DocTIS Research Data and Code Repositories

One of the key outcomes of the DocTIS project is the creation of a valuable collection of molecular datasets, metadata and analytical tools that can support future research in immune-mediated inflammatory diseases (IMIDs). Several resources are already publicly available, while others will be released following publication of the corresponding scientific manuscripts.

To support transparency, reproducibility and long-term reuse of project results, these datasets and analytical resources have been deposited in public or controlled-access repositories.

The resources described below are the result of collaborative work carried out across the DocTIS consortium, involving contributions from the Vall d'Hebron Research Institute (VHIR), the National Centre for Genomic Analysis (CNAG), Cardiff University, the August Pi i Sunyer Biomedical Research Institute (IDIBAPS), Charité – Universitätsmedizin Berlin, the University of Verona, IMIDomics Inc., HudsonAlpha Institute for Biotechnology, Linköping University, Karolinska Institutet and Zabala Innovation.

1 - Publicly Available Resources:

Preprocessed scRNA-sequencing data and metadata for psoriatic arthritis (PsA) patients treated with anti-IL17 therapy

This dataset contains preprocessed single-cell RNA sequencing (scRNA-seq) data and associated metadata from PsA patients classified as responders or non-responders to anti-IL17 treatment at baseline, together with healthy controls.

- Repository: Figshare
- Contact person: Samuel Schäfer (Linköping University)
- Associated publication: *scDrugPrio: a framework for the analysis of single-cell transcriptomics to address multiple problems in precision medicine in immune-mediated inflammatory diseases* (Genome Medicine)
- Status: Data available.

Preprocessed scRNA-sequencing data and metadata for psoriatic arthritis (PsA) patients treated with anti-TNF therapy

This dataset contains preprocessed scRNA-seq data and metadata from PsA patients classified as responders or non-responders to anti-TNF treatment at baseline, together with healthy controls.

- Repository: Figshare
- Contact person: Samuel Schäfer (Linköping University)
- Associated publication: *scDrugPrio: a framework for the analysis of single-cell transcriptomics to address multiple problems in precision medicine in immune-mediated inflammatory diseases* (Genome Medicine)

- Status: Data available.

Preprocessed scRNA-sequencing data and metadata for all IMIDs at baseline and healthy controls

This resource provides harmonised and preprocessed scRNA-seq data from patients across all IMIDs included in DocTIS, together with healthy controls, enabling cross-disease analyses of immune-cell states and inflammatory mechanisms.

- Repository: Zenodo
- Contact person: Holger Heyn (CNAG)
- Associated publication: *Interpretable inflammation landscape of circulating immune cells* (Nature Medicine)
- Status: Data available.

Code repository for the paper "Interpretable Inflammation Landscape of Circulating Immune Cells"

The repository contains the complete analytical workflows and code used to generate the results presented in the publication, facilitating reproducibility and reuse.

- Repository: GitHub
- Contact person: Holger Heyn (CNAG)
- Associated publication: *Interpretable inflammation landscape of circulating immune cells* (Nature Medicine)
- Status: Code available.

2 - Resources Awaiting Release Following Publication:

Raw scRNA-sequencing data and metadata for all IMIDs at baseline and healthy controls

This dataset contains the raw single-cell transcriptomic data and associated metadata generated across all IMIDs included in the project.

- Repository: European Genome-Phenome Archive (EGA)
- Contact person: Sara Marsal (VHIR)
- Associated publication: *Interpretable inflammation landscape of circulating immune cells* (Nature Medicine)
- Status: Data uploaded.

The dataset has been uploaded to the repository but has not yet been formally released.. Following release, access will be provided through EGA under controlled-access procedures.

Raw scRNA-sequencing data and metadata for all IMIDs at follow-up

This resource contains longitudinal scRNA-seq profiles and metadata generated during patient follow-up across all IMIDs included in the project.

- Repository: European Genome-Phenome Archive (EGA)
- Contact person: Sara Marsal (VHIR)
- Status: Data uploaded.

The dataset has been uploaded to the repository but has not yet been formally submitted for release. Submission is planned following acceptance of the associated publication, currently expected by the end of 2026. Once submitted, access will be available through EGA under controlled-access procedures upon request.

Raw bulk RNA-Seq data and metadata for all IMIDs at baseline and follow-up

This dataset contains raw bulk transcriptomic sequencing data and associated metadata generated from patients across all IMIDs, covering both baseline and follow-up time points.

- Repository: European Genome-Phenome Archive (EGA)
- Contact person: Sara Marsal (VHIR)
- Status: Data uploaded.

The dataset has been uploaded to the repository but has not yet been formally submitted for release. Submission is expected following acceptance of the associated publication, currently anticipated by the end of 2026. Once submitted, access will be provided through EGA under controlled-access procedures.

Preprocessed bulk RNA-Seq and additional scRNA-seq datasets

This resource includes preprocessed bulk RNA-seq data for all IMIDs at baseline and follow-up, healthy controls, and scRNA-seq data generated from rheumatoid arthritis donors at baseline and follow-up.

- Repository: Zenodo
- Contact person: Sara Marsal (VHIR)
- Status: Data uploaded.

The dataset has been uploaded and will be released following acceptance of the associated publication, currently expected by the end of 2026. Once released, the data will be available under request.

Code repository supporting the combinatorial therapy analyses

This repository contains the analytical code used in the combinatorial therapy study described in the manuscript by Martínez-Mateu et al., currently under review.

- Repository: GitHub
- Contact person: Sergio Martínez (IMIDomics)
- Status: Code uploaded.

The code has already been uploaded and will be publicly released once the associated publication has been accepted.

Together, these resources represent one of the most comprehensive molecular and computational collections generated within the DocTIS project, encompassing single-cell and bulk transcriptomic datasets, metadata and analytical workflows across multiple immune-mediated inflammatory diseases. By making these resources available to the scientific community, the consortium aims to facilitate further research into disease mechanisms, treatment response and precision medicine approaches in IMIDs.