Personalised Medicine by using an Advanced Point-of-Care Tool for Stratified Treatment in High Risk Cardiovascular Patients (Grant Agreement No 101095432)

D1.5 Data Management Plan

Date: 2024-02-29

Version 1.0

Published by the PoCCardio Consortium

Dissemination Level: PU - Public

Co-funded by the European Union's Horizon Europe Framework Programme for Research and Innovation under Grant Agreement No 101095432

Document control page

Document file: D1.5 Data Management Plan
Document version: 1.0
Document owner: **IN-JET**

Work package: WP1 Project management
Deliverable type: R - Document, report

Document status: ☒ Approved by the document owner for internal review
☒ Approved for submission to the EC

Document history:

Version	Author(s)	Date	Summary of changes made
0.1	Trine F. Sørensen (IN-JET)	2024-01-26	Table of contents, initial input to chapters 1-4, 6
0.2	Jesper Thestrup (IN-JET)	2024-02-21	Chapter 6 added
0.3	Heiko Bugger (MUG), Barbara Luegger (MUG), Harald Sourij (MUG)	2024-02-26	Fair principles of data management in the clinical trial. Minor changes to chapter 5 (FAIR Data)
0.4	Trine F. Sørensen (IN-JET)	2024-02-27	Final editing. Ready for internal review
0.5	Trine F. Sørensen (IN-JET)	2024-02-29	Internal comments and corrections added
1.0	Trine F. Sørensen (IN-JET)	2024-02-29	Final version submitted to the European Commission

Internal review history:

Reviewed by	Date	Summary of comments
Mieke Adriaens (UGent)	2024-02-28	Minor corrections
Andrea Berghold (MUG)	2024-02-28	Minor corrections

Index:

1 Executive summary

The DMP will be at the core of the research data management in the PoCCardio project. The DMP follows the guiding principle of “as open as possible, as closed as necessary” for all data produced and used during as well as after the end of the project. The present deliverable is the first version of the PoCCardio Data Management Plan (based on the Horizon Europe Data Management Plan Template) which will be implemented in the project. The DMP is a living document and it will therefore be continuously updated in line with project developments or changes that either affect existing data or lead to new data. Two additional versions of the DMP will be submitted in M36 and M60.

At this stage in the project (M3), no data has yet been collected or generated and this first version of the DMP is limited to high-level description of the planned and expected data, both re-used and newly generated data. Re-used data originates from previous collaborations between technical and clinical partners and will be re-used for the clinical validation of the PoC tool (in WP6). New empirical data will be collected and generated primarily in the clinical trials in WP7 and WP8 which will begin in M8 (July 2024). It will be collected from trial participants (1836 individuals) and include personal information, medical history, clinical parameters, examination results, and laboratory measurements. Identifiable data will be pseudonymized and stored securely as numerical data or as text information in an eCRF database. Specific details on clinical data management and analysis, including further details on open access management, of the clinical outcome trial (WP7/8) will be described in a dedicated clinical data management plan (D9.1) due in M15 (February 2025).

Data management in PoCCardio will follow the FAIR principles and this deliverable describes how it foresees the implementation of FAIR data management. To ensure that data is **Findable**, the Zenodo repository will be used and Digital Object Identifier (DOI) will be assigned to all open datasets, results, and publications. The use of Zenodo supports the principle of **Accessible**. Zenodo is open, free and can be universally accessed by any browser. Access to patient data will be restricted based on privacy, ethical or security risk. Intellectual property or patent protection will also put some restrictions on accessibility. However, metadata will be openly available under a CC0 license and will be retrievable by their identifiers via the Zenodo repository. The **Interoperability** of data will be ensured by using SNOMED CT for data and metadata and open data will be released in standard, machine-independent file formats (.csv, .txt, etc.) to facilitate exchange and re-use. Datasets deposited on the Zenodo repository can be **Re-used** by anyone who has gained password-protected access for the Zenodo repository. Data will be released using the most permissible license possible.

The online repository JIRA will be used for describing, monitoring and managing datasets. All datasets will be documented and managed in accordance with the data management plan. The key to understand the requirements and needs of the DMP is a data life-cycle. The life-cycle tells where the data originates from, how and where it is stored, processing details, how it will be used and finally whether and how it will be disposed or archived. This life-cycle is integrated into a “workflow” in JIRA thereby ensuring that all datasets follow the same life-cycle.

