CALIBER Resources, Fees and the Project Cycle

2018

A fee of £25,000 per Independent Scientific Advisory Committee (ISAC) approved project will apply to cover CALIBER Data Lab's data management and support (cohort definition, data extraction) and use of specialised infrastructure such as the Data Portal and the safe haven environment. We encourage researchers to consider these costs in grant applications.

There is no additional fee to access CPRD licensed data for any organisation who holds a site licence directly with CPRD to access the data required for the specific collaborative ISAC-approved project. If the collaborating research organisation does not hold a licence to CPRD data, data access licence fees will be applied.

A guide to fees can be found here. Costs will vary according to the type of organisation/ affiliation. To discuss costs or if you have any questions, please email Natalie Fitzpatrick, Data science facilitator.

What does a researcher get for this? The £25k fee provides access to the following resources across the CALIBER project cycle (see figure below) relevant to building the CALIBER community of investigators. To note the £25k fee does NOT include access to CPRD data.

Project set-up
- Agree deliverables and timescales
- ISAC preparation: we will provide resources to guide researchers through the approvals process and forms, including sample ISAC applications
- Data sharing agreement and third party agreement (if non-UCL)*
- Farr Institute data safe haven registration
- Identification of training requirements
- Organise UCL visitor status (if applicable)

Data preparation and release of data
Preparation of cohort: development of phenotype and identification of data items. For example:
- Identify study population and initial set-up
- Identification of relevant data elements
- Initial implementation and consistency checks
- Revised/ adjusted implementation of data extraction algorithm (as required) and consideration of additional codes and cases if necessary
- Final script implementation: develop finalized algorithm and coding list(s) and generate final cohort
- Transfers project-specific data onto data safe haven
- Follow-up, as required, including adjustments to final cohort, request for additional covariates, requesting guidance on covariate selection or amending errors

CALIBER methods and tools
Central online repository for the storage and distribution of algorithms, documentation and metadata including:
- Access to specialist tools including the CALIBER data portal, an open source, online resource consisting of >500 phenotype algorithms

- Statistical packages including the CALIBER Health Records Research Toolkit (a collection of R packages to assist in epidemiological studies using EHR databases for example to create research cohorts and phenotypes)
- Tools to aid research cohort creation are available including computable research cohort definitions and online tool, and tools for manipulating OPCS4 and ICD-10 codes

CALIBER integrated training and infrastructure

- Information governance
- Use of the data safe haven and high performance computing (e.g. Legion)
- CPRD e-Learning module**
- Methods in coding and analysing large, linked datasets
- Tools to support development of CALIBER community, including live interactive demonstrations of data and monthly CALIBER project clinics
- Additional bespoke community training may be arranged as per individual training requirements at additional cost

CALIBER community

Researchers will join the growing CALIBER community which currently consists of researchers from over ten countries, 13 Higher Education Institutes, 25 research groups and Industry partners to bring together research across scientific and clinical disciplines

Knowledge transfer and building the resource

- Implementation of phenotyping algorithms
- Provision of a Digital Object Identifier (DOI)

There is no fee for initial scoping discussions and preliminary feasibility work (e.g. simple counts).

Pilot data analysis for grant submissions (i.e. beyond simple frequency counts) should be discussed with the data science facilitator. We aim to provide such services without charge, however requests will be considered on a case-by-case basis and charges may apply where requests are complex or time-consuming in nature.

Requests from non-academic and industry organisations will be subject to an additional 40-50% on top of the total cost above.

For more information about costs and collaboration, or to arrange an informal discussion, please contact Natalie Fitzpatrick, Farr Institute data science facilitator by email at n.fitzpatrick@ucl.ac.uk.

---

* Any party using CPRD data must have a direct or sub-licence to CPRD data
** As part of CPRD's standard terms any CPRD licensee using CPRD data must complete the CPRD e-learning module

UCL

Left column flowchart (Point of entry legend):

- POINT OF ENTRY
- APPROVALS AND GOVERNANCE
- DATA SPECIFICATION AND EXTRACTION
- TRAINING AND ANALYSES
- KNOWLEDGE TRANSFER
- PROJECT CLOSURE

CALIBER Project cycle

Cycle steps (clockwise from top):

- Researcher initiates idea and identifies funding
- CALIBER review for feasibility and project approval
- CALIBER supports researcher application for ISAC approval
- ISAC approval
- Researcher signs Data Sharing Agreement
- Researcher signs Sub-Licence agreement (non-UCL investigators)
- CALIBER submits data linkage form to specify variables required
- CPRD carries out linkage and transfers data to UCL
- CALIBER preparation and extraction of data
- Researcher receives data
- Researcher undergoes integrated training
- Researcher carries out statistical analyses in UCL safe haven
- Researcher shares analytic scripts, codes and EHR phenotypes on CALIBER portal
- Researcher prepares outputs (publications, conference abstracts)
- Researcher destruction of data

CPRD: Clinical Practice Research Datalink
EHR: Electronic health record
ISAC: Independent Scientific Advisory Committee

CALIBER cost model (v6 17 December 2018)