector to answer questions posed by the research community and Scottish Government officials.
- The first datasets (see Appendix 2) are predominately health data related directly to the COVID-19 outbreak. With the exception of finalising permission to access primary care data with the GP Community all datasets included in original proposal are available.

⁵ Team is led by Public Health Scotland’s electronic Data Research & Innovation Service (eDRIS). It works closely with the Edinburgh Parallel Computing Centre (EPCC, who provide the National Safe Haven infrastructure) and colleagues in Scottish Government, including National Records Scotland, and NHS Scotland.

• ⁶ The COVID-19 Research Database is a deidentified but linked resource allowing eDRIS to respond quickly to requests for project data and to update that data regularly. It sits within the National Safe Haven security perimeter but is separate from the normal preparation space for non-COVID-19 projects (c. 150+ annually) but, as far as possible, it uses all of the existing processes for data management, preparation, and supply to researchers – giving confidence to data providers.

- Safe Haven is also hosting data for some national UK projects (e.g., some of the ISARIC work).
- Datasets that we are currently working on adding from across the Scottish Government include: education, other child-related data, and care home data.
- Efforts are underway to explore data federation with the SAIL TRE in Swansea. Furthermore, a parallel activity around financial data is developing which may, in the future, provide key economic data.
- Access to the data is through the normal eDRIS projects mechanism with an accelerated approval process put in place to support high priority COVID-19 efforts. These permissions have been granted by data controllers to host data only for the duration of the pandemic and to inform the recovery phases.
- A Scottish Government Data Taskforce Delivery Group is meeting on a weekly basis to review progress and to help to unblock issues.

Specific developments in Wales TRE:

- Welsh Government has invested in the Secure Anonymised Information Linkage (SAIL) databank www.saildatabank.com since 2008 as a component of the national research infrastructure. It is hosted on the UK Secure eResearch Platform (UKSeRP) managed by Swansea University and provides global access to accredited research projects. No data needs to leave the system.
- All data supplied to SAIL are first de-identified by the NHS Wales Informatics Service using robust encryption. Any potential use of data is scrutinised by information governance specialists and members of the public to ensure that privacy is protected and analyses are in the public interest.
- An extensive list of datasets supporting this work is to be found in Appendix 2. Permissions are being requested to extend the list further (e.g., NICOR, Ambulance service, ONS Census data) and also increase frequency of data flows (e.g., primary care, birth data).
- Work is progressing on developing federated data analysis capability with other TREs across the UK and internationally.
- Adopting Welsh Government's *One Wales* philosophy a COVID19 Technical Advisory Group was created to provide scientific input. This includes collaborative a cross-institutional data linkage team to harness the skills of data scientists, biological scientists, clinicians and policy makers to provide timely evidence to inform policy and practice and minimise impact of the epidemic on the population.

Authors:

John Aston, Home Office (SAGE sponsor)
Charlie Davie, DATA-CAN
Tom Denwood, NHS Digital
Ben Gordon, Health Data Research UK
Ronan Lyons, SAIL Databank
Carole Morris, NHS National Services Scotland
David Seymour, UK Health Data Research Alliance
Rhoswyn Walker, Health Data Research UK

Caroline Cake, Health Data Research UK (lead)
John Deanfield, NICOR
Clara Fennessy, Health Data Research UK
Sara Hiom, CRUK
Andrew Morris, Health Data Research UK
Mark Parsons, Scotland National Safe Haven
Cathie Sudlow, BHF Data Science Centre
Nilesh Samani, British Heart Foundation

Appendix 1 - List of Priority Questions

The questions are being prioritised by the HDR Urgent COVID Review Group⁷ and categorised by SAGE priorities: Direct Impact (DI), Indirect Impact – Healthcare Pressures (I-HP), Indirect Impact – Socio-Economic (I-SE), and Other Conditions (OC). Questions of direct relevance to the SAGE unanswered science question list, (as shared on 24 April) are shown in **bold**. The top prioritised questions (prioritised 8 out of 10 and above) and their status is provided below. Questions that had already prioritised via NIHR are shaded in green, questions related to Cardiovascular Disease are shaded in orange, cancer in blue and care homes in yellow. We will also provide a short overview of the emerging insight for each of the priority questions where applicable. The full list of 65 active questions is available in our [Matchmaker Tool](#).

