



CPRD Research Data Governance Feasibility Study Guidance Completing the Feasibility Study Application Form

This guidance document has been produced to help applicants complete the Feasibility Study Application Form. Applications not completed in accordance with this guidance will be returned as invalid.

Key Information:

Feasibility studies using CPRD data are not subject to a full scientific protocol review under the CPRD Research Data Governance (RDG) process, provided they meet the definition and scope of a feasibility study, and subject to the restrictions detailed below. Approval for feasibility studies may be requested by submitting the 'Feasibility Study Application Form' and any other required documentation to the RDG Secretariat (rdg@cprd.com).

- A feasibility study is a study where the intended purpose is to assess the feasibility of conducting a future study. This could include assessing the feasibility of a future observational study using CPRD data, or a prospective observational study involving enhanced data collection (questionnaires or bio-samples), or an interventional study (for e.g. pragmatic trial). Applicants may also use CPRD data to assess the feasibility of a future study that does not involve using CPRD data or services;
- Unlike a simple feasibility count, a feasibility study could include percentages (rather than just counts) to assess the distribution/prevalence of key events, exposures, and outcomes in CPRD data. However, a feasibility study approved under this route should not include any hypothesis testing or tests of significance;
- By submitting a feasibility study application, applicants are:
 - declaring that they have read and understood the guidance on completing a Feasibility Study Application Form;
 - confirming that the submitted feasibility study application, and any supporting documents, are accurate;
 - agreeing to abide by all contractual obligations in relation to access to CPRD data and any other linked data where applicable;
 - agreeing to publication of summary information, should the application be approved, in accordance with CPRD's Transparency Policy;
- The application form should be completed in Arial 10pt font only. Superscripts, subscripts, and footnotes should be avoided. The use of special characters and symbols should be avoided unless necessary. The use of graphs, charts, graphics, and images is not permitted;
- Each section of the application form should be completed in full. Applications with incomplete sections will be returned as invalid;
- Direct communication between applicants and reviewers is not permitted. All communication regarding applications should take place via the RDG Secretariat, and only individuals named in the application may make enquiries.

Restrictions

- Sample sizes for feasibility studies are limited to 50,000 patients. Where more than 50,000 patients meet the cohort definition criteria, a random sample of 50,000 will be provided to applicants. This restriction on sample size does not apply to multi-study licence holders or for clients who are commissioning CPRD to undertake the feasibility study on their behalf;
- Applicants will be able to request data based on up to 3 of the following eligibility criteria:
 - Cohort definition based on a single code list representing a diagnosis/prescription (code list to be provided by the applicant as an exported text file);
 - age-group (specified as an age range or threshold e.g. 40 years or younger at date of diagnosis/prescription or start of the study period etc.) and;
 - sex.
- Feasibility studies are permitted using the following data sources:
 - CPRD primary care data (CPRD GOLD and CPRD Aurum);

- ONS Death Registration Data;
 - HES Admitted Patient Care, Accident and Emergency, Outpatient, and Diagnostic Imaging Dataset;
 - Mental Health Services Dataset (MHDS);
 - CPRD Mother Baby Link;
 - Pregnancy Register;
 - Practice Level Index of Multiple Deprivation (Standard), and;
 - Patient Level Index of Multiple Deprivation.
- For clients without a multiuse licence for a specific dataset, data will be released under a Standard Dataset Agreement, with non-negotiable Terms and Conditions for Feasibility Studies.
 - Findings from approved feasibility studies can be shared with third parties and published in internal reports, peer-reviewed academic journals or included in conference presentations without the need for a full RDG protocol submission. However, findings should only be presented in the context of a feasibility study – for example, it would not be appropriate to present the findings in a publication as an incidence or prevalence study.

CPRD Transparency Policy

CPRD publish summary information of each feasibility study approved under its RDG process. For further information on CPRD’s Transparency Policy, please go to <https://cprd.com/protocol-list>.

Completing the Feasibility Study Application Form

SECTION A: GENERAL INFORMATION ABOUT THE PROPOSED FEASIBILITY STUDY
<p>Question 1: Study Title</p> <p>The study title should provide a concise overview of the aim/s of the feasibility study. For clarity, applicants must include “feasibility study” in their study title.</p> <p>The study title must be no longer than 255 characters, including spaces. Studies with titles more than 255 characters will be returned as invalid.</p>
<p>Question 2: Chief Investigator</p> <p>The Chief Investigator will take responsibility for ensuring that the research is undertaken with full adherence to CPRD guidelines. The CV of the Chief Investigator to determine their suitability for the study.</p> <p>The full name, job title, organisation name & e-mail address for correspondence of the Chief investigator must be included in the form.</p>
<p>Question 3: The Corresponding Applicant</p> <p>The name of a Corresponding Applicant, who will be the direct point of contact for the RDG Secretariat, must also be provided for each application. It is acceptable for the Chief Investigator to be the corresponding applicant.</p>
<p>Question 4: Other investigators/collaborators</p> <p>Anyone who will have access to CPRD data, or who will contribute meaningfully to the feasibility study must be named in the study application.</p> <p>The name and institution/organisation email address of all study investigators/collaborators must be stated on the application form. Each investigator/collaborator must also be copied in (cc, not bcc) when a new study is submitted for consideration.</p> <p>Copy and paste a new table for each additional investigator/collaborator.</p>

Curriculum Vitae (CV) requirements

All investigators/collaborators named on the study application form **must** submit a short curriculum vitae (CV) alongside submission of the study application if they have not already done so previously. CVs must be completed using the CV proforma provided on the CPRD website (<https://www.cprd.com/research-applications>).

