Guide for applicants for study data access

Document Ref.	IG-P30 Data Access Guide
Version:	V01.0
Dated:	1 October 2021
Document Author:	IG Lead
Document Owner:	IG Lead
Classification	External

1 Purpose

This guide is for both internal and external applications to use the IBD Registry's data for analysis. This Guide contains useful information for the applicant and outlines the application assessment process from the perspective of the applicant.

2 Audience for this guide

This guide is aimed at individuals within the IBD Registry or external organisations interested in applying to use IBD Registry data in reports, publications, analysis or research. All potential applicants should read this guide thoroughly before application.

3 The data

3.1 Types of data

The IBD Registry holds different types of data that you can apply for. Before completing the Data Request Form, please familiarise yourself with these definitions and decide what sort of data is required for your project.

3.1.1 Aggregated data / statistics

3.1.1.1 Aggregated anonymised information (non-personal data)

Aggregate information is anonymous information that is grouped together to generate statistics. Line-level data is not available; instead aggregate data collects, combines and communicates details in terms of totals or summary. We will maintain patient privacy by small numbers suppression.

This type of information provision does not require a data sharing agreement, but it does require an assessment that the processing has been lawful and compliant. This assessment will be part of our standard assessment and decision-making process.

3.1.2 Line level-data

3.1.2.1 De-identified data (personal data)

Names and obvious identifiers removed. However, still has a unique identifier attached to reach record. This means that de-identified data has the potential to be linked with another dataset. Where this is the case, it will be controlled with a data sharing arrangement. If de-identified data is requested, but linkage is not required, the IBD Registry will investigate as to whether anonymised data would be more appropriate.

3.1.2.2 Anonymised data (risk of being personal data)

Fully anonymised line-level data cannot be linked to any other datasets as all identifiers are removed. This may means that not all data fields are available as anonymised data.

Releasing line-level data carries a high risk that the data may become identifiable (personal data) and so similarly strict conditions and controls would apply when sharing anonymised data as when de-identified data. Requests for anonymised data will be considered very carefully to ensure that the data carries as low as risk as possible of being reidentified.

3.2 Datasets

The IBD Registry holds several datasets that are available to apply for:

- Main Registry data - available as aggregated anonymised
- COVID-19 dataset – for research into COVID-19 and IBD only; available as line-level linkable or aggregated anonymised
- From 2022 onwards – Main Registry data will be available as line-level linkable or aggregated anonymised