2 Introduction

This deliverable presents the initial PoCCardio Data Management Plan (DMP) which constitutes the cornerstone for good, systematic, and responsible management of research data in the project. The DMP follows the guiding principle of "as open as possible, as closed as necessary" for all data produced and used during as well as after the end of the project. Data management in the project is based on the data life-cycle approach and the FAIR principles (Findable, Accessible, Interoperable and Re-usable).

The DMP will guide the description of all data (datasets) that will be generated or collected as well as information on the purpose of the data collection, their relation to the objectives of the project, the types and formats of the data, data re-use, data origin, their expected size, and data utility. It will also provide details on data interoperability and practical data management procedures, including templates for documenting datasets.

The DMP is based on the Horizon Europe Data Management Plan Template (EC 2022) and it will be updated when significant changes arise during the project's lifetime. The next DMP will be submitted in November 2026 (M36 of the project).

2.1 Purpose, context and scope

This first version of the PoCCardio DMP focuses on the data management methodology and defines the guiding principles for data management in the project to ensure that data is collected, processed, and used ethically, securely and in line with the FAIR principles. At the time of writing, the PoCCardio project is in its 3rd month and data collection activities are currently still in the planning phase. The data summary provided is thus based on prospective data. The following versions of the DMP will contain descriptions of the generated or collected datasets.

2.2 Content and structure

The remainder of this document outlines the stages in the data life-cycle in Section Three. Section Four provides a summary of the data that will be generated, collected, or re-used in the project. Section Five provides details on how the FAIR principles will be applied and Section Six describe how datasets will be managed internally using an online repository (JIRA). A conclusive summary of the Data Management Plan is presented in Section Seven.

3 Data Life-Cycle

The key to understand the requirements and needs of the DMP is a data life-cycle. The data life-cycle approach is used to describe whether data is collected or generated and where it originates from, how and where it is stored, processing details, how it will be used and finally whether and how it will be disposed or archived. It is important to note that there may be iterations or overlaps depending on the specific data.

3.1 Collect and/or Generate

The data life-cycle starts with an initial step of data identification. The collection of data will be restricted to data necessary for the objectives of the project. In PoCCardio, the protocols for the clinical trials identify what type of clinical data will be collected from subjects (patients) in the trials and what data is expected to be generated as a result of the data analysis. The collection of personal data will be subject to informed consent.

3.2 Storage

How and where data is stored will be determined based on the type of data that was collected /generated. All types of storage will have suitable security measures that secure data from un-authorised access, that ensure data quality, and that ensure that data can be stored in proper formats.

3.3 Processing

The processing phase involves basic processing and preparation of the data for future use in the project and beyond. Processing also includes pseudonymization, anonymization, and/or general data pre-processing, fusion, and aggregation.

3.4 Usage

The usage stage covers the various ways in which data is used for the purposes of the project. Data usage include viewing, analysing, (extended) processing, visualising etc. This phase also includes the definition of who can use the data and for what purpose. In PoCCardio, the data usage is largely defined by the tasks in the projects' workplan.

3.5 Disposal/Archiving

When data is no longer needed for the purposes of the project, it will be archived or disposed of (destroyed). When there is no longer any use for the data or it has reached the end of its retention period, it will be destroyed. The destruction of data is particularly relevant for personal data which were collected and processed for a specific purpose and with a specified retention period. The disposal of personal data must be done to protect subjects' privacy and to comply with regulations, e.g. the GDPR. Destruction methods may include data shredding, data wiping, and/or physical destruction of storage media.

The archiving of data is done to preserve data for e.g. compliance or legal purposes. Archiving data must be done in a secure and resilient manner to protect its long-term quality and accessibility. Archiving data ensures that it can be shared or re-used in the future, e.g. the archives can become a source of data for another data life-cycle, in a similar manner as in the case of the *Share* phase. The project or project members can decide that some data or publication archived is not suitable anymore for publication. In this case the data and related archive can be destroyed as well.

