CLS data linkage application form

The CLS Data Access Committee (CLS DAC) welcomes applications to perform additional data linkages as part of this programme of work.

***Please read the*** [***CLS data linkage guidelines***](#_Guidelines_for_proposals) ***prior to completing this form.*** *You must complete the most recent version of this form, which can be found on the* [*CLS data linkage webpage*](https://cls.ucl.ac.uk/data-access-training/data-enhancements/)*.*

## Main data applicant

For applications by students, the main data applicant must be the project supervisor.

|  |  |
| --- | --- |
| **Name:** |  |
| **Affiliation and work address:** |  |
| **Academic email address:** |  |
| **Website (e.g., ORCID or institutional website):** |  |
| **Telephone:** |  |

## Research team and collaborators

Please list the names of all members of your research team who need access to the data, their affiliation, and their contact email addresses.

|  |  |  |
| --- | --- | --- |
| **Name** | **Affiliation** | **Email address** |
|  |  |  |
|  |  |  |
|  |  |  |

You may add rows for additional users if required.

## Type of data to be linked

Health

Education

Economy

Geographical

Other (please specify):

|  |
| --- |
|  |

## CLS cohort(s)

Please indicate which CLS cohorts you propose to link data to:

National Child Development Study (NCDS, or the 1958 Birth Cohort Study)

1970 British Cohort Study (BCS)

Next Steps (Next Steps)

Millennium Cohort Study (MCS)

## Title of the research project

|  |
| --- |
|  |

## Plain-language abstract summarising the research project

Please provide a short description of the proposed research you plan to do with the data, including the main research question(s). Maximum 150 words.

The plain-language abstract will be published online and should be written in accessible language that cohort members would readily understand. For further guidance on writing a plain-language abstract, please see Appendix 1 of [the CLS Data Access Framework](https://cls.ucl.ac.uk/wp-content/uploads/2020/12/CLS_Data_Access_Framework.pdf).

|  |
| --- |
|  |

## Project keywords

Please add at least five keywords related to your project, for publication online.

|  |
| --- |
|  |

## Description of the research project

Please provide a full and detailed description of the purpose for which the data are requested, describing the aims of the study/research, methodology, and ethico-legal considerations. Please specify if the research project is part of a larger programme.

|  |
| --- |
| 1. Research project description *(up to 500 words)* 2. Brief methodology description *(up to 500 words)* 3. Ethico-legal issues, in particular regarding sensitive or socially controversial topics. *We strongly advise that you consult section 6.8 of* [*the CLS Data Access Framework*](https://cls.ucl.ac.uk/wp-content/uploads/2020/12/CLS_Data_Access_Framework.pdf) *for guidance.* |

## External data to be linked

Please provide the name of the source and/or the Data Controller/Owner of the external data to be linked:

|  |
| --- |
|  |

Please describe the data to be linked:

|  |
| --- |
|  |

## Consent needed for data linkage

Will the data be based on consent for data linkages that have been collected in the study to date, or will this rely on future collection of consent?

Based on existing consents

Requiring new consents to be collected

Requires unconsented linkage

Please explain your answer further. If applicable, please include a specific justification for any unconsented linkages that are requested:

|  |
| --- |
|  |

## Cost of data linkage

Please provide an estimate of the likely cost of data linkage, for example any charges made by the Data Controller in order to carry the linkage out:

|  |
| --- |
|  |

## Has the project been funded?

Funding is not mandatory but provides an additional reassurance for the project.

Yes. Please provide the name of the funder and the end date of the funding:

|  |
| --- |
|  |

No. If the project is not funded or you are waiting to hear from a funder, please provide a statement on whether the project will still go ahead without funding and how you will proceed:

|  |
| --- |
|  |

## Has the project received ethical approval?

Ethical approval is not mandatory for this DAC application but might be required by your institution.