MM Ref.	SAGE Category /SAGE Question	Date of input	Question	Prioritisation Score (Median)	Time since input (days)	Moved stage? ↑/~/↓	Current funnel stage	(Expected) data requested	Question posed by	Support to answer question	Emerging Insight ⁸
RQ18	DI Treatments & Preventative Measures (SQ26)	02/04/2020	RECOVERY Can Lopinavir-Ritonavir vs Interferon β vs low dose corticosteroids be effective in treating COVID 19 test +ve hospitalised patients?	Auto prioritised	33	—	7	Complete	Peter Horby, Professor of Emerging Infectious Diseases and Global Health, University of Oxford,	NHS Digitrials (Martin Landray)	05/05/2020: >8,900 patients recruited. Initial results expected in June. The key issue at present is obtaining data on the number of cases admitted per hospital per day so that recruitment can be benchmarked and remedial actions targeted. The mortality data should be flowing this week.
RQ05	DI, OC	02/04/2020	ISARIC-CCP What are the clinical characteristics of COVID19 positive patients; what are the determinants (genetic, other omic, prior medical history, other) of good and poor outcome; and how can knowledge of this help to target clinical and public health strategies?	Auto prioritised	33	—	7	England, Scotland population/ demographic datasets	Cathie Sudlow, HDR UK Scotland, Edinburgh; and BHF Data Science Centre,	Cathie Sudlow	28/04/2020: Ongoing support with linkage in Scotland.

⁷ Urgent COVID Review Group Membership - provided in Appendix 3 [21 April SAGE paper](#)

⁸ Pre-peer reviewed findings

MM Ref.	SAGE Category /SAGE Question	Date of input	Question	Prioritisation Score (Median)	Time since input (days)	Moved stage? ↑/-/↓	Current funnel stage	(Expected) data requested	Question posed by	Support to answer question	Emerging Insight ⁵
RQ01	DI Virology (SQ11)	02/04/2020	COG-UK Can study of the whole virus genome enable scientists to monitor changes at a national scale, reveal how the virus is spreading and whether different strains are emerging?	Auto prioritised	33	—	7	Population clinical datasets (e.g. HES)	Ewan Harrison, HDR UK Cambridge; HDR UK fellow	Ewan Harrison	28/04/2020: 10,567 Viral Sequences now available online https://www.cogconsortium.uk/data/ This is the largest national repository of viral genome sequences in the world. Continuing to explore linkage to clinical data in each devolved nation.
RQ34 and RQ47	DI, I-SE	06/04/2020	<p>How can we explain the differences in COVID-19 cases and deaths by key socio-demographic factors of age, sex, socioeconomic status, geographical location and ethnicity (BME groups)? Specifically, why do BME groups appear to have increased risk of severe COVID outcomes (e.g. ventilation and mortality)? How do these vary across the UK and compare against international cases? How can this information be used for more effective stratification and strategies?</p> <p>Initial aim: To investigate and quantify variation in the incidence (test-positive rate) and outcome of COVID-19 on individuals from Black, Asian and Ethnic Minority (BAME) populations in the UK.</p>	10	29	—	8	University Hospitals Birmingham NHS Foundation Trust & DECOVID Partner Trusts	Rhoswyn Walker Chief Science Strategy Officer, HDR UK; Aligned question submitted by Prof. Melinda Mills, Leverhulme Centre for Demographic Science, University of Oxford	Elizabeth Sapey, HDR UK Pioneer Health Data Hub & University Hospitals Birmingham NHS Foundation Trust	28/04/2020: Clinical data from >2200 patient with confirmed SARS-CoV-2 infection plus demographics, ethnicity, baseline co-morbidities, social deprivation index and outcome (death within the censor date) included in a preliminary observational analysis. Current evidence suggests those of South Asian ethnicity may be at risk of worse COVID019 outcomes (including more likely to be younger and have no co-morbidities), but further studies needed to understand this. Potential to replicate this with DECOVID partners in London Foundation Trusts. Further details can be found here .
						—	7	NHSD and PHE datasets); CHES, SGSS, NICOR	PHE, Jem Rashbass, NHSD		

