Data and Connectivity National Core Study

Feasibility assessment: rapid access to admissions data

24th February 2021

Summary of Requirement

As we enter the next phase of the pandemic, SPI-M requires data to enable research modelling of the impact of vaccinations. Matt Keeling (Warwick) has identified the following specific needs:

- More models within the SPI-M consortium, that **give greater diversity** by accessing **different data sources** – **this gives greater robustness to predictions.**
- **Incorporate more data on the robustness of vaccination – especially with hospitalization and mortality as key endpoints.**
- **Ability to model vaccine escape** e.g. lower immunity to the South African Variant
- **Move away from scenario modelling towards optimizing for a fixed goal; for example, a goal of minimizing restrictions to keeping hospital admissions below a threshold.**

This specification lays out the requirements for a data feed(s) which would allow **timely identification of acute/secondary care admissions**. Ideally, such data would be incorporated into the "vaccine pathway" that provides data on demography to event-based information on vaccination status, antigen status (PCR/LFTs) co-morbidities (from GP record), ethnicity, hospitalization date, in-hospital outcome (discharge or death).

Current data feeds (e.g. SUS in England) are compromised by being dated (submitted and available for analysis at least three weeks' post discharge), incomplete (e.g. ECDS only captures admissions via A&E) or lacking sufficient granularity (e.g. NHSE SitRep data reports number of admissions, not diagnosis and cannot be linked at individual level).

Working in close partnership with colleagues in NHS and NHSD we are looking at innovative options, including national data feeds, but there is a critical need to rapidly establish whether complementary data flows will also enable timely identification of acute/secondary care admission to support research and modelling related to the roll out of the C-19 vaccination programme. This may be on a large sub-set of the population (eg at regional level). Such research will inform policy related scenarios on subsequent lifting of restrictions. Ideally this data would not require new data collection.

Groups are invited to submit evidence that they may be able to deliver on this need.

Proposals must be submitted by email with a clear subject heading (**Rapid Access Admissions data**) to **dataconnectivity@hdruk.ac.uk** and received by **12noon GMT 5th March 2021.**

Purpose of data feed

To understand the safety and effectiveness of the COVID-19 vaccination programme, its impact on policy decision making, and wider research, there is a need for rapid access to acute/ secondary care admissions data. This data will be used to answer key research questions and meet the needs of epidemiological modelers amongst others:

Modeling requirements

- How effectively do vaccines reduce hospital admission?
    - How does this vary by ethnicity, age, gender, co-morbidities, previous infection?
    - How does this vary between vaccines? and/or with dosing intervals?
    - How does this vary over time and with different variants?
- What impact do vaccines have on outcomes following hospital admission: length of stay, ventilation, death?

Safety of vaccines requirements

- Does waning of immunity from the vaccine lead to the occurrence of vaccine-associated enhanced disease (VAED)? (sub-group analysis as above)
    - Is there any relationship to prolonged dose interval?
- What is the incidence and type of Adverse Events of Special Interest (AESI) resulting in acute/ secondary care admission)? NB AESI as defined by the MHRA and focused on AESIs resulting in acute/ secondary care admission, see **Appendix 1**.
    - How does this vary by ethnicity, age, gender, co-morbidities, previous infection?
    - How does this vary between vaccines?
- What are the rates of serious anaphylaxis post-vaccination? (sub-group analysis as above)

The Data and Connectivity National Core Study, under the leadership of Professor Sir Munir Pirmohamed, is working with multiple partners across the four nations including NHS England, NHS Digital, OpenSafely, Q Research and others to enable multiple data sources for research that may allow these questions to be answered. Current national data feeds allow us to answer some, but not all, of these questions. Specifically, near real-time data on hospitalisation are poor with most data sources limited to subsets of patients e.g. those in ICU, or capturing patients post-discharge which creates an unacceptable time delay in the availability of data for research; a delay which will potentially be greatest for patients admitted with the most serious conditions. Ben Goldacre of OpenSAFELY has written a brief paper- concluding “we don’t know whether a perfect dataset exists”, We therefore wish to explore whether existing national data feeds or complementary data sources (e.g. Great Northern Care Record, NW eHealth, One London or others) may be enabled for research purposes to answer these questions.