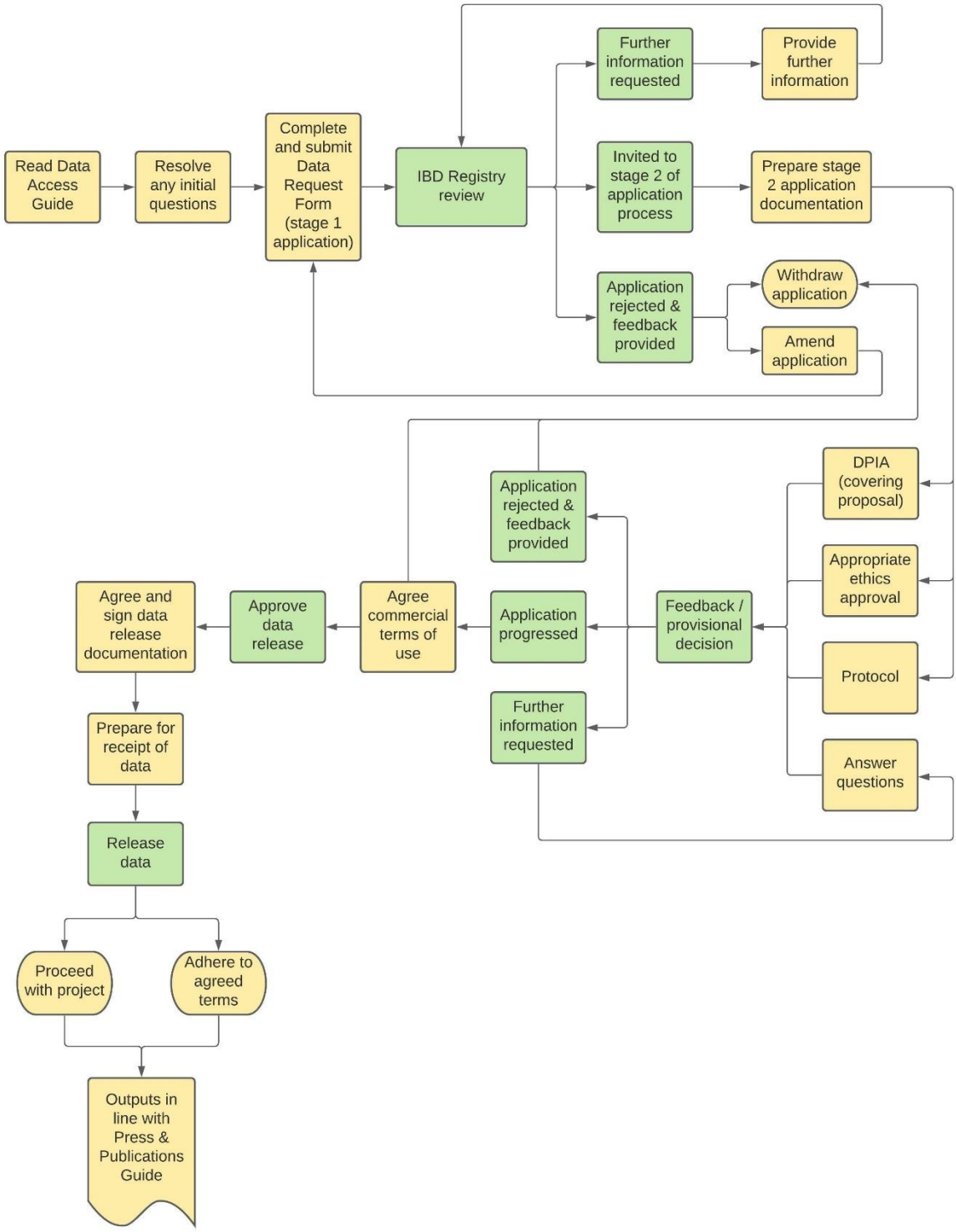
3.3 Variables

When requesting data, you will need to provide the IBD Registry with a list of the statistics or variables that you are interested in.

A list of the data collected by the Registry is set out in our Data Submission Framework. A summary version of this is listed on our website, and the full specification is available on request from the Registry.

4 Application process

The process for applying for data is summarised in the flowchart below. The duration and complexity of each stage will depend on the type of project you are undertaking and the data you are requesting.



4.1 Data request form

4.1.1 Application

The applicant decides whether they want to apply for aggregated data or line-level data. The relevant stage 1 Data Request Form should then be completed by applicant and submitted to the UK IBD Registry (IBDR) (see accompanying forms).

This application form is assessed for its feasibility by the IBDR Data Access Group. It is assessed from the following perspectives:

- Information Governance
- Operations
- Clinical/Academic
- Patient/Public

In order to be able to assess your request, the IBD Registry needs to understand the following points.

- Does your project constitute research according to the HRA?
- What organisation is the project sponsor?
- If so, what is the current status of your ethics?
- Do you require anonymised aggregated data (e.g. statistics) or line-level data?
- If line-level, do you intend to link this data to one or more other datasets

This will have an effect on what data we may able to release to you and what further checks will be required if your application is approved to move to the second stage. The application form will ensure that we have the information that we require in relation to the above points.

4.1.2 Assessment

If the IBDR Data Access Group are in principle satisfied with the stage 1 application, the IBDR invites the applicant to submit further information to complete their application – this will include a data protection impact assessment (DPIA) if required, the full protocol, confirmation of ethics approvals/requirements, as well as answers to additional queries that the IBD Registry team may have.

The other possible outcome is that the stage 1 application is rejected or that the applicant is requested to make changes before resubmission.

