



**UK Longitudinal Linkage Collaboration**

Population Health Sciences

Bristol Medical School

Canyng Hall

39 Whatley Road

Bristol BS8 2BN

# The UK Longitudinal Linkage Collaboration (UK LLC)

# DATA ACCESS AND ACCEPTABLE USE POLICY

**PUBLIC**

**Version 1**

**8 March 2022**

---

## Contents – UK LLC Data Access and Acceptable Use Policy

<b>1</b>	<b>PURPOSE .....</b>	<b>4</b>
<b>2</b>	<b>INTRODUCTION TO THE UK LLC.....</b>	<b>4</b>
<b>3</b>	<b>PRINCIPLES .....</b>	<b>5</b>
3.1	Five Safes.....	6
3.2	Commitment to LPS participants .....	6
3.3	Commitment to data owners.....	7
3.4	NCS data sharing principles.....	7
3.5	Commitment to researchers .....	8
<b>4</b>	<b>DATA TYPES AND DISCOVERY .....</b>	<b>8</b>
4.1	LPS data.....	8
4.2	Linked data.....	9
4.3	Synthetic data (not yet applicable in the UK LLC).....	10
4.4	Derived data.....	10
4.5	Harmonised data.....	10
4.6	Metadata.....	10
4.7	Data quality, coverage and completeness.....	11
4.8	Data discovery.....	11
<b>5</b>	<b>ELIGIBILITY AND REQUIREMENTS TO ACCESS DATA .....</b>	<b>11</b>
5.1	Eligibility.....	11
5.2	Application governance (legal documentation).....	12
5.3	Requirements.....	14
5.4	Intellectual property .....	14
5.5	Transparency and a community of researchers .....	14
5.6	LPS requirements .....	14
5.7	NHS Digital requirements .....	15
5.8	Penalties for misuse .....	15
<b>6</b>	<b>CHARGES TO ACCESS DATA.....</b>	<b>15</b>
<b>7</b>	<b>MAKING AN APPLICATION .....</b>	<b>15</b>
7.1	Stage 1.....	16
7.2	Stage 2.....	16
7.3	Stage 3.....	16
7.4	Appeal process.....	17
7.5	Amendment process.....	17
7.6	Application timescales .....	17

---

<b>8</b>	<b>WORKING IN THE UK LLC TRE</b> .....	<b>17</b>
8.1	Two factor authentication .....	18
8.2	Permitted access period .....	18
8.3	Project names across the project lifecycle .....	18
8.4	Project folder .....	18
8.5	Project Participant IDs.....	18
8.6	Project Git .....	18
8.7	Project DOI .....	18
8.8	User support .....	19
8.9	Audit.....	19
<b>9</b>	<b>OUTPUTS</b> .....	<b>19</b>
9.1	Derived data.....	19
9.2	Manipulated data.....	19
9.3	‘Data Out’ requirements and responsibilities.....	19
<b>10</b>	<b>PUBLISHED OUTPUTS</b> .....	<b>20</b>
10.1	Authorship and acknowledgements .....	20
10.2	Open Access publication policy.....	20
10.3	Research papers.....	20
10.4	Presentations and other publications.....	21
10.5	Media communications .....	21
<b>11</b>	<b>CONFIDENTIALITY AND DATA SECURITY</b> .....	<b>21</b>
11.1	User security requirements .....	21
11.2	IT equipment and connecting to the UK LLC.....	21
11.3	Reporting breaches and risks.....	21
<b>12</b>	<b>APPENDICES – DATA OWNERS’ SPECIAL CONDITIONS AND REQUIREMENTS</b> .....	<b>22</b>
12.1	Appendix 1: NHS Digital .....	22

## 1 PURPOSE

This policy document describes the **rules, processes and procedures** involved when:

- 1) Applying to access data in the UK Longitudinal Linkage Collaboration (UK LLC) Trusted Research Environment (TRE).
- 2) Accessing the UK LLC TRE as an approved user.
- 3) Producing outputs based on data held in the UK LLC.
- 4) Ensuring the reproducibility of research based on data held in the UK LLC.

**This policy will help to ensure a transparent, consistent and safe approach for using the UK LLC resource to conduct COVID-19 (C-19) research that delivers public benefit.**

## 2 INTRODUCTION TO THE UK LLC

Established in October 2020 at the request of the UK Government's Chief Scientific Adviser, Sir Patrick Vallance, the UK LLC was established by the COVID-19 Longitudinal Health and Wellbeing National Core Study (LHW NCS), one of the UK Government's six NCS for C-19 research. The UK LLC integrates data from many major UK Longitudinal Population Studies (LPS) into a centralised research resource for investigation into high priority C-19 research questions.

**The UK LLC has three main objectives:**

- 1) To bring together de-personalised data from more than 20 of the UK's LPS and systematically link data to their participants' C-19 relevant health, administrative and environmental records in the UK LLC's TRE, which is based on the [Secure eResearch Platform](#) (SeRP UK) infrastructure.
- 2) To permit safe access to the UK LLC TRE to large numbers of UK-based researchers so they can conduct public benefit C-19 research investigations using data from multiple LPS, the NHS and other government agencies, other research studies and newly commissioned data.
- 3) To offer the UK LLC TRE as a long-term solution for data linkage and secure pooled analysis for the longitudinal community, in order to support a broader range of research studies.

Initially supported by HM Treasury (01/10/2020–01/03/2021), UK Research and Innovation (UKRI) is now the accountable body for the LHW NCS and UK LLC. The Medical Research Council administers the funding on behalf of UKRI (grant code MC\_PC\_20030).

Ethical approval for the UK LLC was obtained from the NHS Health Research Authority's Haydock Research Ethics Committee (REF: 20/NW/0446, IRAS code: 290946).

The UK LLC resource is open to applications from any UK-based bona fide researcher conducting research into C-19 – this includes C-19 as a virus, the wider impact of the pandemic on the health and wellbeing of the UK population, and the delivery of services during the lockdowns and throughout the pandemic and subsequent period of recovery.

### Management of the UK LLC resource

The UK LLC is a collaborative endeavour of the UK longitudinal research community. It is led by the University of Bristol and University of Edinburgh, in collaboration with Swansea University, University of Leicester, University College London and the contributing LPS.

The governance structure of the UK LLC is shown in figure 1. Membership of the various groups and committees includes key external organisations, including the contributing LPS, LPS participants and

members of the public to advise on progress, scientific outputs, communication materials and long-term strategy of the UK LLC. The role of the UK LLC Data Access Committee is explained in section 7.

Contributing LPS input into the design of the UK LLC and monitor its operations through the Vanguard group of study principal investigators (PIs) and data managers. LPS data managers also form part of the UK LLC Data Access Committee and each study considers proposals for their data to be used in each research project.