CPRD hold copies of all CVs submitted as part of the application process. Investigators/collaborators who have submitted a CV will be provided with a CV number by the RDG Secretariat. This CV number should be used in any future applications. Investigators/collaborators should only submit another CV form if they wish to update the CV held on record by CPRD.

SECTION B: ACCESS TO THE DATA

Question 5: Data Access Arrangements

State the method that will be used to access the data for this study – a CPRD study-specific dataset agreement or a multi-study licence agreement. If an existing license is to be used please indicate the licence institution name and address.

Question 6: Site location of data

Provide the address of the site location/s where the data is being processed, stored and analysed.

SECTION C: INFORMATION ON REQUESTED DATA

Primary care data collected by CPRD can be linked to a number of other patient level datasets, (including Hospital Episode Statistics, ONS Death Registration Data, and the Mental Health Services Data Set...) and is only available for English practices that have consented to participate in the linkage scheme. These linkages cover approximately 75% of the contributing CPRD English practices, or roughly 58% of contributing CPRD UK practices.

Question 7: Primary Care data

Vision and EMIS are different clinical software systems used by general practices in the United Kingdom primary care setting. CPRD has historically collected data from Vision primary care practices, which is referred to as the GOLD primary care data. More recently, CPRD has been able to release data collected via the EMIS software system under the Aurum primary care data.

Question 8: Requests to access linked data

The following linkages are permitted for feasibility studies:

- HES Admitted Patient Care, Accident and Emergency, Outpatient, and Diagnostic Imaging Dataset;
- Mental Health Services Dataset (MHDS);
- Pregnancy Register;
- Practice Level Index of Multiple Deprivation (Standard), and;
- Patient Level Index of Multiple Deprivation.

Question 9: Requesting CPRD to extract the primary care data

Applicants may request that CPRD perform the data extraction. If answering 'yes', applicants will need to complete a Data Specification Form following approval of the Feasibility Study.

Question 10: Commissioning CPRD to conduct the Feasibility Study

Applicants may commission CPRD to conduct the Feasibility Study. Please note that the restriction on sample size does not apply to clients who are commissioning CPRD to undertake the feasibility study on their behalf.

Applicants commissioning CPRD to conduct the Feasibility Study must have spoken to a member of the CPRD Observational Research team beforehand. Please email rdg@cprd.com for further information.

Question 11: Patient identifiers

Investigators must state whether any person named in the study has access to the data in a patient identifiable form, or any associated identifiable patient index. If yes, details must be provided.

If the answer to this question is 'Yes', applicants **must** provide a re-identification and risk management plan as an appendix and refer to it here.

SECTION D: FEASIBILITY STUDY INFORMATION

A. Lay Summary (Max. 250 words)

Please provide a succinct overview of your proposed research in non-technical language.

The lay summary is published on the CPRD website for the benefit of patients and the public, to inform them of how CPRD data are being used and to what benefit.

Applicants proposing to use CPRD data for a feasibility study should clearly state the aim of the proposed study for which purpose the feasibility study is being undertaken (e.g. 'the aim of this study is to assess the feasibility of a cohort study to be undertaken in CPRD data investigating the association between drug X and outcome Y'; or 'the aim of this study is to assess the feasibility of a randomised controlled trial assessing comparative effectiveness of different dosing schedules of treatment X within the UK primary care setting').

The lay summary should not include any technical details, such as study design or statistical methods, should avoid jargon, and must be capable of being understood by a member of the public without a scientific or medical background. Abbreviations should be avoided.

The lay summary should provide an overview of the research without the need to refer to the technical summary.

Studies with lay summaries over 250 words will be returned as invalid.

B. Technical Summary (Max. 300 words)

The technical summary is primarily written for other researchers and clinicians. There should be enough technical detail to allow another researcher to obtain a clear idea of your study aim and methods.

The technical summary should provide a succinct overview of the overarching study aim and objectives, primary exposure(s), and outcome(s), if relevant, study design, and methods including the main analyses to be conducted.

The technical summary should also specify how linked datasets will be used, for example "Hospital Episode Statistics (HES) data will be used to determine hospitalisations."

Studies with technical summaries in excess of 300 words will be returned as invalid.