MM Ref.	SAGE Category /SAGE Question	Date of input	Question	Prioritisation Score (Median)	Time since input (days)	Moved stage? ↑/-/↓	Current funnel stage	(Expected) data requested	Question posed by	Support to answer question	Emerging Insight ⁵
			As above			—	8	ISARIC CCPUK study and CO-CIN data set		Calum Semple, Liverpool	28/04/2020: ISARIC CCP-UK & CO-CIN study led by Calum Semple, commissioned by SAGE, paper reviewed at SAGE on 28 April.
						—	5	ONS, ICNARC, DECOVID, DISCOVER-NOW		Ben Humberstone, ONS DISCOVER-NOW	
						—	8	Individual patient electronic medical records (demographic, clinical, laboratory, radiological, comorbidity and outcome data)		Imperial College London et al.	05/05/2020: Additional emerging insight has been shared by Imperial College London, who have shown that people of black ethnicity may be at increased odds of mortality (Further info here).

MM Ref.	SAGE Category /SAGE Question	Date of input	Question	Prioritisation Score (Median)	Time since input (days)	Moved stage? ↑/~/↓	Current funnel stage	(Expected) data requested	Question posed by	Support to answer question	Emerging Insight ⁵
RQ03	DI Immunology and serology (SQ4, SQ6)	07/04/2020	How do we support the scale-up of COVID-19 testing, by making sure that the data that is provided on the confirmed state of COVID-19 diagnosis and antibody levels is robust and reliable?	9.5	28	—	7	Tissue directory	Philip Quinlan, Head of Digital Research Service at University of Nottingham, UKCRC Tissue Directory Coordinating Centre	Philip Quinlan	05/05/2020 - Also see Section B above 28/04/2020: Ongoing, results so far: Helping PHE to plug gap in samples from those with mild symptoms. Connection has been made with project in Barts hospital, which involves the taking of blood samples from staff members (some of whom also displayed mild symptoms). Working with SME (Tissue Solutions Ltd) who are assisting with logistics to connect collection and sending samples for research. 20 requests for samples so far.
RQ50	DI Immunology and serology (SQ4, SQ6)	20/04/2020	How long will immunity last following infection, and what level of protection is available against different mutant strains?	9	15	—	3		Chris Wigley & Parker Moss, Genomics England		05/05/2020 - See Section B above
RQ04	DI Epidemiology (SQ1)	07/04/2020	How can we accurately measure the ongoing prevalence of COVID-19 in the population following identification of a "good enough" antibody diagnostic? (This requires representative and random sampling from the whole or at risk sub-populations)	9	28	—	3		Rhoswyn Walker, Chief Science Strategy Officer,		05/05/2020 - Also see Section B above