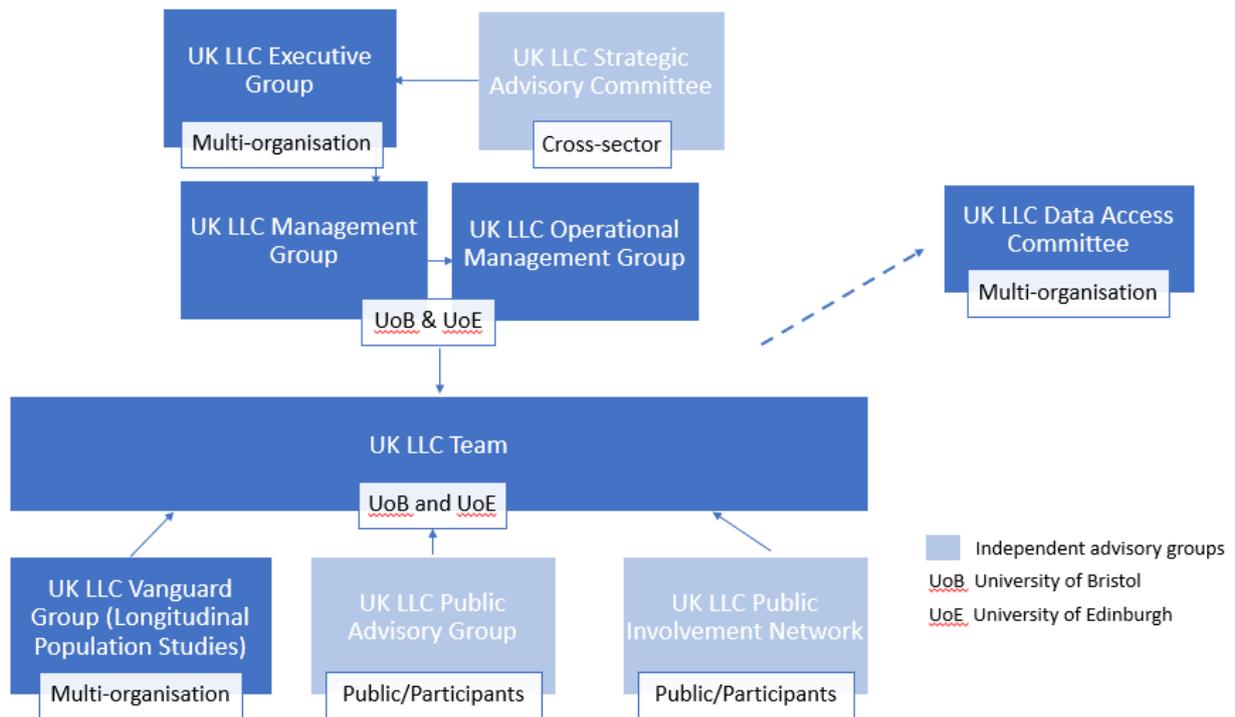
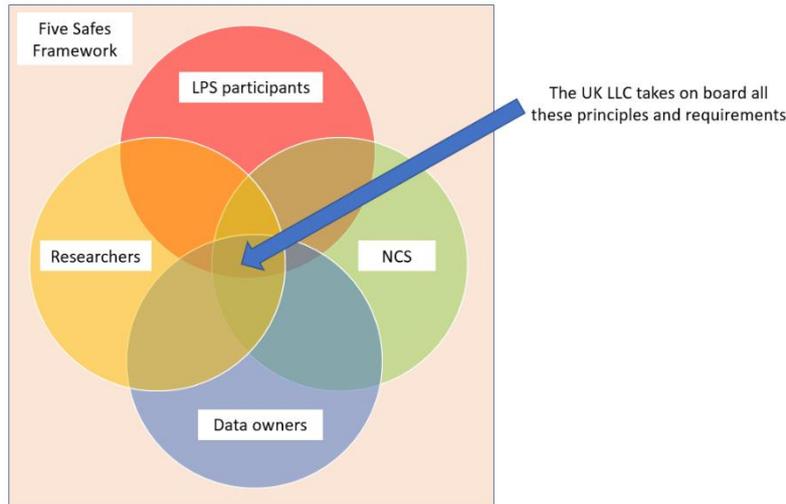


Figure 1 The governance structure of the UK LLC, including membership of each committee or group

### 3 PRINCIPLES

The UK LLC is committed to protecting LPS participants’ data and to supporting researchers to access de-personalised LPS data linked with health and other records, in an efficient, safe and consistent manner to conduct C-19 research that delivers public benefit. Our principles are based on those established between LPS and participants across the longitudinal community and the specific requirements of the NCS programme to work in an open and reproducible manner, making best use of existing infrastructure resources and to involve the public/LPS participants in our design and operations (figure 2).



**Figure 2** Commitment to LPS participants, the data owners, the wider NCS programme and researchers, and the adoption of the Five Safes Framework, form the basis of the UK LLC's modus operandi

### 3.1 Five Safes

Applications to access data in the UK LLC TRE are assessed against the ['Five Safes' framework](#) to determine whether they meet the requirements and commitments agreed with participants of LPS, the data owners and the NCS. The Five Safes are: (i) Safe data; (ii) Safe projects; (iii) Safe people; (iv) Safe settings; and (v) Safe outputs. Section 5.1 provides further detail.

### 3.2 Commitment to LPS participants

Those leading, designing and managing the LHW NCS and its contributing LPS, and the UK LLC, have made a series of commitments to LPS participants as to how their data will be handled and used (Table 1)<sup>1</sup>. Applications to access the UK LLC TRE must meet these conditions. Some LPS have additional specific commitments and agreed ways of working with their participants. The UK LLC manages a review process that enables each LPS to decide whether their data can be included in each application. This may mean a successful application is not provided with data from each contributing LPS.

<sup>1</sup>See: <https://www.ucl.ac.uk/covid-19-longitudinal-health-wellbeing/statement-our-study-participants-about-use-your-data-and-nhs-data-research>

**Table 1:** UK LLC commitments made to participants of contributing LPS

<p><b>Our commitment to you:</b></p> <ul style="list-style-type: none"><li>• We never use personal identifiers such as your name or address in any UK LLC research.</li><li>• We only use personal identifiers such as your NHS number to link the information you give us to your health and other records.</li><li>• This de-personalised data is used solely for research in a secure, confidential space, called a 'Trusted Research Environment (TRE)'. Results cannot leave the environment until an independent checker has confirmed that no individual can be identified.</li><li>• Approved researchers can only access the TRE once they and their research questions have been checked and approved by a rigorous three-stage review process.</li><li>• No data are shared solely for profit-making purposes. We do not sell your data and we never will. Any researcher using your data signs up to this commitment.</li><li>• As a community of studies and data scientists we manage your data ourselves. We do not outsource to private companies.</li><li>• Our ways of working are reviewed by independent ethics committees and volunteer study members.</li></ul>
---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------

### 3.3 Commitment to data owners

As the 'onward sharers' of data, the UK LLC is both legally obliged and morally committed to uphold a number of commitments made to data owners that are documented in the data deposit agreements/data sharing agreements entered into with each contributing data source. The UK LLC commits to:

- Hold data in the UK LLC TRE and ensure access to data is restricted to approved users for approved purposes.
- Ensure the technical and organisational measures are in place to maintain the confidentiality and integrity of the data in the TRE, and protect against unauthorised use.
- Ensure that relevant governance and security terms and conditions set by data owners in our data deposit agreements/data sharing agreements are also applied to research users.