4 Data Summary

The aim of this project is to advance and tailor a point-of-care (PoC) tool, which was recently developed within the framework of HORIZON 2020. Using cutting-edge lab-on-a-chip (LoC) and microfluidic technology, the tool will be further developed to measure qualified biomarkers and polymorphisms from finger prick blood, in order to help identify, classify and monitor cardiovascular patients at high risk. The clinical validation of these biomarkers and the PoC tool will then be performed in a prospective, randomised multinational trial which will include only existing and approved pharmaceuticals. Data derived from the clinical validation study will constitute a rich source for complex AI-powered computational analysis to explore potential predictors for primary and secondary outcome parameters.

The foundation for the technologies to be used in WP2-WP5 is based on the know-how gathered in the H2020 projects, PoCOsteo, ELEVATE, and CHILI, which technical partners participated in together with clinical partners MUG and TUMS (in PoCOsteo). Moreover, data from a previous clinical study carried out in the PoCOsteo project by PoCCardio partners MUG and TUM will be re-used for validation of the PoC tool in WP6.

All data collection activities will be limited to the project scope and restricted to those needed for the fulfilment of the project objectives.

4.1 Re-use of existing data

The PoCOsteo project has developed the point-of-care tool which will be further developed and tested in PoCCardio. The PoCOsteo clinical study was designed as a population-based study on 1500 individuals in Bushehr, Iran, and a clinic-based study on 1000 individuals in MUG. Data from the PoCOsteo lab validation of the tool, including real blood samples, will be re-used in the PoCCardio project for the following purposes:

- To develop a generic electrochemical platform for the simultaneous, quantitative electrochemical detection of multiple biomarkers (PoCCardio-BM) and evaluate the performance of the platform along with its reproducibility and stability over time in WP2
- To develop a generic platform (PoCCardio-GM) for the multiplexed simultaneous detection of single nucleotide polymorphisms (SNPs) in genomic DNA extracted from a finger prick blood sample and to carry out an analytical evaluation and validation of PoCCardio-GM in WP3
- For microfluidic manifold development and development and realisation of a sensor cartridge design for genomic and proteomic sensing in WP4
- To develop the technical specification of the PoC tool in WP5
- To carry out a clinical validation of the PoCCardio through POCT-Cardio-Val trial clinical trial in WP6. The validation protocol designed for PoCOsteo will be the foundation of this validation study.

4.2 Generated data

Newly generated data will be collected from each individual who will participate in the clinical trial of the PoCCardio project and include personal information, medical history, clinical parameters, examination results, and laboratory measurements. Data will be collected from participant interviews or via electronic patient information systems which store medical data, data of laboratory measurements, or medical device measurements of each participating individual. The details about data specification, data storage and data management of the clinical outcome trial will be summarised in a dedicated clinical data management plan for the trial.

Likewise clinical and laboratory data will be collected in the early validation study.

While serum and plasma samples will be stored at the MedUni Graz Biobank, specific biomarker analyses plans, including multi-omics analyses, are yet to be developed throughout the project.

4.2.1 Sources and Types of data

Data will be stored as numerical data or as text information in an eCRF database. Sources of data include structured interviews or electronic patient information systems. Identifiable data will be pseudonymized by allocation of numerical personal identifiers. Data will be disseminated or communicated preferentially as numerical data both in scientific and non-scientific media (e.g. journals, presentations).

In the clinical trial, data from 1836 individuals will be collected. The size of the complete dataset will probably achieve GB or TB size.

4.3 Data Utility

The data will be useful to the general research community, in particular to scientists interested in cardiovascular risk diagnostics and management. The results of the clinical trial will help shaping treatment recommendations for people experiencing an acute coronary syndrome. In addition, medical companies and the general public will benefit from research outputs of the trial. Finally, the introduction of the PoC Device, based on these data, may be beneficial for personnel running any health care institution and patients receiving medical treatment related to cardiovascular risk management.

5 Fair Data

5.1 Making data findable, including provisions for metadata

Whenever PoCCardio results are made open, they will be assigned a Digital Object Identifier (DOI) allowing content to be uniquely citable. Open datasets, results, or publications which will be deposited within a repository (the plan is to use Zenodo, https://zenodo.org/) will automatically be designated with a DOI and will be subject to Zenodo DOI versioning. Any PoCCardio outputs which are stored in other repositories (project internal repositories, institutional, scientific publishers) will at a minimum be given a Uniform Resource Identifier (URI) and, if supported by the repository, a DOI.

