CLS Data Access Committee

Terms of Reference

Document Information	
Document Name	DAC_Terms_of_Reference
Author(s)	Aida Sanchez
Version	5
Issue Date	December 2020
Approved By	CLS Data Access Committee
Review frequency	Yearly

Document History		
Version	**Date**	**Summary of change**
1	Dec 2008	Draft document
2	Dec 2016	Draft document that listed all ToR
3	Sept 2018	Final document with updated information on how the CLS DAC operates
4	Sept 2020	Reformatting, addition of the CLS fast track approvals, addition of the sharing of NHS Digital data, NHS Scotland data and minor text updates
5	Dec 2020	Add genetic data sharing approvals

Contents

1. Scope

This documents sets out the responsibilities, membership and mode of operation of the CLS Data Access Committee (CLS DAC).

2. Responsibilities

(i) To agree policies and procedures for accessing all types of data collected by CLS, as set out in the CLS Data Access Framework. This includes mechanisms whereby CLS research data is available to researchers through designated archival services operating to approved standards of access and security, such as the UK Data Service (UKDS).

(ii) To ensure that of all of the different CLS data access routes, such as the UK Data Service, other national repositories or CLS data release, are promptly reviewed, fully documented and reported to the committee on a regular basis.

(iii) To apply the **CLS Data Classification Policy** to CLS research data, and to review any changes regarding data classification schemas already applied.

(iv) To ensure that applications for **data deposited at the UKDS under Special Licence (Tier 2a) and Secure Access (Tier 2b**), which are signed off by the CLS Research Data management team on behalf of the PI of the relevant study, are treated equitably across studies, and to provide advice and guidance where new issues have arisen.

(v) To consider applications for **CLS research data not available via public repositories such as UKDS**, SAIL Databank or EGA, as described in the guidelines included in the CLS Data Access Application document. This will include applications for main survey data, paradata or genetic data .

(vi) To consider applications for **novel data linkages** to the CLS studies. This may involve linkage with new geographical indicators or with new

administrative data such as health, education, finance or other. These applications will be dealt with as described in the guidelines included in the CLS Record Linkage Application document.

(vii) To consider applications for **data enhancements** to the CLS studies – either for new data collection from survey participants, requests for additional questionnaire/survey time within an existing survey instrument or for novel legacy data projects that involve not transcribed or digitised legacy data. These applications will be dealt with as described in the guidelines included in the CLS Data Enhancements Application document.

(viii) To consider any emerging data access policy issues arising and to refer these to the CLS Strategic Advisory Board for advice.

(ix) To address the CLS assessment criteria set out in the CLS Data Access Framework in determining whether data access should be granted.

(x) If appropriate, to take the advice of third party specialist knowledge, particularly where an application has not already been through established peer review mechanism.

(xi) The Committee will provide regular updates to the CLS Strategic Advisory Board, including a summary of decisions taken, as well as any underlying issues or concerns that the members of the DAC wish to report to the board.

(xii) CLS DAC will publish the information about any linked health data dissemination to which NHS Digital is the data provider, on the NHS Digital release register. This should include the name of the organisation the data was provided to, the purpose (summary of the project) and details of the data released.

(xiii) Any requests for NHS Scotland data that cannot be handled by the UKDS will need to be evaluated by the CLS DAC and referred to the Scottish Public Benefit and Privacy Panel for Health and Social Care (PS PBPP).

3. Membership

- Chair – CLS Director and SIRO
- Principal Investigators of the four CLS studies: NCDS, BCS70, Next Steps and MCS
- CLS Senior Data Manager
- CLS Records Linkage Manager
- A member of UKDS: Deborah Wiltshire, Senior User Support & Training Officer
- Staff from across CLS/UCL may attend to advise on specialized research areas such as genetics, incidental findings and geography.
- An ethics expert to advise on socially controversial research or on genetics.

4. Mode of operation

4.1 CLS DAC Secretariat

The CLS DAC Secretariat will deal with the administration and documentation of all of the DAC applications and DAC related issues. Their responsibilities will be to:

1. Manage the DAC related queries from external and CLS data users
2. Manage all received DAC applications, from receipt to approval, gathering of feedback the committee members and corresponding with the applicants on the CLS DAC decision
3. Design and maintenance of a register of applications and decisions
4. Organise all DAC-related data release and keep an up to date record of the status of each application
5. Generate the monthly UKDS data sharing report
6. Organise the monthly DAC meetings and writing the minutes
7. Design and maintenance of the DAC application forms

8. Manage the content management of the DAC page on the CLS website in collaboration with the CLS Communications team

4.2 Standing agenda items

The Committee will consider at least the following items at each meeting:

- Monthly UKDS data sharing report
- CLS DAC applications
- Changes to data classification

4.3 Changes to data classifications

In determining new and amended CLS data classification the Committee will consider the issues enumerated in the terms of reference point (viii) above. Decisions of the Committee on the data classification will be by majority vote. In the event that either a) a majority decision amongst Committee members is not reached or, b) the majority view of the Committee and the PI of the relevant study do not agree, the Chair of the CLS DAC will refer the decision to the CLS Strategic Advisory Board.

