**CLS Data Linkage Request**

# **CLS External Data Linkage Application Form**

*DAC REF:*

The CLS Data Access Committee (CLS DAC) welcomes applications for data enhancements that relate to a new data collection; addition of additional questions within an existing survey instrument; digitisation, transcription and/or coding of legacy data.

***Please read the*** [***CLS Guidelines for Record Linkage***](#_Guidelines_for_proposals) ***prior to completing this form.***

## Main data applicant

For PhD projects, the main data applicant on the application must be the supervisor.

|  |  |
| --- | --- |
| **Name:** |  |
| **Affiliation and work address:** |  |
| **Academic email address:** |  |
| **ORCID link:** |  |
| **Telephone:** |  |

## Research team and collaborators

Please list the names of all member(s) of your research team who need access to the data.

|  |  |  |
| --- | --- | --- |
| **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 study or studies to which records are proposed to be linked.

National Child Development Study (1958 Birth Cohort)

1970 British Cohort Study

Next Steps

Millennium Cohort Study

## Title of the research project

|  |
| --- |
|  |

## Plain language abstract of the research project

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

|  |
| --- |
|  |

## 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 and the potential impact of the research for policymakers and/or practitioners. Where research is part of a larger programme, please include details below.

|  |
| --- |
|  |

## External records to be linked

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

|  |
| --- |
|  |

Please describe the data to be linked:

|  |
| --- |
|  |

## Consent needed for linkage

Will the records be based on consent for record 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:

|  |
| --- |
|  |

## Ethical issues

Please specify any other ethical issues arising:

|  |
| --- |
|  |

## Cost of record linkage

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

|  |
| --- |
|  |

## Funding

An indication of potential funders and whether any preliminary approaches have been made to funders:

|  |
| --- |
|  |

CLS’ policy is that the Principal Investigator of the study would normally be the Principal Investigator for any funding application to enact record 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 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 1.1.

|  |
| --- |
|  |

## Data outputs

Please describe what derived variables or other data outputs are planned as part of your linkage proposal. At the request of CLS, relevant data outputs and documentation must be sent to CLS.

|  |
| --- |
|  |

## Plans for research outputs resulting from the study

|  |
| --- |
|  |

## 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

All data from the CLS cohort studies are documented and archived for potential use by the wider research community. Therefore, I understand that the data that I request, and any data or documentation I return to CLS, may be shared by CLS for wider use via the UK Data Archive.

I agree that I will use the newly created data according to the CLS conditions listed below:

* Use the data only for the research project approved by CLS DAC
* Preserve the confidentiality of, and not attempt to identify, individuals, households or organisations. Data cannot be linked or combined to other information in a way that could re-identify the research participants, even if access to that data has been formally granted
* Store the data 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.
* Be aware of, and follow, any information security guidelines provided by the applicant’s organisation.
* Ensure that the means of access to the data (such as passwords) are kept secure and not disclosed to anyone else.
* Notify CLS of any non-compliance they are aware of. Any incidents of unauthorised access to, processing of, or disclosing of, the research data must be reported immediately
* Not use the data for commercial purposes.
* Notify CLS of any errors discovered in the data by emailing *clsfeedback@ucl.ac.uk*.
* Understand that CLS does not transfer any interest in intellectual property to the Recipient Institution.
* Understand that the data are provided without warranty or liability of any kind.
* Ensure that the data must not be passed to a third party (i.e., to anybody that is not included in this list of applicants on this project, nor is a direct employee of one of these applicants). This would include any sharing of publicly available individual level data.
* 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.
* 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, preferably to be immediately followed by running a secure erasure programme.
* 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.

**Signed: Print name:**

**Date:**

# **Guidelines for proposals of CLS Record Linkage**

## Record linkage via the CLS DAC

The Centre for Longitudinal Studies (CLS) is responsible for four national cohort studies that follow the same group of people from childhood and throughout their lives:

* the National Child Development Study (NCDS), or the 1958 Birth Cohort Study
* the 1970 British Cohort Study (BCS70),
* the Millennium Cohort Study (MCS)
* Next Steps

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

The CLS Data Access Committee (CLS DAC) welcomes proposals for novel record linkages to its four cohort studies. 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:

* Geographical data
* Education
* Health
* Economy
* Other

## Request process

To apply for a novel record linkage to CLS cohorts, 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 [Principal Investigator (PI) of the relevant cohort study](http://www.cls.ioe.ac.uk/page.aspx?&sitesectionid=794&sitesectiontitle=Data+enhancements) to discuss their request:

* NCDS and BCS70: Alissa Goodman, [alissa.goodman@ucl.ac.uk](mailto:alissa.goodman@ucl.ac.uk)
* Next Steps: Lisa Calderwood, [l.calderwood@ucl.ac.uk](mailto:l.calderwood@ucl.ac.uk)
* MCS: Emla Fitzsimons, [e.fitzsimons@ucl.ac.uk](mailto:e.fitzsimons@ucl.ac.uk)

1. Upon agreement with the PI, please complete the [CLS Data Linkage Application Form](#_CLS_External_Data)
2. Send the completed application form to [clsfeedback@ucl.ac.uk](mailto:clsfeedback@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.

Researchers must allow plenty of time between submitting their application and when they plan to undertake research on the data they are applying for. In cases where significant resources of the CLS 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 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/).

**Principal Investigators and co-PIs**

Data linkages should name the PI of the cohort study as the Principal Investigator of the project. The team or researcher proposing the enhancement would normally be Co-Investigator(/s) of the project. This allocation of PI and Co-I roles does not reflect the extent of the scientific contributions of the investigators. Instead, 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.

## CLS DAC approval process

All proposals for data linkages will be discussed at an internal meeting of the CLS DAC, attended by the PIs, a CLS Records Linkage manager and the Senior Data Manager. The CLS DAC meets typically every month.

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

## Enacting the approved record linkage

Any ethics-related applications, applications to data controllers, or to funders for record linkage will be made by the Principal Investigator of the relevant study, in collaboration with the proposer of the enhancement.

Linkage of data from an external source to the individual cohort member records’ IDs will be enacted according to the protocol agreed between CLS and the data controller. CLS will be responsible for any transfer of data between a data holder and CLS.

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

All researchers working on CLS data will protect the privacy of cohort members, and will not attempt to identify individual cases, or copy any data to other people who may not have signed the appropriate undertaking. All data released from the project to applicants (outside of CLS or a nominated data controller) will be de-anonymised.

All researchers making a proposal for a record linkage will be asked to sign a declaration that the linked data will be documented and archived for use by the wider research community as soon after data collection and preparation as practically possible (usually within 12 months).