### 3.4 NCS data sharing principles

The NCS have committed to a set of principles to enable delivery and share data in a way that builds public trust:

- Work collaboratively to actively share data to allow the scientific community to pool expertise, draw fresh insights, and increase our collective understanding.
- Demonstrate active and ongoing engagement with patients and the public in the design, development and governance of their activities, to provide assurance that these activities are in the public interest.
- Be transparent in the use of personal data and respect the privacy and confidentiality of individuals, complying with legal requirements and ethical expectations at all times.
- Make research outputs, observations, code and tools generated from the studies open-source, rapidly and freely accessible as a public good.

**PUBLIC**

---

- Ensure all data and associated code and tools generated through the studies are Findable, Accessible, Interoperable and Reusable (FAIR).
- Demonstrate value for money by using existing UK infrastructure and research investments as far as possible and using open competitions where necessary to develop new infrastructure capability.

**3.5 Commitment to researchers**

The UK LLC is committed to:

- Processing applications in a fair, transparent, consistent and timely way.
- Maintaining clear communication channels with researchers throughout the application process.
- Maintaining the integrity of the data and to evaluate this and document our findings.
- Providing clear documentation and guidance to researchers and providing advice when questions arise.
- Working alongside researchers to facilitate team data science for the aim of public benefit research into the impact of C-19 and the recovery of the UK population from the pandemic.

To facilitate efficiency and effectiveness of the resource, and to uphold the transparent reproducible research principles of the NCS, the UK LLC intends to build a library of core 'research tool' products generated by UK LLC data managers and as the products of each UK LLC project. This will include tools such as protocols, syntax, code lists and documentation. These will be made available in the public domain.

The UK LLC is specifically designed to facilitate research using linked longitudinal records in alignment with whole/devolved population electronic health record sources (e.g. OpenSAFELY, NHS Digital TRE, SAIL Databank, EAVEII) with LPS data. To support this, the UK LLC will extend the research tools library to include code lists and potentially data structures found in other TRE settings. This will allow analysts to efficiently move from one whole population environment to conduct 'deep dive' analysis within linked LPS data. This process is facilitated by the open sharing and exchange of research products across the NCS.

**4 DATA TYPES AND DISCOVERY**

The UK LLC allows approved users access to the source data in the TRE to create anonymous outputs for the purpose of the approved research. The source data comprises: (i) LPS data; (ii) Linked data; (iii) Synthetic data; (iv) Derived data; and (v) Metadata relating to the source data.

The UK LLC uses all LPS data and linked data under licence. Neither the recipient institute nor the researcher gain ownership of these data and will need to comply with any specific requirements or conditions imposed by the original owner of the data.

**4.1 LPS data**

LPS data are collected by the contributing LPS, directly from, or about, LPS participants and are deposited in the UK LLC TRE by the contributing LPS. All LPS data are provided as de-identified coherent 'logical blocks' of data. These blocks can relate to a data collection assessment (e.g. a 'wave' of data collected by fieldworkers or through a questionnaire) or a set of measures collected over different timepoints (e.g. mental health assessments collected longitudinally across questionnaires). Where possible, these logical blocks are designed to match the existing datasets used by LPS to share data directly or that are used in other repositories (e.g. UK Data Service).

The variable names and labels match the LPS' existing format. The UK LLC is not seeking to redevelop metadata resources and signposts applicants to the relevant LPS resources and existing catalogues (e.g. CLOSER Discovery) via the UK LLC website and the UK LLC Data Request Form. All LPS datasets are provided to the user with the same structure as defined by the contributing LPS.

## 4.2 Linked data

The UK LLC systematically links to participants' health and geo-coded environmental data. The extent of the linkage is controlled by each contributing study based on study and participant permissions. This means that not all LPS are linked to all data sources, or that all participants are linked to their records. The linkages are determined by data sharing contract which is subject to change and is subject to any licence conditions imposed by the owner of the data within the linked datasets.

### 4.2.1 Linked health data

The UK LLC has entered into a data sharing agreement and contract with NHS Digital, permitting the sharing of linked English NHS records. The UK LLC is negotiating access to linked health records from the NHS in Scotland, Wales and Northern Ireland; more information will be provided in due course.

The data sharing agreement and contract with NHS Digital enables the UK LLC to host, process, curate and onwardly share (within the confines of the UK LLC TRE) linked health records. Access to these records is only available for applicants requiring linked LPS data (i.e. for projects only requiring NHS records, please contact the NHS data services directly). The UK LLC is required to onwardly assert all conditions and contractual terms imposed by NHS Digital (see Appendix 1). Applications are assessed against the NHS Digital conditions by the UK LLC Linked Data Review Panel. The UK LLC Data Access Agreement (DAA) which must be signed by each institute hosting approved users prior to them gaining access to the UK LLC TRE, refers to the NHS Digital terms and conditions.

The UK LLC has a responsibility to minimise NHS Digital data shared through the UK LLC TRE. Not all variables are included: the UK LLC only hosts C-19 relevant data; direct identifiers are excluded; and potentially identifiable information (such as practice codes or geographical IDs) are masked prior to deposit in the UK LLC. The UK LLC does not have access to the information needed to reverse this masking.

The UK LLC respects the rights of the public to opt-out of the use of their data for research purposes, and the [NHS National Opt Out](#) (England and Wales only) is respected unless specific consent is in place.

### 4.2.2 Linked administrative data

The UK LLC is in the process of negotiating access to linked administrative records from key UK government departments including the Department of Work and Pensions. More information will be provided in due course.

### 4.2.3 Linked geographical and multi-level modelling data

The UK LLC is a de-identified resource where inclusion of real-world geographical indicators are kept to an absolute minimum. The only real-world indicators available are LPS participants' UK region and devolved nation of residence, and service provision. These data are essential for reporting purposes and will be inherently identifiable in the case of devolved regions, because of the different sources of linked data. All other levels of geographical indicator are encrypted and neither the UK LLC nor the approved user have access to the encryption key. Encrypted identifiers will be made available where required for multi-level modelling or other geospatial purposes.

***Linked geospatial records and GIS approaches***

The UK LLC has developed a mechanism to geocode LPS participants' addresses using third parties and masking address data to counter confidentiality concerns. Only some contributing LPS participate in this feature of the UK LLC and based on LPS-level governance controls some LPS permit geocoding to property-centroid and some to postcode-centroid.

The UK LLC Data Team (University of Bristol) use this mechanism to add commonly used natural environment and built environment indicators to the UK LLC resource (e.g. deciles of indices of multiple deprivation, urban/rural indicator). It is possible to add additional geospatial indicators to the UK LLC resource to meet project needs and to enable GIS approaches (subject to disclosure risk checks and processing). Please get in touch to discuss options.

***Modelled geospatial exposure data***

The UK LLC has commissioned the University of Leicester's Centre for Environmental Health and Sustainability to model a range of environmental exposure estimates relating to air pollution, noise and green space.

**4.3 Synthetic data (not yet applicable in the UK LLC)**

Synthetic data are representations of the source data that are artificially generated rather than relating to real individuals. Synthetic data are typically created using statistical techniques, so that while the data are not 'real', they may retain a similar underlying structure and distribution. The University of Bristol retains ownership of the synthetic data and, for the avoidance of confusion or mistakes, the distribution and handling of synthetic data must still use rigorous and secure processes.