The IBDR's Data Access Group assesses the full application submission. If the group is satisfied with the application materials (scientific and methodological merits, feasibility, IG compliance, data flows and risk), then the access request will be approved. If personal data is to be shared – or data that is at risk of becoming personal data - the IBDR completes a data sharing agreement (DSA) with external applicants prior to releasing data. Note: any high risk applications will be referred for scrutiny at a higher level.

5 Additional documentation

5.1 Legalities and ethics

5.1.1 Legal & ethical permissions required to process UK IBD Registry data

The UK IBD Registry has to ensure that all external researchers or clinicians requesting Registry data for a study have the necessary permissions and approvals in place prior to receiving this data.

This section details what these permissions are and the evidence you must submit alongside your data request form and data protection impact assessment (DPIA).

5.1.2 Audit/service evaluation or research?

Defining the purpose for which you want to use UK IBD Registry data is key, because this determines what permissions you need to have in place to receive the data.

The UK IBD Registry shares data for two study purposes:

(i) **Audit and service evaluation/quality improvement** – audit and service evaluation/quality improvement projects seek, respectively, to measure clinical practice against published guidelines and evaluate and improve the quality of the service provided. These projects may be comparative, but they do not seek to establish generalisable conclusions

(ii) **Research** – research projects aim to generate new knowledge that can be generalised.

Key resources to help you categorise your project are:

- The Health Research Authority's (HRA) website and tool (http://www.hra-decisiontools.org.uk/research/).
- Twycross & Shorten's paper: 'Service evaluation, audit and research: what is the difference?'

If you believe your project is audit, you must complete the HRA's audit and research tool and submit the output alongside your data application.

The IBD Registry holds some data that is permissioned for use in research, and some data that is not. It is crucial for us to know what the usage of data will be in order for us to provide you with the correct cut of data.

5.1.3 Common law duty of confidentiality

All projects that use patient data, whether research or not, must be able to provide evidence that they have satisfied the common law duty of confidentiality. This duty states that beyond its use for the direct care or treatment of patients, permission must be given for all use of a patient's confidential information.

Permission can be granted in two forms:

(i) Patient consent for use beyond their direct care e.g. audit and research. For consent to be valid the patient must have been informed of how the data will be used and why, have given a positive and unambiguous statement confirming their consent (ideally in writing) and have the right to withdraw at any time. Projects that rely on patient consent to receive and process confidential information, e.g. clinical trials, must provide a copy of the patient information sheet and the template consent form.

(ii) Secondary use permissions under s251 of the NHS Act (2006) – where it is not possible or feasible to collect patient consent for a project, applicants may use permissions granted under s251 of the NHS Act (2006) (s251 permissions) to collect patient data. s251 permissions are granted by the HRA's Confidentiality Advisory Group (HRA CAG) –see their guidance.

5.1.4 Applying for data for your study with the right permissions

The IBD Registry currently receives s251 support (ref: 18/CAG/0131). However, this is under non-research permissions. This means that if you are applying for data from us for a research project we are not permitted to share s251 data with you.

The IBD Registry also holds consented patient data. However, a participant can provide consent to be part of the Registry, but withhold consent for their data to be used for research.

For these reasons, when you apply for line-level data from the Registry, you are required to state whether or not the data will be used for research or non-research and to provide proof.

5.1.5 Research ethics

The UK IBD Registry does not currently hold research ethics permissions for the UK IBD Registry to populate and maintain a research database. (We are currently applying for ethics for the purposes of providing high quality data for studies).

Applicants applying to the UK IBD Registry for data for research purposes must be able to either:

(i) Provide evidence that their research project has been granted research ethics permissions by a national or local research ethics committee or

(ii) Provide evidence that their research project does not require research ethics approval. This can be determined using the HRA's ethics tool (http://www.hra-decisiontools.org.uk/ethics/).

The UK IBD Registry may still require an internal ethics review, depending on the nature of the application.

Applicants should be aware that their institutions may have specific rules regarding what kinds of study require ethical approval and all applicants should consult their institutions' policies.

For projects that require ethics approval, no data will be released until the IBD Registry receives proof that this approval is in place.

