IBD Registry

Data Submission Specification 2022_L

Document Ref.	2022_L Data Submission Specification
Version:	V03.0.0
Dated:	31 Dec 2022
Document Author:	LD SG JB
Document Owner:	SG
Classification	PUBLIC

1 Introduction

The IBD Registry is the curated database of IBD data collected in the UK, primarily from IBD clinical teams but also from patients.

The ethos of the Registry is to be open to the widest possible participation by clinical teams, and wherever possible to avoid or minimise additional data entry work by clinical teams. We describe this approach as 'system agnostic' and 're-using captured point of care data'.

Data Submission into the Registry can be from any electronic clinical system (ranging from a Trust-wide eHR system to a basic local clinical system) that can produce a Registry-compliant set of files for submission.

The **Dataset Definition** describes the standard for the Registry dataset. Any clinical system that can extract the identified data in a set of files in this format will be able to upload data into the Registry.

The detailed **Submission Framework** describes the standard for submitting data to the Registry. Any clinical system with data in a set of Registry submission-compliant extract files will be able to upload data into the Registry.

2 Audience

This document is designed for use by clinical teams and their supporting hospital IT teams. Third-party developers may also use this document as an introduction to our dataset – a full dataset definition is available upon request.

3 Participating in the Registry Project

The IBD Registry encourages all hospitals in the UK to join the Project by collecting and submitting the Registry data on their IBD patients.

4 Joining the Registry

Information on the steps required to participate in the IBD Registry are set out in the Information Pack available by email from our Support team . If you would like to be sent the detailed Submission Framework (2022_L), please email the same team.

5 Registry, the IBD Audit and IQICC

One of the missions of the IBD Registry is to facilitate the continued collection of data for national clinical quality improvement in IBD. The Royal College of Physicians' IBD Audit team and the Registry worked together in 2016-17 to transition the collection of the IBD Biologics Audit data to ensure the IBD Registry's data framework, together with compliant data capture tools (such as the Registry WebTool), were able to capture the data required for this audit. During 2022, the BSG IBD Section led on the national revision of the KPIs used in this original audit, resulting in a wider but simpler set of indicators for 2023 as part of the new Improving Quality in Crohn's and Colitis (IQICC) programme. The tools will be updated to reflect data capture for these changes.

Users of these Registry-compliant data capture tools are able to submit their data for the Biologics Audit (currently) and IQICC (later in 2023) along with their Registry data. The

specification for this is incorporated into the Registry submission dataset (available upon request).

We suggest third-party software developers test their data submission files for accuracy before using them to submit real data; access can be arranged to a test IBD system by contacting support@ibdregistry.org.uk.

Any queries on this should be directed to support@ibdregistry.org.uk in the first instance.

6 The Data Submission Process

6.1 Submission schedule

The Registry works to a schedule of quarterly data submissions in the months of January, April, July and October each year. Submission dates are on our website.

Hospitals wishing to submit live data to the Registry are encouraged to organise their registrations and approvals and then make a first submission some time before a quarterly update schedule to familiarise themselves with the process and test that their data can be successfully submitted. Please contact support@ibdregistry.org.uk if you wish to arrange this.

6.2 Data submission platform

The IBD Registry has its own data access portal and data handling platform hosted at AIMES, from July 2022. Previously, our data submission was with NHS Digital's Clinical Audit Platform.

Anyone that submitted data prior to July 2022 that has not yet registered with our data access portal will need to do this before they are able to submit data.

6.2.1 Registering for our Data Access Portal

Hospitals need to register the details of individuals who need access to the data access portal via our online form before they are able to submit their data.

6.2.2 Hospital Codes

The IBD Registry uses IBD audit codes to identify the source of any data submission.

The code needs to be present on each row of each extract file. When setting up your local tool, you will be provided with this code by the IBD Registry.

When you sign up to use our data access portal, you will also be provided with a very simple "provenance file" consisting of just one line of data. This file will need to be uploaded along with your data every time you submit data to the Registry.