**4.4 Derived data**

Derived data are individual-level data generated from analyses or functional combinations of the UK LLC source data. An example of a derived data variable is body mass index (BMI), it being a function of height and weight. The UK LLC produces some derived data products for commonly used information and these will be catalogued and made available for request. Approved users must report any data they derive during their research projects and provide the syntax and documentation for these so they can be added to the UK LLC resource (by the UK LLC Data Team). The UK LLC manages the source data of these derived data and the information used to produce the derived variable remains the property of the original owner.

**4.5 Harmonised data**

The UK LLC will produce and maintain a harmonised set of core socio-economic status (SES) and demographic indicators. It is not the intention of the UK LLC to conduct any further harmonisation of data. However, we welcome applications that involve harmonisation approaches and all documented harmonised datasets will be added to the UK LLC resource as derived data once completed. Insights from the production of the core SES and demographic indicators will be made available to help future users.

**4.6 Metadata**

Metadata is information that describes the LPS data, linked data and/or derived data, and the way in which the data were generated. Metadata include variable names and labels, data collection protocols, information about the tools and techniques used to collect the data, the staff involved, the time and place of data collection, details of quality control and error – i.e. the 'provenance' of the data, which allow future researchers to understand and accurately use these data in a secondary context. The contributing LPS and external third parties provide metadata to support approved users

and prospective users and the public. The UK LLC generates additional metadata when processing LPS data and linked data and will make these accessible in a clear and understandable format.

#### 4.7 Data quality, coverage and completeness

A summary of UK LLC data quality, coverage and completeness:

- Most data in the UK LLC TRE are accessed and used under licence and are a secondary reuse of datasets owned and/or controlled by other organisations. This means that some data may be removed with limited notice and that some variables may change over time.
- The UK LLC cannot guarantee the quality of the underlying data but will provide assessments of completeness and coverage and will make assessments of linkage quality.
- The denominator of the UK LLC TRE (i.e. the sum total of participants across contributing LPS) will change over time. This will be recorded in UK LLC documentation and it will be clear which 'denominator version' is in place for each project's dataset.
- Linked records are refreshed on a regular basis. Each refresh may be based on a different underlying denominator where new LPS and/or participants are added and some participants may be removed.
- Users are provided with a snapshot of data current at the time of project approval; updates (due to data refreshes or amendments) may not include data for all individuals.

#### 4.8 Data discovery

The UK LLC is designed to be a discoverable and accessible resource for C-19 research. A summary catalogue of the data held in the UK LLC is available at the [HDR UK Innovation Gateway](#). An overview of available data can be found on the [UK LLC website](#).

## 5 ELIGIBILITY AND REQUIREMENTS TO ACCESS DATA

To ensure the appropriate and fair use of LPS participants' information, the UK LLC operates a rigorous three-stage application assessment process (see section 7) that considers the eligibility of applicants and their projects, as well as the requirements of the contributing LPS and owners of linked data.

### 5.1 Eligibility

The UK LLC welcomes applications from any **UK-based bona fide researcher** (including PhD students, on condition that appropriate supervision is in place from an applicant who is a senior researcher) who is conducting a **C-19 focused research project** for public benefit (and not for commercial gain). A **research project** is defined as a specific research activity addressing a pre-defined purpose. A project may consist of activities investigating a range of hypotheses, but these must all be specified and related to the project theme. The purpose of a project may be to conduct applied research, methodological research or resource enhancements (e.g. the creation of a harmonised dataset or outcome adjudication).

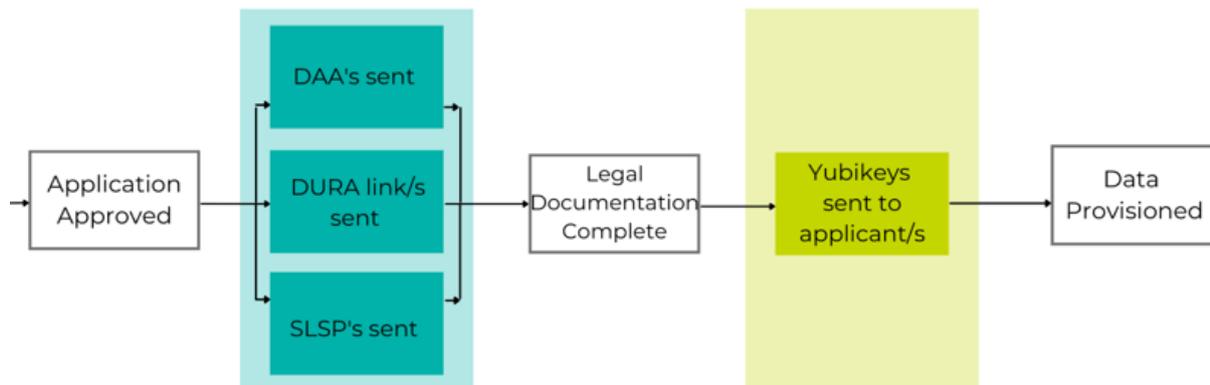
All applicants and their applications are assessed against the '[Five Safes](#)' Framework:

- 1) **Safe people:** we will determine if the **applicants** (those applying to access the UK LLC TRE, not their wider team) are trained and authorised to use the data safely. It is important that all applicants have the technical skills necessary to use the data effectively and have passed the ONS Safe Researcher Training course.

- All applicants must provide a short CV demonstrating their ability to conduct the proposed research.
  - All applicants must have [ONS Approved Researcher Accreditation](#).
  - All applicants must be part of a research organisation.
  - All applicants must sign a UK LLC Data User Responsibility Agreement (DURA).
- 2) **Safe projects:** applicants need to demonstrate that their proposed research and data request is appropriate, ethical and likely to deliver clear public benefit that will be disseminated.
- The scope of all projects must be C-19 research, either regarding the virus itself and its outcomes, efforts designed to ameliorate the virus, or the wider effects of the pandemic and (all aspects) on health and social wellbeing.
  - Applicants must provide relevant information about their protocols, including information about the purpose of their research, the data and methods to be used and the intended outputs and means for effective dissemination.
  - All applicants must commit to publish their workings (syntax, code lists, documentation, protocol) and findings into the public domain as quickly as possible.
  - For all approved projects, the institute(s) that employs the applicant(s) must sign a Data Access Agreement (DAA) with the University of Bristol (the owner of the UK LLC).
  - The UK LLC TRE hosts data under licence, where conditions are defined in data deposit agreements or data sharing agreements. All applications are assessed against these conditions either by the UK LLC (where this permission has been delegated) or by the data owner. Successful applicants are bound by the conditions detailed in this policy, in the DAA and in the DURA.
- 3) **Safe data:** all organisations providing data to the UK LLC must ensure their data are de-identified. They achieve this by removing names, addresses and any other details that would directly identify a participant. The UK LLC performs a series of checks throughout the data processing pipeline to ensure confidentiality. Only the minimum required data are provided to approved applicants.
- 4) **Safe setting:** the UK LLC operates a TRE model using infrastructure provided by SeRP UK (Swansea University). All data in the UK LLC TRE are sensitive. To maintain the confidentiality and security of the LPS participants' information, data are only made available *within* the UK LLC TRE. The TRE is a locked down environment for analysis which hosts statistical packages and other software needed for research.
- 5) **Safe outputs:** all outputs (e.g. tables, statistical results) must be population level and not individual level data. All outputs are checked by trained staff prior to release. Once approved, the results are sent to the researcher, who can then publish or present their analyses.