5.2 DPIA

In order for data to be provided, a relevant Data Protection Impact Assessment (DPIA) must be in place and approved by the IBD Registry. For frequent sharing scenarios, a generic DPIA will already be in place. For scenarios where an element of the project is not covered by an existing DPIA (e.g. where the applicant is requesting that data is released into an analysis environment that has not been previously approved by the IBD Registry), a DPIA template will be provided to the applicant to complete. This will then be reviewed by the IBD Registry. Where possible, the IBD Registry will work with the applicant to provide guidance as to what measures may need to be put in place to keep the risk level down to an acceptable level. If the risk associated with the project is considered too great, the data application may be declined.

5.3 Protocol

In order to assess the study effectively, where line-level data is being requested the IBD Registry require to see a copy of the protocol. It is not a requirement to submit this with the initial data request form, however it will be required before a final decision can be made. We suggest that applicants submit the draft protocol to us *before* submitting it for ethical approval, as this will allow you to consider any suggestions or feedback that the IBD Registry's Data Access Group may have.

6 Release of data

Plain English summaries of each successful application are listed on the IBDR website.

6.1 Terms of use

Only de-identified (pseudonymised) data are released for study or research purposes. For line-level data, a data sharing agreement will be put in place and signed by both parties prior to data release. This will outline the full terms of use of the data.

There must be no attempt to re-identify the data subjects and protection of participant identity is required throughout. Data may not be linked unless pre-agreed.

Included in the terms of use will be the requirement to appropriately acknowledge or credit the IBD Registry in any reports of publications. Recipients will be provided with a Press & Publications Guide to ensure the requirements are clear.

6.2 Storage of data

Before releasing data to an external organisation, the IBD Registry must be satisfied that an appropriately secure environment is in place. We may decide that we would like to release data to an external organisation, but be unable to because the recipient does not have an appropriately secure environment in place. The potential data recipient may apply to the IBD Registry for use of its Trusted Research Environment (TRE) for the specified project. This is the preferred scenario for the IBD Registry as the security of this environment is already covered in an existing DPIA. In these instances, the terms of its use will be negotiated and a separate Data Processing Agreement will be established.

7 Fees and charges

Whilst the IBD Registry does not charge for the data itself, it must recoup costs for the preparation of the data, the governance assessments and resulting agreements, plus the provision of the analysis environment, on behalf of third parties. Therefore, an appropriate fee will be charged to third parties for information, studies or research.

Fees are set on a case-by-case basis according to the following principles:

- The fee charged will be based on the time and resource provision it takes to prepare the data, as well as the type of organisation requesting it.
- An element of the fee may also include a charge for the investment over time of building up the Registry as a resource
- Each data sharing or access agreement must be cost-neutral or cost-positive for the Registry.
- The fee will take into account any fixed costs as well as variable costs around the amount of data that has been requested and therefore the work related to extraction and linkage (if applicable). Whether aggregated data only or patient-level data has been requested will have an impact on the cost.
- The fixed cost element will vary depending on the type of organisation making the request, such as:
    - NHS/Public Sector
    - Academia/Charities
    - Commercial/Industry

8 Next steps

The documents referred to in this guide that you need to progress your application are all available on the IBD Registry website.

If you have any queries about your application, please email analysis@ibdregistry.org.uk

The IBD Registry has enabling research as part of its core purpose. We hope you find this guide helpful. We are keen to help you progress your study.

9 Appendix: document information

9.1 Revision History

Version	Date approved	Revision Author/Group	Summary of Changes
01	1 Oct 2021	IG Lead	Original document

9.2 Review and Approval

Changes to this document must be reviewed by the following:

Reviewer Role	Review/approval
SIRO	Review & approval

Final approval before release is required from the SIRO.

9.3 Document Status

This is a controlled document. Whilst this document may be printed, the electronic version is the controlled copy. Any printed copies of the document are not controlled.

This is a **mandatory** protocol document. Staff are required to follow the guidelines set out in this document. There may be occasions when it makes sense to deviate from these guidelines. In this cases, written approval from the document owner must be sought, and the deviation recorded together with the reasons.