Any appeal against the CLS DAC's decision will be considered by the Strategic Advisory Board.

4.4 Approvals of CLS DAC applications

An application is completed by the individual/ persons making the request. The CLS DAC has a dedicated email address, clsfeedback@ucl.ac.uk, for any queries. Applications to be considered by the Committee are sent to this address.

There are three types of DAC applications:

1. Data access
2. Novel record linkage
3. Data enhancement

The Committee will meet monthly, at which applications requiring a decision from the CLS DAC will be considered.

The DAC applications will be circulated to the Committee a week before the DAC meeting, in order to allow enough time to the committee members to read and evaluate the application in advance.

Decisions of the Committee on whether to grant access to applications, in the light of the data classification made and other considerations, will be by majority vote. In the event that either a) a majority decision amongst Committee members is not reached or, b) the majority view of the Committee and the PI of the relevant study do not agree, the Chair of the CLS DAC will refer the decision to the CLS Strategic Advisory Board.

Decisions about which requests will be approved will be based on the information provided in the application form and the set of criteria set out on the CLS Data Access Framework, as follows:

1. Sensitivity of data;
2. Data minimisation: the amount of data requested must be justified by the purpose stated within the application
3. Risk of re-identification of study respondents by researchers
4. Welfare of one or more study respondents;
5. Impact of the conduct or reporting of the research on public perception or the future viability of the data collection.
6. Data security
7. Organisational Security assurance for those applications which are requesting linked NHS Digital health data
8. Legal basis for processing personal sensitive data for those applications which are requesting linked NHS Digital health data
9. Where relevant, sample depletion);

10. Public benefit, including potential scientific and wider impacts of the proposed research
11. Benefits to healthcare provision, adult social care or the promotion of health, for those projects which are applying to access linked health data for those applications which are requesting linked NHS Digital health data.

4.5 Fast track approvals

The Committee have delegated to the CLS Research Data management team (RDM) the capability to evaluate and approve applications for data deposited at the UK Data Service under Special Licence (Tier 2a) and Secure Access (Tier 2b). The RDM will seek advice and guidance of the Committee where new issues have arisen.

The RDM approval on behalf of Committee is referred to as the CLS 'fast track' approval.

In addition, from September 2020 the fast track approval will be applied to some CLS DAC applications. This RDM approval on behalf of the Committee will only take place for CLS DAC applications that are straightforward, and will only apply to:

a) Simple CLS DAC requests for Tier 1 data (EUL) or Tier 2a data (Special Licence) not yet deposited at the UKDS.
b) Amendments to existing CLS DAC applications, such as the changes to the research users or straightforward requests for additional data.

The RDM team will record the CLS DAC approved applications accordingly and will report them at the next monthly DAC meeting.

This additional approval system will free up time in the DAC meetings to discuss more complex DAC applications and strategic issues arising.

4.6 Approvals of access to genetic data linked to survey/biomedical data

When users wish to access genetic and survey/biomedical data, this can potentially increase the disclosure risk, and so such applications demand careful linkage of the

relevant data to enable secure analysis at an individual level. A research group may even require access to survey data, biomedical phenotypes, GWA genotypes, cell-line DNA, and blood samples. All of these must be linked together at an individual level in a manner that prevents end-users (DAC research applicants) from identifying individual participants, either from the resources they have been awarded, or by joining their data together with another end-user who has been awarded a different set of data.

In order to achieve this level of security, CLS DAC requests for genetic data combined with phenotypic data are subject to separate data minimisation and pseudo-anonymisation arrangements that require the creation of a bespoke phenotypic dataset identified by newly created IDs, and the data are prepared and released from a modular manner.

a) Data minimisation: bespoke phenotypic datasets

As part of the CLS DAC Data Access application, researchers will need to submit a list of survey data variables names they need to link to the genetic data. This data minimisation strategy offers additional protection to the genetic data, given their potential sensitivity and classification as Tier 2a personal data.

These variables need to be publicly available under the UKDS End User Licence. The applicants will have to provide a summary of how this list of variables fits in with the project, but they don’t need to justify how they intend to use every single variable (e.g. as exposure, confounder, outcome).