## 5.2 Application governance (legal documentation)

The UK LLC hosts sensitive de-identified data from many contributing data owners. Whilst all use of the data occurs within the UK LLC TRE, there is an expectation from data owners that those accessing the TRE will maintain adequate levels of security whilst doing so. Once an application is approved, there are several governance steps that must be completed prior to data being accessible to approved users in the UK LLC TRE (figure3).



**Figure 3** Once an application is approved, application governance steps (legal documentation) must be completed to ensure LPS participant data are protected – only then are data provisioned to approved users in the UK LLC TRE via use of a YubiKey

**5.2.1 Data Access Agreement (DAA)**

The UK LLC DAA defines the terms and conditions for accessing and using data within the UK LLC. These terms and conditions are non-negotiable and signing a DAA is a pre-condition to gaining access to the UK LLC TRE.

For each approved application, there must be a DAA in place with every institute involved, i.e. the institutes employing the approved users of the UK LLC resource (individuals involved in the research who will not access the data do not need to be covered by a DAA). Where an individual is employed by multiple institutes, the DAA needs to be in place with the institute sponsoring the research and taking accountability for it.

**5.2.2 Data User Responsibilities Agreement (DURA)**

Each approved user of data in the UK LLC resource must complete a DURA for each application they are named on. The purpose of this document is to make clear to each user the responsibilities they hold. The terms and conditions of the DURA are non-negotiable and signing a DURA is a pre-condition to gaining access to the UK LLC TRE.

**5.2.3 System Level Security Policy (SLSP)**

Approved users need to demonstrate sufficient levels of security policy and practice to assure the UK LLC that they are compliant with expectations. This can be evidenced through providing proof of ISO 27001 Information Security Standard or NHS Digital Data Security and Protection Toolkit (DSPT) accreditation (where the approved users’ activity is within scope of the accreditation). Where this is not possible, each institute must evidence this through completion of the UK LLC System Level Security Policy (SLSP) form.

The scope of the SLSP is limited to only the equipment used by the approved users, the setting in which they work and the network they use. It does not extend to the UK LLC TRE itself or their wider institute. Typical contemporary University IT or Departmental Policy and Practice with adequate user training and awareness is likely to meet the requirements of the UK LLC SLSP (e.g. conforming to the Cyber Essentials scheme).

**The UK LLC provide a template SLSP document for applicants to provide guidance and make sure this is an efficient process.**

### 5.3 Requirements

The following requirements need to be taken into account when applying to access the UK LLC TRE.

### 5.4 Intellectual property

The University of Bristol owns the UK LLC and is Data Controller for the resource, managing the rights to, or owning itself, all rights to the **source data** (see section 4). The majority of data held within the UK LLC are accessed under DAAs and typically the owners of the data retain all intellectual property rights – **no user will gain any intellectual property over data or participant-level research outputs (e.g. derived variables) as a result of their project (given the data are accessed under licence, this is a non-negotiable feature of using the UK LLC resource).**

Approved users are granted a non-exclusive, revocable, non-transferable, licence to use the source data for their approved project, until the end of their approved access period, so long as the University of Bristol maintains a relevant data sharing agreement with the third-party data owner, and retains the rights to sub-license the source data.

### 5.5 Transparency and a community of researchers

The NCS are designed to maximise the benefits of transparent ‘Team Data Science’<sup>2</sup>. Across the NCS – and in the UK LLC – it is required that:

- Publicly accessible registers of all UK LLC approved users, projects and outcomes will be maintained to ensure participants of contributing LPS and the wider public are aware of how data within the UK LLC TRE are being used. A record of all applications and their outcomes will be kept in the public domain (in the [Data Use Register](#)) and fed back to contributing data owners (LPS and organisations providing data).
- UK LLC approved users will be required to make their project workings visible and sharable via creative commons licence (or similar). The UK LLC will maintain a [UK LLC Git Hub repository](#) where, for each project, users will make their syntax, code lists, protocols, methods and derived variables accessible and understandable to future users. This will take the form of an internal Git repository within the UK LLC TRE for active projects and an external Git repository for completed projects.
- Users are strongly encouraged to make use of research notebooks such as Jupyter notebook or R Markdown to document their syntax and workings (both systems are provided within the UK LLC TRE).
- Any derived data, along with relevant and adequate documentation, must be made available to the UK LLC Data Team for incorporation into the UK LLC resource. Failure to make derived data available will risk future access to the resource.

### 5.6 LPS requirements

Each contributing LPS sets out conditions in the data deposit agreement regarding data use in the UK LLC TRE. Many LPS reserve the right for their data access committees (DACs) to assess applications against their study-specific application review frameworks. Requirements for LPS acknowledgements are set out in the UK LLC Publication Checklist.

---

<sup>2</sup>Ford E, Boyd A, Bowles JK, Havard A, Aldridge RW, Curcin V, Greiver M, Harron K, Katikireddi V, Rodgers SE, Sperrin M. Our data, our society, our health: A vision for inclusive and transparent health data science in the United Kingdom and beyond. Learning health systems. 2019 Jul;3(3):e10191.

## 5.7 NHS Digital requirements

The linked health records in England are accessed under a data sharing agreement and data sharing framework contract with NHS Digital. These impose conditions as to how the UK LLC may use NHS data and, through our application review and governance process, we 'onwardly share' these conditions to users.

To use NHS Digital data users must demonstrate:

- How their proposed research clearly articulates 'benefits to healthcare provision, adult social care and/or the promotion of health' (this is a legal requirement and condition of access).
- A realistic and comprehensive plan for how their findings will be disseminated to relevant stakeholders to achieve the stated benefits.
- That the use of NHS data is relevant to the stated purpose. Within this, the applicant must justify how their application requires NHS data linked to LPS data – and why their purpose cannot be satisfied through using either source alone.
- Each and all of the requested data are relevant to the stated purpose, i.e. a justification at an item level, where item relates to the broad datasets requested – e.g. primary care records, mortality register records – or to a range of data defined by a code list (e.g. a list of respiratory health codes).

**Applicants must clearly and specifically answer each of these points. Applications which do not achieve this cannot be processed.**

The full list of NHS Digital requirements can be found in **Appendix 1**.