C. Outcomes to be Measured

This section should clearly list the key outcome variables, separated by semicolons. For example: "Ischaemic stroke; Composite of all bleeding; Major bleeding; Gastrointestinal bleeding; Clinically significant non-major bleeding; Myocardial infarction; All-cause mortality; Intracranial haemorrhage."

This section should not include statements relating to the study aims and objectives.

D. Information on the Study Population

It is important to ensure that the application clearly defines the study population. The following areas listed below should be addressed as relevant:

- a) Describe the source/target population:
 - cohort definition (age, sex, key condition/exposure of interest)
 - whether only permanently registered acceptable patients will be included
 - whether only up-to-standard follow-up will be considered;
 - state the recruitment period and state the definition of the start and end of follow-up for patients, including whether the CPRD death date should be used in defining the end of follow-up.
- b) Describe the study population in terms of inclusions, exclusions, and the data used for each (clinical, referral, test, therapy, immunisation). Reference should be made to provisional code lists for inclusion & exclusions specified;
- c) Provide any minimum requirements for previous follow-up time;
- d) Information on the exposure window(s) of interest, where appropriate, defining clearly time which will be considered "exposed" or "non-exposed";
- e) For studies requiring linked data, please make clear the restrictions imposed by patients' eligibility for linkage and the coverage period for each linked dataset requested. Further information on patient's linkage eligibility and the coverage periods of linked data sources may be requested from the CPRD (enquiries@cprd.com).
- f) For studies requiring linked data, please make clear whether the study population should be defined among all eligible patients in CPRD or only among patients eligible for linkage to your linked data source of interest.

Appendix 1: Examples of feasibility studies that can be approved under this route

EXAMPLE 1

Future Study: To investigate the association between LABA/LAMA prescribing and rate of decline of lung function in patients with moderate to very severe COPD (GOLD stages II to IV).

Feasibility study: To explore definitions of chronic obstructive lung disease (COPD) severity based on diagnosis, treatment and/or spirometry data

Identify the target population

- Number of patients with a diagnosis of COPD during the study period
- Number of patients with spirometry data recorded
- Number with valid spirometry data recorded
- Proportion of patients with FEV1/FVC < 70
- Patients treated with COPD therapy
- Patients with a diagnosis of asthma

Implement algorithm for classifying COPD severity (GOLD stages II to IV) based on treatment and spirometry data.

- Distinguish pre and post-bronchodilator spirometry data
- Implement GOLD categories using data using spirometry and other data recorded in primary care.

EXAMPLE 2

Future Study: Prediction of exacerbation onset in Chronic Obstructive Pulmonary Disease (COPD) patients

Feasibility Study: To examine the recording of COPD exacerbations in primary care and understand the extent to which sub-group analyses may be possible.

Identify the target population

- Number of patients with a diagnosis of COPD during the study period
- Numbers eligible for linkage, number with at least 12 months of prior UTS follow-up
- Numbers with current/ex-smoker status,
- Proportion of patients with FEV1/FVC < 70 at diagnosis and 6-month post diagnosis.
- Numbers with a record of asthma and COPD

Develop algorithm for identifying COPD exacerbations

Proportion of patients treated with combined therapy (dual, triple therapy) for at least 3 months/6 months

- Proportion of patients receiving add-on therapy and cumulative dose
- Proportion of patients switching therapy, type of therapy, dosage and duration
- Proportion of patients with at least 2 exacerbations during dual therapy
- Proportion of patients with at least 2 exacerbations during triple therapy

Frequency of exacerbations among patients with COPD

Proportion of patients treated with combined therapy (dual, triple therapy) for at least 3 months/ 6 months who have a record of an exacerbation

Identify COPD exacerbations in the hospital setting.

EXAMPLE 3

Future Study: To investigate the association between COPD medication adherence and resource utilisation

Feasibility Study: To assess the completeness of recording of COPD prescribing data and explore methods for estimating treatment adherence in light of the data that are available.

Identify the target population

- Number of patients with a diagnosis of COPD during the study period
- Numbers eligible for linkage to Hospital Episode statistic
- Numbers with at least 12 months of Up-to-standard (UTS) follow-up prior to COPD diagnosis
- Numbers with at least 24 months of follow-up following COPD diagnosis
- Numbers initiating maintenance medication during the first 365 days of the 24-month post-index period (have at least one prescription for COPD maintenance - long-acting muscarinic antagonists [LAMA] and inhaled corticosteroid-long-acting beta agonist [LABA] fixed-dose combinations)
- Evaluate dose of maintenance therapy by treatment type

Explore approaches for measuring COPD treatment adherence

- Proportion of days covered (PDC)
- Medication Possession Ratio (MPR)

Examine the characteristics of adherent/non-adherent

- Proportion adherent/non-adherent
- Mean Proportion of days covered/ MPR
- Baseline characteristics of adherent and non-adherent groups (Mean, median, and interquartile range)