MM Ref.	SAGE Category /SAGE Question	Date of input	Question	Prioritisation Score (Median)	Time since input (days)	Moved stage? ↑/-/↓	Current funnel stage	(Expected) data requested	Question posed by	Support to answer question	Emerging Insight ⁵
RQ32	DI, OC	06/04/2020	Understanding vulnerable patients: How are underlying conditions defined, and what is the impact of infection on a range of outcomes, and what are the benefits of 'shielding' and other preventive interventions?	9	29	—	4		Harry Hemingway, Professor of Clinical Epidemiology at UCL	Harry Hemingway & Spiros Denaxas	28/04/2020: Initial understanding of underlying conditions supported by HDR UK COVID-19 Phenotype open resource. covid19-phenomics.org
RQ41	DI Clinical Health Care Management (SQ20)	07/04/2020	Are there any treatments which show evidence of improving outcomes for patients infected with coronavirus? Clinicians are having to make real-time decisions today, on the best possible treatment options for critically ill patients without robust evidence of harm or potential benefits of the therapeutic interventions. Better use of routine medication data could provide additional evidence to inform these decisions prior to the definitive outcomes of clinical trials.	9	28	—	3		Liz Sapey, Alastair Denniston, Tanya Pank Hurst, PIONEER & INSIGHT Hubs		
RQ59	DI, I-HP, ISE, OC Clinical Health Care Management (SQ23)	21/04/2020	What are the direct and indirect effects of SARSCov2/COVID-19 on incidence, presentation, diagnosis, management, and prognosis of cancer patients and which patients are most susceptible to these direct and indirect effects?	8.5	14	—	3	Linked datasets to including cancer registry, primary care and hospital records	Jonine Figueroa, Academic researcher, University of Edinburgh	Alvina G. Lai and Harry Hemingway UCL & HDR UK London Charlie Davie, Mark Lawler & the DATA CAN Health Data research Hub team	05/05/2020: Analyses of cause-specific mortality in Scotland have found an increase in the number of cancer deaths from 1 March when the first case of COVID 19 was reported in Scotland, peaking at the end of March, one week after lockdown with around 13% excess cancer deaths compared with the average for the same dates over the last 5 years, and declining thereafter. 28/04/2020: Investigated the potential increase in excess cancer deaths both directly

											and indirectly related to the COVID-19 crisis. As a result of the emergency, they estimate at least 6,270 additional deaths in newly diagnosed cancer patients alone. Over 78% of excess deaths occur in cancer patients with ≥ 1 comorbidity, highlighting a particularly vulnerable group of patients. The researchers analysed recent weekly data from major cancer centres in the UK and found a 76% decrease in urgent referrals from GPs for people with suspected cancers and a 60% decrease in chemotherapy appointments compared to pre-COVID-19 levels. Further details and pre-print can be found here .
RQ52	DI, OC Immunology and serology (SQ4, SQ6)	20/04/2020	What is the most effective approach (ease, cost and impact) for identifying people as recovered/immune so that they can come out of lockdown?	8.5	15	—	3		Chris Wigley & Parker Moss, Genomics England		
RQ11	DI, I-HP, OC Clinical Health Care Management (SQ16 & SQ23)	02/04/2020	Where hospitals have EHRs is it possible to provide real time data on outcomes per COVID-19 admission by age and by co-morbidities by hospital? To understand whether there are hospitals that appear to have better outcomes for particular co-morbidity sub-groups (indicating potentially more effective interventions to learn from)?	8.5	33	—	7		Jose Sousa, CTU Manager, School of Medicine, Dentistry and Biomedical Sciences, Queens University Belfast		

MM Ref.	SAGE Category /SAGE Question	Date of input	Question	Prioritisation Score (Median)	Time since input (days)	Moved stage? ↑/-/↓	Current funnel stage	(Expected) data requested	Question posed by	Support to answer question	Emerging Insight ⁵
RQ29	I-HP, OC Clinical Health Care Management (SQ16 & SQ23)	02/04/2020	What is the influence of COVID 19 epidemic in the UK and the NHS response to this on presentation, management and prognosis of non-COVID disease, in particular cardiovascular diseases such as MI and stroke?	8.5	33	—	6-7		Cathie Sudlow, HDR UK Scotland, Edinburgh; and BHF Data Science Centre	BHF Data Science Centre	05/05/20: An academic/NHSD/NHSE/ NICOR group have shown an approximate halving of acute coronary syndrome presentations to hospital in recent weeks compared with the same time last year. Approvals are being sought for broader set of analyses across all four nations and covering a broader set of cardiovascular conditions and procedures.
RQ48	DI, I-HP Clinical Health Care Management (SQ23)	12/04/2020	Cancer and COVID-19; how do we manage cancer optimally through a public health crisis?	8	23	—	3		Mark Lawler, Faculty of Medicine, DATA-CAN		
RQ27	I-SE, OC	13/04/2020	Can we predict the likelihood of ICU admission from COVID-19 patient demographics, pre-existing conditions, symptoms and clinical data collected on hospital admission?	8	22	—	3	Patient level deidentified hospital records for COVID patients	Marko Balabanovic, Medopad		