## 5.8 Penalties for misuse

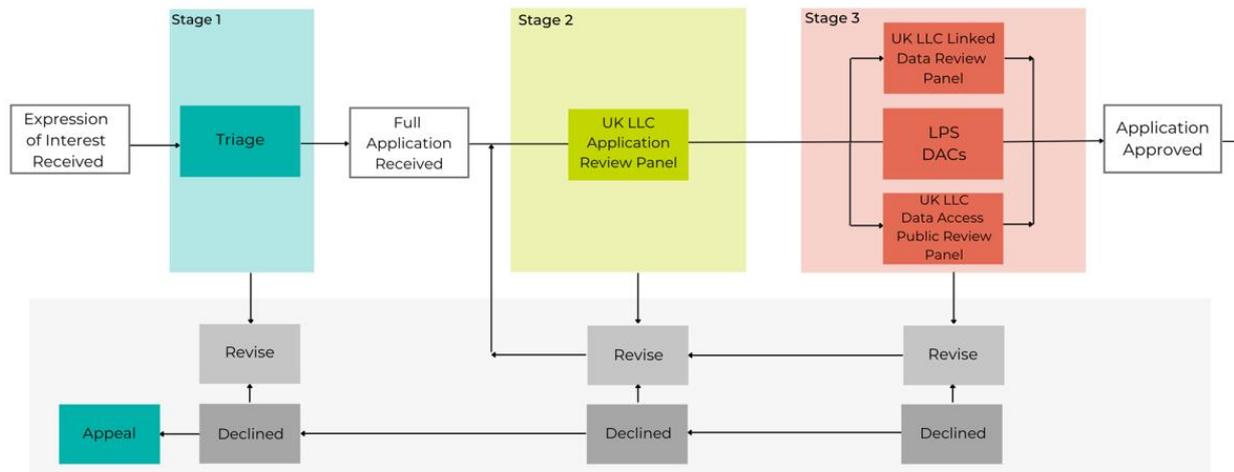
Users must adhere to the terms of the DAA and DURA at all times. The DAA is a legally binding contract. Current and future access is at risk for you, your team and your institute if you do not abide by the terms of these agreements. Serious breaches of these terms may be subject to prosecution to the full extent of civic or criminal law.

## 6 CHARGES TO ACCESS DATA

The UK LLC is currently funded through the COVID-19 LHW NCS programme to support users to conduct C-19 research. At present, there is no charge for data access and provision. The UK LLC is in discussion with the research funders around the future of the resource and funding options.

## 7 MAKING AN APPLICATION

The process to apply to access the UK LLC TRE is comprised of three stages (figure 4). Applicants with multiple projects must submit multiple forms; one per project.



**Figure 4** A summary of the three-stage process to apply to access data in the UK LLC TRE. Once an application is approved, the applicant completes paperwork (see section 5.2) and then moves to the data provisioning stage of the process (see section 8)

### 7.1 Stage 1

Expressions of Interest (EoIs) must be submitted through the UK LLC enquiry form on the [HDR UK Innovation Gateway](#). The EoIs are triaged by the UK LLC Applications Team (University of Edinburgh). Incomplete or insufficiently justified EoIs are returned to applicants, who have the opportunity to revise and resubmit. Applicants whose EoIs are approved by the UK LLC Applications Team are invited to submit a full application and data request form.

### 7.2 Stage 2

An initial assessment of full applications and data request forms is conducted in stage 2 by the internal UK LLC Application Review Panel. This review assesses each application against the Five Safes Framework detailed in section 5.1 and checks that all requested fields are filled out appropriately.

In particular, applications are reviewed to ensure that the data being requested are appropriate for the research question, that public involvement plans have been detailed and a lay summary provided.

Any questions or concerns are sent back to the applicant(s) and resubmitted applications are re-reviewed by the UK LLC Application Review Panel. Applications that are declined are permitted to go through the appeal process outlined in section 7.4.

### 7.3 Stage 3

Approved applications proceed to stage 3. The majority of applications are from researchers wishing to access LPS data linked with health (or other) data. However, a small number of researchers apply to access only LPS data. The review process at stage 3 differs slightly for LPS only data, versus LPS and linked data.

Applications for **LPS data alone** are reviewed by:

- 1) The individual **LPS DACs** – each LPS DAC is responsible for reviewing the application against their study-specific application review framework, unless they have formally deferred decision making to the UK LLC. The latter LPS receive regular updates from the UK LLC. The UK LLC

**PUBLIC**

---

collates and records the LPS' responses and shares this information, plus any stated conditions, with the applicant.

- 2) The **UK LLC Data Access Public Review Panel** – the panel focuses on the lay summary and public involvement sections of the application. Applicants have a five minute slot to present their research to the panel at an online meeting, with a short Q&A session thereafter.

Applications for **LPS data linked with health (or other) data** are reviewed by:

- 1) The individual **LPS DACs** (see above).
- 2) The UK LLC Data Access Public Review Panel (see above).
- 3) The **UK LLC Linked Data Review Panel** – the panel focuses on reviewing the application against the linked data owners' requirements (see Appendix 1).

Applicants must minimise their data request to only include data that are relevant to their project. They must specifically justify the inclusion of data of 'elevated sensitivity' such as socially sensitive data (e.g. mental and sexual health data) and demographically sensitive data (e.g. ethnicity).

Approved applications are passed to the UK LLC Data Team (University of Bristol) who work with successful applicants to complete the necessary legal documentation detailed in section 5.2 prior to despatch of YubiKeys and data provisioning.

#### **7.4 Appeal process**

Declined applications may go through the appeal process. Applicants must address all comments, questions and criticisms fully for the appeal to be considered, with all changes to the application highlighted. Amended applications are then submitted for stage 2 approval, and, if successful, progress to stage 3 approval.

#### **7.5 Amendment process**

Amendments to approved applications are considered by the UK LLC Application Review Panel. The DAA contains provisions for most project amendments to be made without needing to amend the DAA. However, amendments will need to be confirmed by the study PI before the amendment can take effect.

#### **7.6 Application timescales**

For the parts of the process under UK LLC control we set the following timescales (see figure 4):

- **Stage 1.** We aim to review Eols within one week of receipt.

Following receipt of a completed application form and data request form:

- **Stage 2.** We aim to review applications within two weeks.
- **Stage 3.** We aim to review applications within four weeks.

On completion of required legal documentation (DAA and DURA, and SLSP where needed):

- We aim to provision data within two weeks.

**When seeking funding** please enquire at the earliest opportunity ([access@ukllc.ac.uk](mailto:access@ukllc.ac.uk)). You must make clear that any access to data in the UK LLC TRE is subject to approval.

## **8 WORKING IN THE UK LLC TRE**

An application is referred to as a project at the point data are prepared for provisioning within the TRE.

### 8.1 Two factor authentication

Once the legal documentation is completed, each approved user will be sent a YubiKey two factor login authentication device to protect access to the TRE.

### 8.2 Permitted access period

The permitted access period runs from the date the DAA is signed until the end date stated on the application. If approved users wish to extend their access period, they can submit an amendment (see section 7.5) and, if approved, the access period will be adjusted accordingly.

All users in the TRE must have **current ONS Approved Researcher Accreditation**. If accreditation expires before the end of the agreed access period, access to the TRE for that user will be automatically terminated. It is the responsibility of the user to maintain their ONS Approved Researcher Accreditation and to communicate successful renewals to the UK LLC in good time ([access@ukllc.ac.uk](mailto:access@ukllc.ac.uk)), so that their access to the TRE is not interrupted.

### 8.3 Project names across the project lifecycle

Each approved project will be allocated a unique project reference number. It is important for public/participant transparency, the reusability of content and governance compliance that these project numbers are used in a clear and consistent way across the project lifecycle. **All research users are required to follow the UK LLC naming convention for labelling research products (e.g. syntax and code lists); this will help us archive these and make them available for wider reuse.**

### 8.4 Project folder

Each project will be allocated a 'project folder' within the UK LLC TRE (named according to the project number e.g. llc\_1234). This is a secure folder which can only be accessed by approved project users, certain UK LLC staff and the TRE hosts (SeRP UK). The project folder will be used to store all workings relating to the project.