Administrative metadata will be generated by Zenodo and further defined in the clinical DMP (D9.1) to be submitted at M15 (February 2025). Similarly, standards to be used for labelling of Omics data and of clinical data will be defined in the clinical DMP and are currently still subject of discussion.

The consortium also considers to harmonise sample information according to the MIABIS standard (https://www.bbmri-eric.eu/howtomiabis/). The Minimum Information About Biobank data Sharing (MIABIS) aims to standardize data elements used to describe biobanks, research on samples and associated data. The MIABIS Community Standards work on several granularity levels, with the aim to support interoperability between biobanks sharing their data. General attributes to describe biobanks, sample collections and studies at an aggregated/metadata level are defined in MIABIS Core 3.0. MIABIS Core 3.0 represents the minimum information required to initiate collaborations between biobanks and to enable the exchange of biological samples and data. The aim is to facilitate the re-use of bio-resources and associated data by harmonizing biobanking and biomedical research.

Keywords will be generated during the course of the project to facilitate data search in Zenodo based on the specific nature of the data being archived in the repository.

Metadata will be accessible via Zenodo as well as in scientific publications. Data uploaded in Zenodo will automatically be allocated a DOI and will thus be indexed.

5.2 Making data accessible

5.2.1 Repository

Zenodo shall be chosen as the default repository for the PoCCardio project. It is an OPENAIRE, Horizon 2020, and EU Research Innovation and funding program compliant repository which generates a DOI for any files uploaded to the site. Any file type can be uploaded to Zenodo including data, publications, or software. Data deposited is stored in the CERN cloud infrastructure. The work package leaders of the PoCCardio project will define the content to be shared on Zenodo.

Open datasets, results, or publications which are deposited within the Zenodo repository will automatically be designated with a DOI and will be subject to Zenodo DOI versioning.

5.2.2 Data

Data will be made available as open as possible and as closed as necessary, according to *Rules on Open Access to Scientific Publications and Open Access to Research Data in Horizon 2020* whenever possible. Restriction of data access may be related to privacy, ethical or security risks with respect to patient data, or if internal aims, contractual obligations with participating research sites or intellectual property protection of PoCCardio (e.g. successful publication in scientific journals; dissemination or marketing goals; patents) may be jeopardized by early or extended access to datasets. On Zenodo, exclusively datasets based on anonymized data can be shared, whereas pseudonymized data cannot be shared due to data protection restrictions.

In cases where only some clinical trial data can be shared on a repository, the PoCCardio project welcomes proposals for any research question related to the data and bio-samples collected. Proposal will be reviewed by the responsible work package leaders or Trial Steering Committee, respectively, and if approved, details about joint data analyses will be planned and clarified.

With respect to a potential embargo on making data openly available, the decision and information will be provided on an individual basis, depending on intellectual property protection and future directions of the project. Nevertheless, open access policy will be supported as much as possible and according to journal policies and *Rules on Open Access to Scientific Publications and Open Access to Research Data in Horizon 2020*. In the PoCCardio grant agreement, a definitive delay of data access on a repository of at least 6 months after generation of data was defined, irrespective of any other rules or restrictions that may apply.

Data decided to be shared shall be accessible through Zenodo, which is freely accessible to anyone who is granted a password-protected repository access.

Both during and after the end of the project, data are generally accessible via Zenodo access, and provision of data on this repository is defined and/or restricted by the PoCCardio consortium. For internal members of PoCCardio, the data controller (e.g. clinical trial sponsor) can grant to specific datasets stored on NextCloud (https://nextcloud.com) based on the need-to-know principle. With regards to data generated by participating clinical centres within Europe during the conduct of the clinical trial, the participating personnel will have access to data via Clincase software with restriction to participants at their own site. Trial personnel of the sponsor of the trial will have access to trial data for monitoring purposes.

Zenodo will generate a password-protected access for any person that wishes to explore data deposited on the repository. For NextCloud access, persons applying for access will get a personal password-protected access based on their unique identification.

Restrictions for external individuals wishing to access data or folders stored in the repository will be defined by the PoCCardio consortium. Whether a distinct data access committee will be defined is subject of future discussion within the consortium but for the clinical trial the Trial Steering Committee will most likely make decisions on the approval of data requests from researchers.