MM Ref.	SAGE Category /SAGE Question	Date of input	Question	Prioritisation Score (Median)	Time since input (days)	Moved stage? ↑/-/↓	Current funnel stage	(Expected) data requested	Question posed by	Support to answer question	Emerging Insight ⁵
RQ57	DI, OC	20/04/2020	How can we build from existing studies (e.g. Kings and internationally) that have looked at the effect of ACE inhibitors and Angiotensin Receptor Blockers on predisposition to COVID-19, as well determining the severity of COVID, to ensure that all potential confounders are robustly addressed using larger and more diverse datasets? Particularly including older patients who are likely to be on multiple antihypertensives?	8	15	—	4	Linked datasets (primary care, secondary care, COVID lab data, mortality data, intensive care data) across all four nations of the UK	Professor Sir Munir Pirmohamed, University of Liverpool	Cathie Sudlow BHF Data Science Centre	28/04/2020: Small UK study (from Kings) which shows ACEI were protective against severe COVID + two studies from China which show the same. Another study from China shows a small (p=0.06) increase in risk. All of these are on preprint servers and have not been peer reviewed. Difficult to account for all confounding factors in these studies. Good to look at ACE-I and ARBs separately, as well as together.
RQ51	DI, I-SE Immunology and serology (SQ 1 & SQ6)	20/04/2020	Asymptomatic carriers: do they exist, and what proportion of the population fit into this category? Does this differ with age?	8	15	—	3		Chris Wigley & Parker Moss, Genomics England		
RQ62	I-HP	28/04/2020	How has the delivery of primary care to vulnerable patient groups changed during the NHS response to the COVID-19 pandemic, and have these changes led to an increase in adverse events (e.g. hospital attendance, death)? Specifically, patients with chronic conditions (including multimorbidity), mental health problems, frailty, learning disabilities and vulnerable children.	8	7	↑	3	Primary care with linked secondary care - national GPES extract with linked hospital data.	John Macleod, University of Bristol		

MM Ref.	SAGE Category /SAGE Question	Date of input	Question	Prioritisation Score (Median)	Time since input (days)	Moved stage? ↑/-/↓	Current funnel stage	(Expected) data requested	Question posed by	Support to answer question	Emerging Insight ⁵
RQ63	DI, I-SE	28/04/2020	What are the direct and indirect impacts of COVID-19 on care home residents? Given that there are nearly 5,000 excess deaths/week in care homes [ONS stats 28 April], and only about 40% are coded as COVID deaths – what is the explanation for the other 60% of excess deaths in care homes?	8	7	↑	3		Caroline Cake, HDR UK Central		
RQ24	I-SE SQ23 (Clinical/healthcare management)	07/04/2020	What are the long-term mental health impacts of COVID-19 social isolation measures?	8	28	–	3		Rhoswyn Walker Chief Science Strategy Officer, HDR UK Central		05/05/2020: This question was re-prioritised after a further round of scoring. See emerging insights from Generation Scotland – COVIDLife study in Section A above.
RQ07 and RQ14	DI, I-HP Data Science & Engineering (SQ32)	02/04/2020	Can we use data science to support front line decision making in Intensive Care Units? E.g. at the point of peak need, if patient requirements outstrip ventilation capacity, how should hospitals stratify and prioritise patients for ventilation?	8	33	–	4		Simon Ball, Medical Director, University Hospitals Birmingham;	Hubs – Pioneer, Discover NOW	
RQ30	OC	02/04/2020	What is the influence of pre-existing cardiovascular disease on outcomes of COVID-19 infection?	8	33	–	4-5		Cathie Sudlow, BHF Data Science Centre,	BHF Data Science Centre	
RQ22	I-SE Behavioural Science (SQ45)	02/04/2020	What are the psychological, social and economic consequences of policies to limit the spread and flatten the peak of COVID 19?	8	33	–	7		David Porteous and Cathie Sudlow; HDR UK Scotland (on behalf of Generation Scotland and		05/05/20: CovidLife (www.covidlife.co.uk) is an online survey developed by Generation Scotland (GS) (www.generationscotland.org), which is open to anyone aged 18+ who is a UK resident and focuses on