Yes. Please provide the name of the Ethics Committee Board, reference, and date of approval:

|  |
| --- |
|  |

No. Please explain why the project has not received ethical approval:

|  |
| --- |
|  |

## Statement on Principal Investigator role

CLS’ policy is that the Principal Investigator of the study would normally be the Principal Investigator for any funding application to enact data linkages in the CLS cohorts, and for all formal applications for linkage to any Data Controller, and any other associated applications. Please comment if this causes any particular issues for your application:

|  |
| --- |
|  |

## Project dates

Please specify when you estimate that the project will start and finish.

|  |  |
| --- | --- |
| Start date (MM/YYYY) |  |
| Finish date (MM/YYYY) |  |

## Data outputs

Please describe what derived variables or other individual-level data outputs you plan to generate or compute as part of your proposal. Relevant data outputs and documentation must be sent to CLS if requested.

|  |
| --- |
|  |

## Data access

Please specify the server where the data will be accessed and stored and indicate if it is different to the organisation/institution servers specified in section 1.

|  |
| --- |
|  |

## Potential **use by the research community**

Please describe what research questions which could potentially be answered by other researchers using the data created:

|  |
| --- |
|  |

## CLS conditions of data use

This section concerns the conditions of use of Data that form part of the research Project hereafter referred to as 'the Materials'. The term ‘Data’ refers to the Data requested, as well as any new derived variables or data arising from this CLS DAC application. CLS data sharing is governed by the [CLS Data Access Framework](https://cls.ucl.ac.uk/wp-content/uploads/2017/02/CLS_Data_Access_Framework.pdf).

In signing this Agreement, the main data applicant agrees to use the Materials according to the CLS conditions listed below:

1. Use the Materials only for the research Project approved by the CLS DAC.
2. Understand that the Materials requested, and any Data or documentation returned to CLS, may be shared by CLS for wider use via the UK Data Archive or other repositories.
3. Preserve the confidentiality of, and not attempt to identify individuals, households, or organisations. Materials cannot be linked or combined to other information in a way that could re-identify the research participants, even if access to those Materials has been granted.
4. Store the Materials securely in an organisational setting; if Data are placed in a shared directory or on a Local Area Network (LAN), access must only be available via personal authentication
5. Be aware of, and follow, any information security guidelines provided by the Recipient’s Institution.
6. Ensure that the means of access to the Data (such as passwords) are kept secure and not disclosed to anyone.
7. Notify UCL/CLS of any non-compliance they are aware of. Any incidents of unauthorised access to, Processing of, or disclosing of the research Materials must be reported immediately. Such incidents will be considered as data breaches and CLS will report them to the UCL Information Security Group (ISG) and, where the breach is notifiable within the applicable laws, to the Information Commissioner’s Office (ICO) and the data subjects.
8. Not use the Materials for commercial purposes.
9. Notify CLS of any errors discovered in the Materials by emailing [clsdata@ucl.ac.uk](mailto:clsdata@ucl.ac.uk).
10. Understand that CLS does not transfer any interest in Intellectual Property to the Recipient’s Institution.
11. Understand that the Materials are provided without warranty or liability of any kind.
12. The Materials must not be passed to a third Party (i.e., to anybody that is not included in the list of applicants on this Project, nor is a direct employee of one of these applicants). This includes sharing of publicly available individual level Data. If collaborators are based in different organisations, they should sign a separate Data Sharing Agreement (DSA).
13. At the request of CLS, return derived variables and related documentation arising from this request to CLS within 12 months of receipt of the requested Data.
14. Ensure that all reporting of Project results and press releases use careful and balanced language in order to avoid misinterpretation or exaggeration of the findings. Live interactions with the media (including social media) should use careful and balanced language.
15. At the end of the Project, Data must be deleted from the system on which it has been stored using a secure erasure programme. The recycle/trash bin must be emptied and immediately followed by the running of a secure erasure programme.
16. Inform CLS of any publications that arise from this Project, which CLS will cite and publish on the CLS website.
17. Be aware that CLS will publish anonymised basic information about this CLS DAC approved Project on the CLS website.
18. Understand that this Agreement is subject to review and without limitation whenever a change in the law, contracts for services with third Parties, other procedures, or other relevant circumstances take place.
19. Agree to comply with the General UCL Terms and Conditions described in Section 6 and with the Data Protection requirements described further in Schedule 1. In case of a conflict with Schedule 1, the provision of sections 5 and 6 shall prevail.
20. Understand that any non-compliance with these terms and conditions will lead to immediate termination of access to the Data and could result in legal action.