The final phenotypic dataset will be a bespoke dataset that only contains the exact list of variables requested by the applicant.

b) Pseudo-anonymisation: newly created IDs

Once an application has been approved, a new ID will be created to identify the requested phenotypic and genetic data for every CLS DAC research team. This ID will always be different to:

- the public ID used to identify the study data deposited at the UK Data Service or other national repositories.

- the public ID used to identify the genetic data deposited at the EGA

- the IDs released to other CLS DAC applicants.

Generic IDs or ID lookups will never be released to end-users.

4.7 Approvals of access to finite biological samples

Tissue samples collected from the study members of the CLS cohorts are a finite resource. The principles applied in depleting biological samples include the scientific strength of the proposal and the appropriateness of the methodology proposed. CLS holds responsibility for applying these principles to specific applications for the depletion of finite biological resources, as follows:

- All applications to use samples should demonstrate a clear scientific rationale regarding why the study is appropriate to the proposed research, and for non-renewable samples, that the use of samples is justified by the expected contribution to the scientific body of knowledge. Applications that demonstrate a unique dependence on the study, for example use of longitudinal data not widely available, are preferred.
- Appropriate ethical approval must be in place and all applications must comply with relevant legislation, e.g. the Human Tissue Act.
- Scientific strength, novelty and potential health/social impact of the research proposal must sufficiently justify use of longitudinal study samples.
- Evidence must be provided to show methodology is appropriate to the processing history of the samples. e.g. published literature or pilot data.
- The assay test platform should have proven quality assurance measures in place, preferably in accredited facilities according to ISO standards.
- The assay strategy should aim for maximum research impact with minimal depletion of the resource.
- The methodology should include measures to ensure the quality of any remaining sample is not jeopardised and can be used in further assays.

- All data generated from samples must be returned to the study and made available to other users within an agreed timeframe

4.8 Approvals of access to linked NHS Digital health data

When reviewing applications to access linked health data, the committee will consider the NHS Digital requirements as enumerated in the terms of reference,point (viii). These are explained in detail below:

a) **Data minimisation-** CLS DAC will consider if the HES variables requested by the data applicant is justified by the purpose stated in the project proposal, by assessing the relevance of the datasets to the purpose of the project. If CLS DAC considers that the applicant did not justify the amount of data being requested, CLS can request the applicant to either provide further explanation or re-select the variables that fit in with the scope of the research project.

Further information on the NHS digital data minimisation requirement can be accessed via the link below:

https://digital.nhs.uk/services/data-access-request-service-dars/dars-guidance/data-minimisation

b) **Organisation Security Assurance-** CLS DAC will verify that the applicant has provided evidence of having an organisation security assurance (e.g. Security Level Systems Policy (SLSP) or Data Security and Protection toolkit (DSPT) or International Organisation for Standardisation (ISO27001)), as outlined in the '*CLS Licence Agreement for Linked NHS-Digital Data'.* Approval for data access will only be granted to organisations that meet the security assurance requirement. Failure to meet the organisational security requirements will result in CLS DAC requesting for the applicant to provide further evidence, and the project will only be re-submitted to CLS DAC for approval when evidence provided is satisfactory.

c) **Legal basis-** CLS DAC will verify that the organisation requesting access to linked health data has a legal basis for processing the data.

d) **Expected measurable benefits to health and social care-** CLS DAC will review the stated purposes of the application and how it benefits healthcare, adult social care or the promotion of health and will decide whether it is satisfied with the answer provided. If CLS DAC is not satisfied with the statement provided by the applicant, CLS DAC can ask the applicant to provide additional information and the project can be re-submitted for approval on the next CLS DAC meeting or via Chair approval. Approval for data access will only be granted if the Committee is satisfied with the proposed benefits.

For Committee members- Further information on the NHS Digital expected measurable benefits to health and social care requirement can be accessed via the link below. https://digital.nhs.uk/services/data-access-request-service-dars/dars-guidance/expected-measurable-benefits

5. Risks

The Risk Register for CLS is owned by CLS and is shared with the CLS Strategic Advisory Board so that the board can discharge its duty to provide assurance to the funders. The following risks should be monitored by the DAC:

(i) Respondents and/or Policy Community withdraw support due to concerns over security of data, for example in relation to the administrative data linkage aspects of CLS;

(ii) There is a delay to the data being granted by the CLS DAC being made available to applicants through the UK Data Service or other mechanism;

(iii) The access procedures become a barrier to wide use of CLS Data;

(iv) The Committee is overwhelmed by applications for access to CLS Data and does not have the resources to ensure that the data is made as widely available as possible.