MM Ref.	SAGE Category /SAGE Question	Date of input	Question	Prioritisation Score (Median)	Time since input (days)	Moved stage? ↑/-/↓	Current funnel stage	(Expected) data requested	Question posed by	Support to answer question	Emerging Insight ⁵
									other UK cohorts)		psychological, social and economic impacts of 'lockdown'. Invitations to GS volunteers commenced on 17 th April, with additional awareness-raising via social media. By 30 th April, there were 10,000 responses, ~80% from Scotland, ~20% from England, showing an early bias towards females (70:30), and an estimated response so far of 35% from GS volunteers. Initial data show that: while 50% of those in employment before COVID-19 are very confident of post-COVID employment, the rest have concerns, with 15% reporting they are not very confident or not at all confident; 50% of men and 60% of men have made contact with an old friend; there has been a dramatic increase in feeling lonely and socially isolated. A full report will be prepared shortly, posted on the GS web site and shared with researchers, stakeholders and policy makers.

MM Ref.	SAGE Category /SAGE Question	Date of input	Question	Prioritisation Score (Median)	Time since input (days)	Moved stage? ↑/-/↓	Current funnel stage	(Expected) data requested	Question posed by	Support to answer question	Emerging Insight ⁵
RQ23	DI, I-SE Behavioural Science (SQ39)	02/04/2020	Socioeconomic inequalities: Analysis by postcode IMD. What's the best way to provide targeted and tailored messages to diverse communities?	8	33	—	4		Linsey Hovard		
RQ35	DI, I-SE Clinical and Health Care Management (SQ17)	02/04/2020	How can we ensure that we fully understand variations in response to COVID-19 infection at the molecular, environmental, social and economic levels, by effectively coordinating the UK's longitudinal population studies to gain a much richer understanding of disease progression and outcomes?	8	33	—	3		Mary De Silva, Debbie Lawlor, Martin Tobin, John Danesh, Nic Timpson, & David Porteous; Wellcome Trust COVID-19 Longitudinal Population Study Steering Group		
RQ36	DI	02/04/2020	How can we maximise the speed and power of host genomic studies internationally to inform drug development?	8	33	—	3		Martin Tobin + (International) COVID-19 Host Genomics Initiative		
RQ39	OC	02/04/2020	Current guidelines recommend shielding is carried out for patients receiving immunosuppressants, however there is mixed evidence as to whether these patients will have poorer outcomes following coronavirus infection. Could we compare the outcomes data for patients who are receiving / not immunosuppressants and validate whether this population group are more vulnerable?	8	33	—	3		Liz Sapey, Alastair Denniston, Tanya Pank Hurst, PIONEER & INSIGHT Hubs		

MM Ref.	SAGE Category /SAGE Question	Date of input	Question	Prioritisation Score (Median)	Time since input (days)	Moved stage? ↑/-/↓	Current funnel stage	(Expected) data requested	Question posed by	Support to answer question	Emerging Insight ⁵
RQ40	OC	02/04/2020	Are there any concomitant treatments/ongoing prescribed medication which are making the outcomes of coronavirus infection worse for patients? This information would help clinicians to understand if there are any ongoing treatments which should be stopped as a priority when patients present with suspected COVID19. For example, help to better understand existing theoretical associations between anti-hypertensives and NSAIDs and COVID-19 outcomes	8	33	—	3		Liz Sapey, Alastair Denniston, Tanya Pank Hurst, PIONEER & INSIGHT Hubs		

Appendix 2a – National Linkages (Wales)