6.2.3 Using the extract filter tool

Before you upload your files to the data access platform, you must use the extract filter tool first. This tool processes your existing extract files and produces new files for

upload containing only data that is permissioned to flow to the Registry. You import your extract files, a file of version 4 consents (sent to you by us), a national data opt-out file and your provenance file, and the tool produces your output files that you can then upload to the platform.

7   The Data Submission Format

This section is primarily for the use of hospitals using third-party databases to submit data to the Registry. The Registry Web Tool, plus a number of compliant third-party systems) have this data submission format integrated into the software and the Registry Export Process. To obtain a copy of our full (detailed) Dataset Definition (2022_L) for implementing in your system, please contact us at support@ibdregistry.org.uk.

Data must be exported from your database in csv format in order to be compatible with the extract filter tool (and the data access portal); no other format can be accepted. Where data items have multi-select coding options, the codes must be separated by a semi-colon (;). All data fields must be provided, even when the value is unknown; when a data item is not available an empty value must be provided.

Each submission should contain all relevant Registry data including any data previously submitted.

The Registry has adopted this approach to minimise any local complications in tracking which data has previously been submitted and situations where previous data might subsequently have been corrected.

There is no need to filter which records go into the extract – all patients' records can go into the extract files. At present the Registry has been granted s251 approval under the NHS Act (2006) to collect patient data from hospitals without needing formal consent; however, patients have the right to 'opt-out'. The IBD Registry's extract filter tool can process the extracted files, consent files (provided by the Registry) as well as national data opt-out file provided by the Trust (once obtained from NHS Digital) and produce a set of files for upload which only contain records permissioned to flow to the Registry. The current s251 exemption runs until the end of 2027.

(NB. Third party software developers should ensure they notify support@ibdregistry.org.uk that they wish to be added to the list of developers to be notified of any changes to the Registry dataset or Data Submission Framework.)

7.1   Data Submission Flow

Overview of Registry data-flow and pseudonymisation:
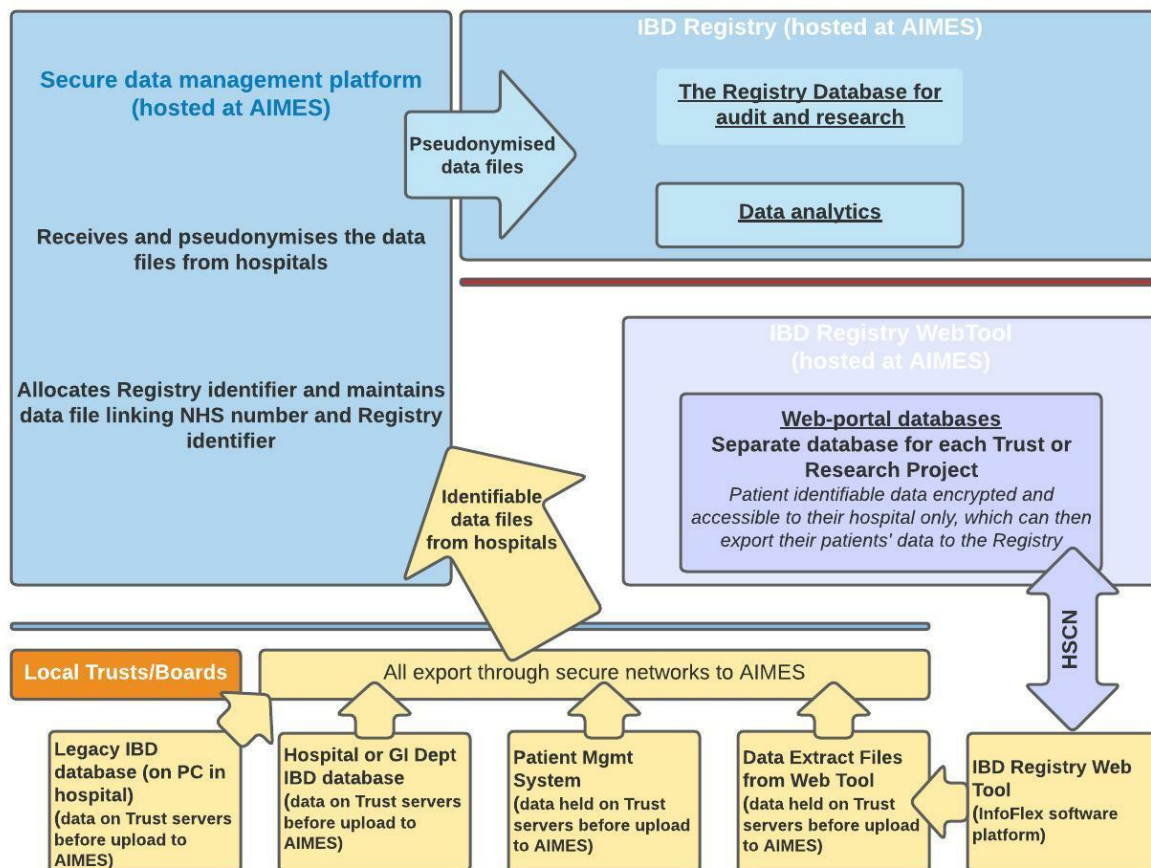
8 Validation Rules