[SIGNATURE BELOW]

**Signed:** Click or tap here to enter text.

**Print name:** Click or tap here to enter text.

**Date:** Click or tap here to enter text.

# CLS data linkage guidelines

## Introduction

The Centre for Longitudinal Studies (CLS) is responsible for four national cohort studies: the National Child Development Study (NCDS, or the 1958 Birth Cohort Study), the 1970 British Cohort Study (BCS70), the Millennium Cohort Study (MCS) and Next Steps, which follow the same group of people from childhood and throughout their lives.

CLS research data are publicly available to researchers through a number of data repositories. The vast majority of CLS research data are available from the UK Data Service (UKDS):

* [NCDS](http://discover.ukdataservice.ac.uk/series/?sn=2000032)
* [BCS70](http://discover.ukdataservice.ac.uk/series/?sn=200001)
* [MCS](http://discover.ukdataservice.ac.uk/series/?sn=2000031)
* [Next Steps](http://discover.ukdataservice.ac.uk/series/?sn=2000030)

The Centre for Longitudinal Studies (CLS) has a programme of data linkages underway, which covers a wide range of external data such as health, education, geographical, and economic indicators linked to its four longitudinal studies. These linked data are largely made available through the UKDS. As part of this programme of work, CLS welcomes proposals to perform additional and novel data linkages of its four cohort studies with external data sources.

The functioning of the CLS DAC is described in detail in the [CLS DAC Terms of Reference](https://cls.ucl.ac.uk/wp-content/uploads/2020/12/CLS_DAC_Terms_of_Reference.pdf).

CLS DAC proposals may refer to linkages with external data sources such as:

* Health
* Education
* Geographical data
* Economy
* Other

## Request process

To submit a data linkage application please follow these steps:

1. Those wishing to submit a proposal for a data linkage on one of the CLS cohort studies should get in touch with the CLS Records Linkage team at [*clsrecordslinkage@ucl.ac.uk*](mailto:%20clsrecordslinkage@ucl.ac.uk%20) to discuss their request.
2. Complete the [CLS data linkage application form](#_CLS_External_Data).
3. Send the completed application form to [clsdata@ucl.ac.uk](mailto:clsdata@ucl.ac.uk).

CLS data can be used according to the consents and ethical approvals of the studies and cannot be varied under any circumstance. The applicant’s funders’ and publishers’ requirements must be compatible with these and cannot require anything in conflict with the CLS conditions of data use ([section 19 of the CLS data linkage application form)](#_CLS_conditions_of).

Researchers must allow plenty of time between submitting their application and when they plan to undertake research on the data requested. In cases where significant resources of the CLS data management team are required to fulfil the request, we suggest an application is submitted at least 6 months before the planned research will take place.

Requests to **access data not publicly accessible or to biosamples** require the submission of the [CLS data access application form](https://cls.ucl.ac.uk/data-access-training/data-access/accessing-data-directly-from-cls/).

Requests to **collect new CLS data** or to **enhance the legacy data resource**, e.g.,by digitisation, coding, or transcription, are governed by separate CLS guidelines and require the submission of the [CLS data enhancement application form](https://cls.ucl.ac.uk/data-access-training/data-enhancements/).

## CLS DAC approval process

All proposals are discussed at a meeting of the CLS Data Access Committee (CLS DAC), which typically meets every month.

The CLS DAC’s decisions will be based on the information provided in the application form and accompanying documentation. The approval criteria are listed in section 8.2 of the [CLS Data Access Framework](https://cls.ucl.ac.uk/wp-content/uploads/2017/02/CLS_Data_Access_Framework.pdf).