Specific details on data management, including further details on open access management, of the clinical outcome trial (WP7/8) will be outlaid in a dedicated clinical data management plan (D9.1) due in M15.

5.2.3 Metadata

Metadata will be made openly available under a CC0 license (https://creativecommons.org/public-domain/cc0/).

Data will be accessible on a permanent basis.

The definition of data formats remains to be defined by the consortium and will be provided in the clinical DMP (D9.1). However, commonly used formats shall be used to allow convenient access for anybody exploring the datasets stored on the repository, including standard formats such as CSV for numerical datasets and TXT for text documents.

5.3 Making data interoperable

It is intended to use the SNOMED international clinical terminology for data and metadata (https://www.snomed.org). Furthermore, standards and SOPs (Standard Operation Procedures) for disease treatment, lab manuals, and others are mandatory for the course of the clinical trial and will be captured in the clinical trial documents. Where data and results can be made openly available, they will be released in standard, machine-independent file formats (.csv, .txt, etc.) to facilitate exchange and re-use.

Should the use of any uncommon ontology or vocabulary be necessary, specific explanations or abbreviations can be found in the reference publication that can be linked to the available datasets as outlined above.

Where data from previous research is used for the published data, qualified references will be used.

5.4 Increase data re-use

Detailed information on data analysis and methodology will be provided in the specific scientific publications. To increase transparency and easy access to this information for a specific analysis or publication, the publication (or the DOI of the publication) will be provided on the research data platform and linked with the data analysis/data sheets, where appropriate.

Whilst some of the data generated and used within PoCCardio will not be openly released, whenever possible data owners are encouraged to release their data openly using the most permissible license possible. Data

can be used by anybody, including third parties, that gain password-protected access for the Zenodo repository. Restrictions to data access are defined by the PoCCardio consortium, as outlined above.

5.5 Other research outputs

Research output other than data may include the development of aptamers, the PoCCardio Device, and artificial intelligence algorithms, among other research outputs. These outputs are subject of intellectual property and may underlie transient or permanent open access restrictions; such restrictions will be defined in WP10 Impact Creation, Dissemination, Exploitation.

Due to restrictions related to intellectual property, FAIR data handling may not be providable but will always be intended as far as possible. Depending on the specific and as yet to be defined other research outputs, the management of these outputs remains to be defined during the course of the project.

5.6 Allocation of resources

There are no immediate costs that arise from making the data open access where possible. Zenodo is a free to use resource and therefore no costs are expected from utilizing this repository. If any costs, direct or indirect, arise as the project progresses, they will be reported in future updates of this DMP. However, there are costs associated with publishing papers in open-access journals which vary based on the journal and will be covered by the project budget.

Data management will be coordinated and monitored at task, work package and project level. The Clinical Trial Lead is responsible for data management on project level. Work Package Leaders and Task Leader are responsible for data management in their respective work packages and tasks. The responsible party for creating and maintaining the DMP for a data set is the partner that generates/collects the data. All individuals participating the project are responsible for familiarising themselves with the data management procedures as laid out in this deliverable.

Access to data deposited on the Zenodo repository has no time limit, thus data should be permanently accessible. Restriction and accessibility of data is decided by the PoCCardio consortium as outlined above.

5.7 Data security

Data will be stored using the NextCloud ('MUG box') software managed by MUG and only accessible via password protection. MUG will provide permanent data storage with daily and weekly data back-up.

Data will be stored safely and permanently in the Zenodo repository.

Atlassian JIRA will be used for internal management of datasets (see Section 6). It is a server-side web installation hosted on IN-JET's Azure virtual servers located in the EU. The solution is fully GDPR compliant. Only the Administrator can assign users and access is protected by password. Atlassian JIRA has a built-in backup feature and IN-JET will provide weekly back-up.

5.8 Ethics

Personal data will not be shared with third parties.

Personal data will be pseudonymised to protect the privacy rights of individual subjects. Unique identifiers will be created to pseudonymise personal data using techniques such as tokenization and/or encryption. Pseudonymised data will be held in an electronic case report form with storage servers at the Medical University of Graz (Institute for Medical Informatics, Statistics and Documentation) which will be protected by appropriate security measures (password protected) to prevent unauthorised access.

The collection, processing and management of personal data is subject to informed consent from the subject and will be in complete compliance with the GDPR.