### 8.5 Project Participant IDs

Each project will also be allocated a unique set of project participant ID numbers. These IDs must not be shared with users outside the project. If a PI has more than one project, separate identifiers will be attached to each set of datasets related to each project. **Researchers must not attempt to join data from different projects.**

### 8.6 Project Git

Each project will be allocated a Git repository (repo) *within* the UK LLC TRE. The project users will be required to add cleaned and documented content to the repo at the point at which data are exported from the UK LLC for publication. The UK LLC resource curators will review this content, request changes where necessary, and copy the non-disclosive materials over to the *external* UK LLC GitHub for public access. **All research users are required to follow the UK LLC naming convention.**

### 8.7 Project DOI

The UK LLC is investigating the use of digital object identifiers (DOIs) for the citation and archiving of project materials and other aspects of UK LLC functioning. Users may in future be required to use set DOIs to cite the UK LLC resource or the research tools they deposit in the UK LLC repository.

## 8.8 User support

The UK LLC Data Team maintain a library of self-service user help tools in multiple formats, including user guides, helper syntax, FAQs and accompanying videos. We aim to help researchers begin their analyses in the TRE as quickly as possible.

If users are unable to find the help they require, the UK LLC Data Team can be contacted at [access@ukllc.ac.uk](mailto:access@ukllc.ac.uk). If users experience issues with the software or the SeRP environment they should contact [helpdesk@chi.swan.ac.uk](mailto:helpdesk@chi.swan.ac.uk) in the first instance. Both mailboxes have a ticketing system to track user requests.

## 8.9 Audit

All users must agree to being audited and will provide full cooperation to internal UK LLC audit processes or audits being conducted by external organisations. All auditors commit to the confidentiality of the research project.

# 9 OUTPUTS

## 9.1 Derived data

Derived data must be made available to the UK LLC data managers with full and clear documentation (including syntax) by the project access end date. **It is not possible for derived data to leave the UK LLC TRE**; although it is legitimate for researchers to claim derivations as an output which are then discoverable and accessible to future users. Derivation syntax will be processed through the 'data out' (see section 9.3) checks and made available via the external UK LLC git repository.

## 9.2 Manipulated data

If the derived data come from linked data that were sourced from NHS Digital, then NHS Digital retain the right to determine whether derived data are derived data or manipulated data. Manipulated data are defined as information created by the approved users using NHS Digital source data that can be identified as originally from NHS Digital and must be returned by the project access end date.

## 9.3 'Data Out' requirements and responsibilities

It is a critical principle of the UK LLC TRE that only anonymous population level aggregate data (e.g. tables of results) can be released from the TRE. It is not permitted to release any individual level data from the TRE. All outputs must be assessed for disclosure risk and transformed to control for identified risks prior to being permitted to leave the TRE. **All users must agree to uphold these principles.**

- Data that are destined for publication should be converted to anonymous aggregate form by the project users.
- All users must use their knowledge from the ONS Approved Researcher training (and other relevant training) to make best efforts to ensure outputs are anonymous before submitting them for 'data out' checks.
- Once submitted to the 'Data Out' process, outputs will be reviewed by trained individuals.
- The findings of the review are binding and research users must correct any identified issues.
- Users must minimise the number of items they submit for data out (there should ideally be two data outputs per publication – one for submission and one to respond to reviewers' comments).

- Failure to adhere to these conditions may result in sanctions (such as users being required to undergo renewed training).

## 10 PUBLISHED OUTPUTS

### 10.1 Authorship and acknowledgements

The UK LLC principles of authorship and acknowledgements are detailed in the UK LLC Publication Checklist. All publications must properly acknowledge the contribution of the collaborating LPS and wider data owners whose data forms the UK LLC resource.

Contributions should be acknowledged using [icmje.org](http://icmje.org) authorship guidelines.

‘UK LLC’ must be included as a key word in all publications.

UK LLC staff are not required to be authors unless they have provided material input into the research and where authorship is justified under [icmje.org](http://icmje.org) guidelines.

### 10.2 Open Access publication policy

All papers published using the UK LLC resource must follow [UKRI’s policy on Open Access](#).

### 10.3 Research papers

The UK LLC requires that publications are placed into the public domain as quickly as possible. All pre-prints/papers/reports being submitted for peer-review or placed into the public domain must be sent to the UK LLC ([access@ukllc.ac.uk](mailto:access@ukllc.ac.uk)), prior to submission, for licence condition and disclosure risk checks.

Data must not be published as part of a study output that could identify an individual or could be used to identify an individual. Users will proactively bring to the attention of the UK LLC any outputs that they consider may identify an individual or individuals. The UK LLC licence checks include ensuring that the UK LLC and contributing LPS and data sources are accurately described and acknowledged. The UK LLC management may withhold the right to publish any paper that threatens the confidentiality of participants of contributing LPS or threatens to bring the UK LLC into disrepute.

UK LLC staff do not conduct peer-review or scientific assessment, although they may offer optional guidance to the authors. Under all circumstances the UK LLC management reserves the right to submit letters or papers for publications in response to any paper to explain study procedures or to express a coherent scientific argument.

**Important:** All UK LLC users must ensure that their publication is compliant with the UK LLC Publication Checklist, which details specific publishing requirements.

The UK LLC should be informed of all publications arising from use of the UK LLC resource ([access@ukllc.ac.uk](mailto:access@ukllc.ac.uk)).

#### 10.3.1 Associated datasets

Some journals request that the datasets associated with any publication are themselves published in public databanks. The UK LLC does not permit this given that the underlying data are sourced from confidential LPS data and linked health and administrative records. Where this is requested from a journal, we suggest that you provide the following statement:

‘Data used in this research are made available via the UK Longitudinal Linkage Collaboration (UK LLC), which is a UK Trusted Research Environment for the use of longitudinal study data with linked routine records. These data cannot be used or shared outside this environment. Researchers can apply to use

the UK LLC using the procedure outlined in the UK LLC Data Access and Acceptable Use Policy (<https://ukllc.ac.uk/>). The UK LLC uses a system of managed open access where researchers can demonstrate their project is intended to improve the public good.'

#### 10.4 Presentations and other publications

The UK LLC does not require to see presentations before they are delivered, but requests that users must only include anonymous data that has been approved through the 'data out' process (see section 9.3). **It is not permitted for users to make any UK LLC held data available in the public domain through presentations or any other form unless it has been through 'data out'**. Presentations must acknowledge the use of the UK LLC resource and its funding/LHW NCS (see the UK LLC Publication Checklist).

#### 10.5 Media communications

Where possible, press releases and other media communications should be developed in conjunction with the UK LLC Communications Team. We reserve the right to publish press releases on certain articles and expect the lead author of the article to agree to the press release with the Communications Team and to be available to deal with media enquiries and interviews. We may also ask authors to prepare a précis of important papers and/or lay summaries to include in reports to funders and future applications for core funding.