Decisions about which linkages to support will be based on the information provided in the application form and a clear set of criteria as follows:

* 1. The scientific merit of the proposed linkage. If necessary, the CLS DAC may seek further expert input on this, though this is not intended to replace a peer review process at the funding proposal stage.
  2. Whether the proposed study needs to take place in the specific cohort study proposed, or could feasibly occur elsewhere.
  3. Respondent burden if consent collection is requested.
  4. The potential wider benefits of the linkage to the scientific community beyond the research set out by the proposer.
  5. Other possible ethical issues.
  6. CLS resources required to deliver the linkage, and any risk posed by the linkage to CLS’ ability to deliver on its core mission or other existing commitments.

Following the CLS DAC discussion, the applicant of the data linkage may be asked to respond to any queries raised by the Committee and to modify their proposal to render it suitable for further consideration.

### Principal Investigators (PIs) and co-PIs

Proposals for data linkages that need to be centrally managed by CLS rather than by the applicant should name the PI of the cohort study as the PI of the project. This will be decided by the CLS DAC at the time of the evaluation and approval of the proposal. In these cases, the team or researcher proposing the linkage would normally be Co-Investigator(s) of the linkage project. This allocation of PI and Co-I roles does not reflect the extent of the scientific contributions of the investigators. Instead, it is intended to ensure that the PI of the cohort study retains full responsibility for the conduct of the study, and that CLS retains appropriate control over issues such as ethical permissions, data security, and the timely documenting and deposit of data. This condition relates to all types of proposals enumerated above.

## Enacting the approved data linkage

Any ethics-related applications, applications to Data Controllers, or to funders for data linkage will be made by the PI of the relevant study, in collaboration with the proposer of the enhancement.

Linkage of data from an external source to the individual cohort member data IDs will be enacted according to the protocol agreed between CLS and the external Data Controller. CLS will be responsible for any transfer of data with the external Data Controller.

A procedure will be documented and agreed at the outset of the project for collaborators to access linked data under the supervision of the Data Controllers and/or CLS as appropriate, in order to undertake quality assurance, or to code up derived variables from linked data.

CLS is committed to the release of data to the research community. Agreement will be reached between the PI of the study and the enhancement proposer, in discussion with Data Controllers, on the scope and content of data to be documented and deposited at the UKDS or another appropriate repository.

## Data release

Once access has been granted, research data will be released via a suitable mechanism, free of charge.

The data release mechanism will depend on the level of disclosivity and sensitivity of the data requested, which will determine the relevant “tier” according to the [CLS Data Classification Policy](https://cls.ucl.ac.uk/wp-content/uploads/2017/02/CLS_Data_Classification_Policy.pdf). The four research data tiers are:

* Safeguarded data – Tier 1a: low level of disclosure and sensitivity. This corresponds to the UKDS End User Licence.
* Special safeguarded data – Tier 1b: medium level of disclosure and/or sensitivity. This corresponds to the UKDS Special Licence.
* Controlled data – Tier 2: high level of disclosure and/or sensitivity. This corresponds to the UKDS Secure Access.
* Special controlled data – Tier 3: very level of disclosure and/or sensitivity. These data must be accessed via the UCL Data Safe Haven.

The CLS DAC considers appeals from researchers regarding the tier level allocated to any CLS data. For further information on how to proceed, please contact [clsdata@ucl.ac.uk](mailto:clsdata@ucl.ac.uk).

Safeguarded data (tiers 1a and 1b) and controlled data (tier 2) will generally be shared via the UKDS. However, depending on CLS capacity, interim arrangements will be made occasionally to provide the applicants with the requested data prior to these becoming available at the UKDS.

CLS data releases will be governed by a CLS data sharing agreement (DSA) and subject to data security arrangements equivalent to those from the UKDS.

In circumstances where access is granted to special controlled data (tier 3), this would only be permitted via the highly secure UCL Data Safe Haven.

All researchers working on CLS data will observe the CLS conditions of data use ([section 18 of the CLS data linkage application form](#_CLS_conditions_of)), and will not attempt to identify individual cases, or share data with other unauthorised person(s).