In order to be registered, a submission needs to have both a provenance datafile and a patient datafile.

8.1 Provenance file

When a set of data files are submitted to the IBD Registry, the system identifies where they are from by the included Provenance file. This is a single line file included with each submission. The key identifier in this file is the site's IBD Audit code.

8.2 Patient file

When data files are submitted to the IBD Registry, the system will validate the files according to the following rules:

Records will be rejected if a required item is missing or has an invalid format or code.

Records which have an invalid format or code for a data item that is not a 'Required item' will be accepted but the invalid item will not be added to the Registry data.

We will contact you to inform you of this.

9 Data Items

A summary list of our dataset items is available to download from
https://ibdregistry.org.uk/data-submission-framework/

Full Dataset Definition for Implementors

If you are a Trust or Health Board, and would like a copy of our full (detailed) Dataset Definition (2022_L) for implementing in your system, please contact us at support@ibdregistry.org.uk.

10 Version Control

Owner	Version	Date	Status
Richard Driscoll	2015_A	12 Feb 2015	First draft
Richard Driscoll	2015_B	13 Feb 2015	Second draft for comment
Richard Driscoll	2015_C	19 Feb 2015	With comments from SC and CM
Richard Driscoll	2015_D	17 Mar 2015	Minor typographical errors corrected, data items to be included in csv files clarified and missing data validation statements added to table on pp 9-14
Richard Driscoll	2015_E	19 May 2015	Addition of Contact Type as required data item and updates to submission process information
Richard Driscoll	2015_F	8 July 2015	Additional information requested by users of the document. Addition of the IBD Audit Biologics Dataset.
Richard Driscoll	2015_G	24 Aug 2015	Addition of the IBD Audit Biologics Dataset. Clearer specification of three levels of data submission. Extra codes for Contact Type
Richard Driscoll	2015_H	20 Sept 2015	Corrections to Biologics Dataset – removal of duplicate item (perianal disease) and submission filename corrected to BiologicsEvent.
Richard Driscoll	2015_H.1	12 Nov 2015	Wording changes to description of the levels of data that can be submitted. Updated data flow visual.
Stephen Grainger	2019_J.1	13 July 2019	Revised governance section. Additions to drug names/codes; additional consent field and ethnicity; other additions to modernise dataset; typographical errors corrected
Liz Dobson	2019_J.2.0	01 June 2021	Description updated to reference planned 2020-21 transition away from NHS Digital. Formatting and layout upgraded.
Liz Dobson	2019_J.2.0.1	11 June 2021	Minor formatting corrections
Liz Dobson	2022_L 3.0.0	31 Dec 2022	Updates relating on dataset update from 2019_J to 2022_L. (Includes changes for 2021_K also.) Removal of a lot of dataset information and replaced with a statement that people can apply to receive the full dataset.

10.1 Version Control Numbering Rules

Any Version that has changes to the dataset that must be implemented will be labelled with the next letter in the alphabetical sequence (the letter I is omitted to avoid confusion with 1).

Any changes that are for clarification of the text will be labelled with a number after the appropriate letter. (e.g H.1)