Datasets available for linkage	Collector	Controller (if different)	Status (RAG'ed)	Metadata on Gateway?	Cohort Coverage	Data Period Coverage (as at 05/05/2020)	Frequency of updated
ADBD Annual District Birth Daily (ONS Births)	ONS	ONS / NWIS	Permission being sought		Wales	n/a	Daily
ADBE Annual District Birth Extract (ONS Births)	ONS	ONS / NWIS	Available		Wales	2019-09-23	Yearly
ADDD Annual District Death Daily (ONS Deaths)	ONS	ONS / NWIS	Available		Wales	2020-04-18	Daily
ADDE Annual District Death Extract (ONS Deaths)	ONS	ONS / NWIS	Available		Wales	2020-04-01	Monthly
CAPD – Cancelled Admitted Procedures Dataset	NWIS		Available in SAIL but not requested for use currently		Wales	2020-03-31	Monthly
CARE Care homes index	n/a derived	Originally derived from CIW	Available		Wales	n/a	One-off
CARS Congenital Anomaly Register and Information Services for Wales	PHW		Available (but needs update)		Wales	2018-12-31	Yearly
CCDS Critical Care Data Set	NWIS		Available		Wales	2020-03-11	Monthly
CCDS Consolidated Death Data Source (combination of WSDS, MPI (Master Patient Index) & ONS deaths)	NWIS		Available		Wales	2020-04-22	Daily
CVSD COVID-19 Sequence Data	Public Health		Permission granted, currently being sequenced and collected		Wales	n/a	To be confirmed
CVST Zoe Symptom Tracker App (Linkable version)	ZOE Global Limited		Permission granted, currently being loaded		Wales linked – but UK wide data	n/a	Daily
CVST Zoe Symptom Tracker App (Unlinkable version)	ZOE Global Limited		Available		UK	2020-04-29	Daily
CVVP COVID-19 Vulnerable People list (12 week self-isolators)	NWIS		Available		Wales	2020-02-29	One-off (updates in the future if completed)

Datasets available for linkage	Collector	Controller (if different)	Status (RAG'ed)	Metadata on Gateway?	Cohort Coverage	Data Period Coverage (as at 05/05/2020)	Frequency of updated
CYFI Cystic Fibrosis Register	UK Cystic Fibrosis Trust		Available (but needs update)		UK	June 2018	6-monthly
DATW – Diagnostics And Therapy services Waiting times	NWIS		Available in SAIL but not requested for use currently		Wales	2020-03-31	Monthly
EDDD Emergency Department Data Daily	NWIS		Available		Wales	2020-04-29	Daily
EDDS Emergency Department Data Set	NWIS		Available		Wales	2020-04-24	Monthly
EDUW Education data on schools and pupils	Welsh Government (WG)	WG / NWIS	Available		Wales	2018 academic year	Yearly
ICNC ICNARC – Intensive Care National Audit & Research Centre	ICNARC		Permission granted, currently being loaded		Wales	n/a	Weekly
LIMS Laboratory Information Management System (COVID test results)	NWIS		Available		Wales	2020-05-03	Daily
NICO – NICOR audits and Registers	NICOR	HQIP/NICOR	Permission being sought		Wales	n/a	To be confirmed
NCCH National Community Child Health database	NWIS		Available (but needs update)		Wales	2019-03-25	Monthly
NHSO NHS 111 Call data (telephony, activity, 999)	NWIS		Permission granted, currently being processed and loaded		Wales	n/a	Monthly
NSWD National Survey for Wales Dataset	WG	WG / NWIS	Available		Wales	Defined sweeps	One-off (yearly sweeps)
ONSC Office of National Statistics Census	ONS		Permission being sought		Wales	n/a	One-off
OPDW Out Patient Dataset for Wales	NWIS		Available		Wales	2020-03-31	Monthly
OPRD – Outpatient Referral Dataset	NWIS		Available in SAIL but not requested for use currently		Wales	2020-03-31	Monthly
PEDW Patient Episode Database for Wales	NWIS		Available		Wales	2020-03-31	Monthly