Specification details

Aim

Enable a near real time flow of admissions data into national Trusted Research Environment(s) which are openly accessible for research access (ONS/NHS Digital/ SAIL/ Scotland/ Northern Ireland) and/or other research environments (OpenSAFELY) where it can be linked to vaccine data other key datasets, is accessible to researchers and allow analysis to help answer key policy questions and faster identification of serious adverse events resulting from C-19 vaccination than current approaches.

Data requirements:

Category	Data items	Notes
Recipient ID and linkage data	Forename	Data required for ID and linkage. NB if NHS # can be provided other variables are not essential.
	Surname	
	Address	
	NHS #	
	Sex	
	Date of birth	Age on admission if not possible
	Ethnicity*	*May be derived by linkage to national vaccine dataset, census or primary care records*
Vaccination status* *May be derived by linkage to national vaccine dataset*	Date of vaccination and dose	
	Type of vaccine	
COVID infection status * *May be derived by linkage to national Pillar 1/2 testing data*	Pillar 1 /2 test date	Where the date and result of multiple COVID tests are known data is known, all should be listed. As a priority this should capture COVID infection status on hospital presentation
	LFT date	
	COVID test result	COVID infection positive/ negative
	Viral variant signature	*High level viral variant signature may be derived by linage to COG-UK data in national TREs*
Co-morbidities*	GP record	*May be derived by linkage to primary care records*
Presentation to hospital details	Date of admission to acute care setting	
	Location (care setting)	Name of secondary care centre, organisation code
	Primary reason for admission - Presentation diagnosis	This information should be provided in descriptive form and coded using SNOMED CT
	Primary reason for admission - Presentation symptoms	This information should be provided in descriptive form and coded using SNOMED CT
	COVID-19 infection status (if know)	
Discharge details* Not required for unfinished episodes *May be derived by linkage to SUS+*	Date of discharge	
	Status on ultimate discharge	e.g Discharged on clinical advice or with clinical consent; Self discharged, or discharged by a relative or advocate; Died; Not applicable (patient still in hospital)
	Destination post-discharge	e.g. Same NHS hospital site; Other NHS / Independent hospital site; Non-hospital destination within the UK (e.g. home, care home, penal institution), non-UK destination (e.g. repatriation); No discharge destination, patient died in unit

*We are keen to explore whether the variables marked as * and shaded are available within the dataset or whether linkage within the TRE will be required.

Data flow times

Update frequency	Data must be submitted at least **weekly**
Time lag	Data must be submitted within **<1 week of admission date**, a shorter submission delay is preferable **NOTE** submissions must include unfinished episodes for which discharge data should be left blank

Data must be shared from the secondary / acute care centre with a national TRE (in England **preferably ONS SRS**) where it can be linked to vaccination, testing and census data and made available to researchers and statutory bodies for monitoring and research. The legal basis for data sharing will need to be confirmed.

Next steps

Brief proposals are invited from groups who might be able to meet this need, answering the questions laid out below.

1. **Proposed approach** to include:
    - Data items which will be provided and consideration of their completeness and accuracy
    - The size and characteristics of the population that will be covered by the dataset
    - Update frequency and time delay for data feed
    - Proposed national TRE to receive data (this will be assumed to be ONS by default in England)
        - The above points should highlight how the data feed adds value beyond existing data flows and access routes.
    - Mechanisms to ensure that the data will be made available to any approved researcher (overcoming any information governance or access barriers)
    - Indication of whether retrospective data would be available OR potential start date for data feed (retrospective data would be used as a comparator and is not essential)
    - Legal basis for sharing data
    - Engagement required from the Data & Connectivity NCS and national TRE providers
    - Evidence of capability to deliver: experience and expertise of the project team, routine processes and technology to support delivery

2. **Indicative timelines** for set-up
    - There is a priority need for these data flows to be rapidly established. It is anticipated that any initiatives will have **started by 1st April 2021** and **data flows will be fully operational by 30th April 2021. A more detailed timeline is provided below.**

3. **Indicative resource requirements**
    - NOTE: Funding is available for resources to:

- Establish the data collection, curation and flow into a national TRE (NHSD, ONS Secure Research System, SAIL, Scottish National Safe Haven, or NI Honest Broker Service)
- Operate the data feed for **18 months** (1st April 2021 – 31st March 2022, funding can not be extended into the 2022_23 financial year)
- Ensure that the data is **openly accessible to any accredited researchers** through the TRE
- Provide data updates **at least weekly**

Proposals must be submitted by email with a clear subject heading (**Rapid Access Admissions data**) to **dataconnectivity@hdruk.ac.uk** and received by **12noon GMT 5th March 2021.**

Applications meeting these submission requirements will be considered the final versions to be used for consideration. We will assess based on the following criteria:

- Coverage, quality and timeliness of the data flow (i.e. value added beyond existing data flows) (50%)
- Capability to rapidly establish dataset in TRE with minimal barriers for onward research use (30%).
- Value for money (20%)

Timelines

Date	Activity
26 February 2021	Call for feasibility assessment launched
03 March 21, 5pm GMT	Deadline for submission of clarification questions from potential applicants
05 March 21, 12 noon GMT	Applications submitted
12 March 21	Successful applications notified
01 April 21	Projects commence
30 April 21	**Live data flows** into TRE which are: • Updated weekly • Accessible to researchers • Listed on HDR UK Innovation Gateway
01 June 21	**Milestone 1:** review of data flows, granularity, quality and access to confirm value added over existing data flows. • Project teams will be asked to provide evidence against the data requirements listed in the proposal. This evidence will be reviewed by TRE providers, D&C NCS and an independent review panel to confirm "Go" /"No go" for the remainder of project
31 March 2022	**Milestone 2:** Annual review of data flows, granularity, quality, access and extent to which data is meeting research needs.
30 September 2022	**Project close**

Appendix 1

Table 1: List of AESI defined for COVID-19 vaccines (May 2020) *Source: COVID-19 Vaccines: Safety Surveillance Manual*
Indicative SNOMED codes can be provided for the below conditions but would be primarily for analysis purposes.

AESI	Brighton Collaboration case definition status	Link to access definition	Recommended length of post-vaccine surveillance
Vaccine-associated enhanced disease	Case definition submitted for publication Sept 2020	Link will be provided-	1 year
Multisystem inflammatory syndrome in children	Under development and targeted for Oct 15, 2020	For all under development – they will be posted at time of submission for publication	1 year
Acute respiratory distress syndrome	Under development and targeted for Oct 15, 2020	-	1 year
Acute cardiovascular injury (microangiopathy, heart failure, stress cardiomyopathy, coronary artery disease arrhythmia, myocarditis)	Under development and targeted for Nov 15, 2020	-	1 year
Coagulation disorder (thromboembolism, hemorrhage)	Under development and targeted for Nov 15, 2020	-	1 year
Acute kidney injury	Planned start in Sept and targeted completion by Nov 30, 2020	-	1 year
Generalized convulsion	Published 2004	10.1016/j.vaccine.2003.09.008	LA vaccines: 4 weeks Others: 1 week
Guillain Barré Syndrome	Published 2011	10.1016/j.vaccine.2010.06.003	4-6 weeks
Acute liver injury	Planned start in Sept and targeted completion by Nov 30, 2020	-	4-6 weeks
Anosmia, ageusia	Planned start in Sept and targeted completion by Nov 30, 2020	-	4-6 weeks
Chilblain – like lesions	Planned start Jan 2021 and targeted completion by Apr 30, 2021	-	4-6 weeks
Single organ cutaneous vasculitis	Published 2016	10.1016/j.vaccine.2016.09.032	4-6 weeks
Erythema multiforme	Planned start Jan 2021 and targeted completion by Apr 30, 2021	-	4-6 weeks
Anaphylaxis	Published 2007	10.1016/j.vaccine.2007.02.064	2 days
Acute aseptic arthritis	Published 2019	10.1016/j.vaccine.2017.08.087	
Meningoencephalitis	Published 2007	10.1016/j.vaccine.2007.04.060	LA vaccines: 4 weeks
Acute disseminated encephalomyelitis	Published 2007	10.1016/j.vaccine.2007.04.060	4-6 weeks
Thrombocytopenia	Published 2007	10.1016/j.vaccine.2007.02.067	4-6 weeks


HDRUK
Health Data Research UK


UK Health Data
Research Alliance


Office for
National Statistics