### 11 CONFIDENTIALITY AND DATA SECURITY

Users must adhere to the requirements listed below. These include, but are not limited to rules regarding the maintenance of confidentiality, IT equipment and reporting breaches or risks.

#### 11.1 User security requirements

- All UK LLC users must keep their UK LLC user access password private and their YubiKey secure and not share these with anyone else.
- All UK LLC users must comply with the data security requirements set out in the DAA, DURA and this policy.
- All UK LLC users must maintain the confidentiality of the UK LLC held data.
- All UK LLC users must report any breaches of data security – including instances of accidentally recognising the identity of any participant.

#### 11.2 IT equipment and connecting to the UK LLC

The UK LLC requires that the device you use to access and work within the UK LLC must be a work provided machine which is automatically maintained so that it is fully patched and up to date with relevant virus protection. Access must be made via a secure private or corporate network. Where using a private network for home working, the user must make all reasonable efforts to ensure this is maintained and the equipment is secure (strong password) and kept patched and up to date.

#### 11.3 Reporting breaches and risks

It is vital that all users report any breach of the UK LLC as soon as possible. Examples include, but are not limited to:

- The loss or theft of a UK LLC YubiKey.
- The loss or theft of the computer used to access the UK LLC TRE.

- Any unauthorised person gaining access to the UK LLC TRE (i.e. a person who is not approved and has not signed a DURA).
- Any sharing of login details or devices that permit access to the UK LLC TRE.

Where inadvertent disclosure has taken place, users must notify the UK LLC as soon as is practicable and in reasonable detail as to how this occurred. They must not attempt to contact the LPS participants.

**The UK LLC is committed to continuous improvement of our systems. We ask that all users notify us about risks and any weaknesses that they identify in our systems, policies and procedures.**

## 12 APPENDICES – DATA OWNERS’ SPECIAL CONDITIONS AND REQUIREMENTS

### 12.1 Appendix 1: NHS Digital

#### 12.1.1 Special conditions

Where the Recipient Institute is accessing NHS Linked Data, the following additional conditions will apply:

- 1) The University [of Bristol] reserves the right to enter any of the premises where the data are accessed at its own discretion and when required by an NHS Contributor.
- 2) The NHS Contributors shall benefit from and can enforce the terms of the Licence directly in accordance with the Contracts (Rights of Third Parties) Act 1999 (without the involvement or consent of the University) against the Recipient Institute.
- 3) The NHS Contributors shall have the right at any time and at their own discretion to perform an audit of the Recipient Institute, provided that the Recipient Institute is given seven (7) days’ notice where permissible and practicable.
- 4) The NHS Contributors shall have the right to suspend, in whole or in part, the Licence in relation to all or part of the Data where:
  - i. The agreement between the University and the NHS Contributor is suspended in whole or in part; or
  - ii. The Recipient Institute has not complied, or is not complying with any of its obligations under the Licence, until the breach is resolved to the NHS Contributor’s reasonable satisfaction.
- 5) The Recipient Institute shall have no rights in or to the NHS Linked Data other than the right to use the NHS Linked Data in accordance with the express terms of the Licence, which shall not exceed the Sub-licence permitted under the Data Sharing Agreement signed between the University and the NHS Contributor.
- 6) The Recipient Institute shall prospectively assign any and all future Intellectual Property Rights in any NHS Contributor Manipulated Data to the relevant NHS Contributor, from creation and shall execute or procure the execution of any document, or shall perform, or procure the performance of any acts as may be required to give this effect.
- 7) Where the Recipient Institute does not maintain a registration under the NHS Data Security Protection Toolkit, the Recipient Institute warrants that the Data Security Policies outlined in the Systems Level Security Policy (available upon request) provide an equivalent level of protection of any Personal Data accessed by the Recipient Institute.

- 8) The Recipient Institute shall not have the right to assign, novate, transfer, charge, dispose of or deal in any other manner with the licence, or any of its rights or beneficial interests under it, or purport to do the same, nor sub-contract any or all of its obligations under the Licence without the prior written consent of both the NHS Contributor and the University.

### 12.1.2 Requirements for onward data sharing

These requirements are used by the UK LLC Data Access Committee to assess applications at stage 3.

- 1) To consider and define criteria by which application subject matter/data content is determined to have an 'elevated sensitivity' and thus requires additional review steps.
- 2) To assess full applications requesting to use linked data provided by the NHS or other government agencies in more detail against the following criteria:
  - i. Has the applicant demonstrated how their research outputs will lead to public benefits?
    - a. Are there clearly articulated scientific and wider societal impacts of the proposed research?
    - b. Are there clearly articulated benefits to healthcare provision, adult social care or the promotion of health?
    - c. Is the applicant committing to publish the findings of the research into the free-to-access public domain?
    - d. Has the applicant provided a realistic and comprehensive plan for how their findings will be disseminated to relevant stakeholders to achieve the stated benefits?
  - ii. Is the data request proportionate?
    - a. Is the data request minimised sufficiently so that all requested data is justified by the stated purpose?
    - b. Is there a clear and specific justification for the use of data of particular sensitivity?
    - c. Is the request to use linked data justified as being necessary to realise the intended benefits of the application?
    - d. Has any request to use linked data demonstrated how the intended purpose is only achievable by using NHS or other government data linked to non-NHS data within the UK LLC (i.e. that the request could not be fulfilled by the NHS or government department directly)?
  - iii. Does the application generate any additional or unacceptable risks of disclosure of participant identity?
  - iv. Does the proposed research have the potential to bring the UK LLC or the owners of the data deposited in the UK LLC TRE into disrepute?
  - v. Does the applicant demonstrate adequate levels of data security at their institution?
    - a. Is there organisational security assurance for applicants requesting linked NHS data?
  - vi. Is there a clear legal basis for this use of UK LLC data?
    - a. Is there a clear legal basis for any processing of linked NHS data?
  - vii. Has the applicant committed that their proposed data use will not be for profit-making purposes?
  - viii. Where the applicant is a student/PhD student, has the applicant demonstrated how their use of the data will be adequately supervised?

- 3) To consider any emerging risks, issues and/or data access policy issues and to refer these to the UK LLC Executive Group and UK LLC Strategic Advisory Committee for advice.
- 4) To, where appropriate, take advice of third-party specialist knowledge to inform decision-making and the evolution of the UK LLC access request process.
- 5) To provide regular updates to the UK LLC Executive Group and UK LLC Strategic Advisory Committee, including a summary of applications made, decisions taken, performance metrics and any arising risks, issues or landscape changes that the members of the UK LLC DAC wish to report.
- 6) For each onward sharing of the linked NHS Digital data, to publish information about the dissemination on the NHS Digital release register. This should include the name of the organisation(s) accessing the data, the purpose (summary of the project) and details of the data released.
- 7) To provide regular updates to the contributing studies summarising the applications made, decisions taken, performance metrics and an overview of projects and project outcomes.
- 8) To publish information into the public domain about applications made, the decisions relating to these and the name of the organisation(s) accessing the data, the purpose (summary of the project) and details of the data released.
- 9) To ensure a clear and reproducible process for decision making is established and made clear to members, applicants, key stakeholders and the public through a 'Standard Operating Procedure' and through set Service Standards (quantifiable metrics for service delivery).