Datasets available for linkage		Collector	Controller (if different)	Status (RAG'ed)	Metadata on Gateway?	Cohort Coverage	Data Period Coverage (as at 05/05/2020)	Frequency of updated
RTTD – Referral to Treatment Times		NWIS		Available in SAIL but not requested for use currently		Wales	2020-03-31	Monthly
WAST	Welsh Ambulance DataSet	WAST		Permission being sought		Wales	2018	Monthly
WCSU Unit	Welsh Cancer Incidence Surveillance Unit	PHW		Available (but needs update)		Wales	2016-12-31	Yearly
WDDS	Wales Dispensing DataSet	NHS Shared Services		Permission granted, currently being loaded		Wales	n/a	Monthly
WDSD	Welsh Demographic Service Dataset	NWIS		Available		Wales	2020-03-31	Weekly
WHSD	Welsh Health Survey Dataset	WG	WG / NWIS	Available		Wales	Defined sweeps	One-off (yearly sweeps)
WLGP	Welsh Longitudinal General Practice (Daily COVID codes only)	EMIS, Vision, Informatica	GP Practices / NWIS	Available		Wales (100%)	2020-05-03	Daily
WLGP	Welsh Longitudinal General Practice	EMIS, Vision, Informatica	GP Practices / NWIS	Increased cadence of existing data source currently in negotiation		Wales (80%)	n/a	Monthly
WLGP	Welsh Longitudinal General Practice	EMIS, Vision, Informatica	GP Practices / NWIS	Available		Wales (80%)	2020-04-01	Quarterly
WRRS	Wales Results Reporting Service	NWIS		Available (but needs update)		Wales	2018-07-30	Monthly

Appendix 2b – National Linkages (Scotland)

Datasets available for linkage	Collector	Controller (if different)	Status	Metadata on Gateway?	Coverage	Frequency of updated
NHS Scotland General Practice	Albasoft ESCRO GP Extraction	GP Practices / NHS Boards	Permission to access still to be agreed with the GP Community	No	Scotland	Daily
NHS Scotland COVID-19 lab testing data	ECOSS database	Public Health Scotland	Available	No	Scotland	Daily
Scotland Prescribing Data	Practitioner and Counter Fraud Services	Public Health Scotland / NSS	Available	Yes	Scotland	Monthly
Death registrations	National Records Scotland (NRS)	National Records Scotland (NRS)	Available	Yes	Scotland	Weekly
Acute hospital records (inpatient, outpatient and ITU)	Public Health Scotland	Public Health Scotland	Available	Yes	Scotland	Monthly
ISARIC-CCP	ISARIC	ISARIC	Currently being collated	No	UK wide	
NRS Stillbirths	National Records Scotland (NRS)	National Records Scotland (NRS)	Available	Yes	Scotland	Monthly
NRS Births	National Records Scotland (NRS)	National Records Scotland (NRS)	Available	Yes	Scotland	TBC
SMR02	Public Health Scotland	Public Health Scotland	Available	Yes	Scotland	Monthly
SICSAG Daily (ICU)	Public Health Scotland	Public Health Scotland / Scottish Intensive Care Audit Steering Group	Available	No	Scotland	Weekly
SICSAG Episodes (ICU)	Public Health Scotland	Public Health Scotland / Scottish Intensive Care Audit Steering Group	Available	No	Scotland	Weekly
Diabetes Covariates	Scottish Diabetes Research Network	NHS Boards	Available	No	Scotland	N/A
Emergency Department attendances (A&E)	Public Health Scotland	Public Health Scotland	Available	Yes	Scotland	TBC
NHS24 (111 Calls)	National Services Scotland	NHS 24	Available	Yes	Scotland	Daily
GP Out of Hours	National Services Scotland	Public Health Scotland	Available	Yes	Scotland	TBC
Scottish Ambulance Service (SAS)	National Services Scotland	Scottish Ambulance Service (SAS)	Available	Yes	Scotland	TBC

Datasets available for linkage	Collector	Controller (if different)	Status	Metadata on Gateway?	Coverage	Frequency of updated
SBR	National Services Scotland	Public Health Scotland	Available	Yes	Scotland	Monthly