The project's Ethics Monitoring Framework in D1.4 Statement of Ethics (for submission in M8, July 2024) will include detailed internal guidelines and protocols for handling personal data and for protecting the subject's rights under the GDPR.

The project will appoint an internal Ethics Board which will have an advisory role for the project. Two external experts will also be appointed as board members. The PoCCardio Ethics Manager will chair the board.

The study protocols for the trials (randomized controlled trial, observational trial) are all subject to approval from the relevant ethics committees.

5.9 Other issues

During the conduct of the clinical trial, data will be entered into Clincase at all clinical centres involved in patient recruitment in a standardized way. Electronic case report form data will be stored at servers of the Medical University of Graz, Austria in a secure way.

6 The Data Management Process

A Data Management Process will be implemented to facilitate the management of all the datasets covered by the DMP. It will consider all aspects and information on the dataset's life-cycle and, in particular, the FAIR principles for the datasets that are generated or re-used in the project.

6.1 Overview

The DMP management process consists of the following phases:

- *Initialize*: The phase when the data management plan methodology is defined and process of data management initialised. This deliverable boosts the initialisation phase and defines necessary concepts and process to start data management. A collaborative DMP management platform is established during this phase.
- *Assess*: The assess phase carefully monitors the phases of the dataset's life-cycle. During the assessment, a dataset template is used to record information about the dataset. Together with the template, the FAIR principles and the Ethical and Security views are evaluated as well. The assessment is done by the project partners with the most intimate understanding of the datasets under consideration.
- *Manage*: The management phase included the DMP management process as well as the datasets' pathway through their life-cycle. The aim is that the process should be transparent and managed in a way to be efficient and lean. It should guide the datasets' life-cycle stages and stimulate proper assessment of the data. Sharing of the data should be encouraged through internal assessment, selection, and recommendation of potential repositories for publishing the data. The guiding is done through project general and technical meetings as well as during dedicated data workshops. The management will involve project partners concerning quality and quantity of the shared datasets and the Ethics Manager concerning security and privacy aspects. The management phase triggers a renewed assessment phase before every reporting phase to re-assess the data templates and to add potentially new data to the collection of the templates. The process is managed by the Clinical Trial Lead.
- *Report*: The *reporting* phase is done at scheduled intervals. Forthcoming revisions of the Data Management Plan are foreseen to be submitted in M36 and M60 of the project.

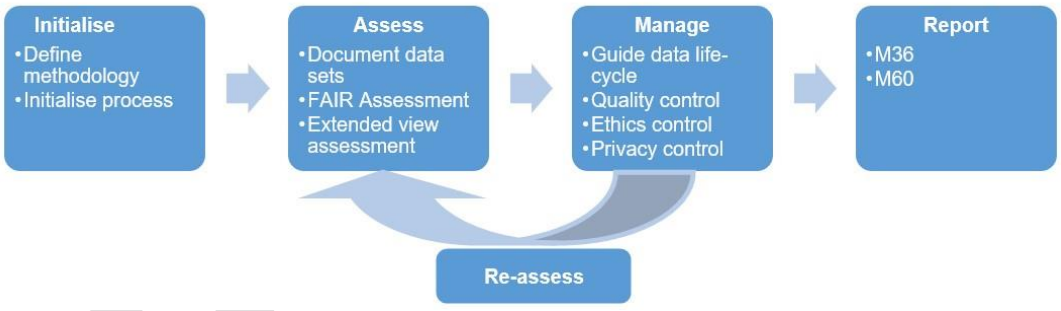


Figure 1 Data Management Process

6.2 Collaborative platform

An online collaborative repository will be used for this management process. It will be based on an Atlassian JIRA server-side web installation (https://www.atlassian.com/software/jira) hosted on IN-JET's Azure virtual servers located in the EU. The solution is fully GDPR compliant.

JIRA is an online collaborative platform where project users can log in and manage the DMP processes. Users can create and view datasets under the DMP scope and continuously track the status of the datasets and any modifications made. The dataset entries are formed around a DMP template, which contains the various database fields that define all aspects of the datasets.

JIRA provides front-end editing and management capabilities to the database content. It contains the following main components: Database, Screens, Workflow, and Analytics and these components and their use, are described in details.

6.3 Design and Implementation

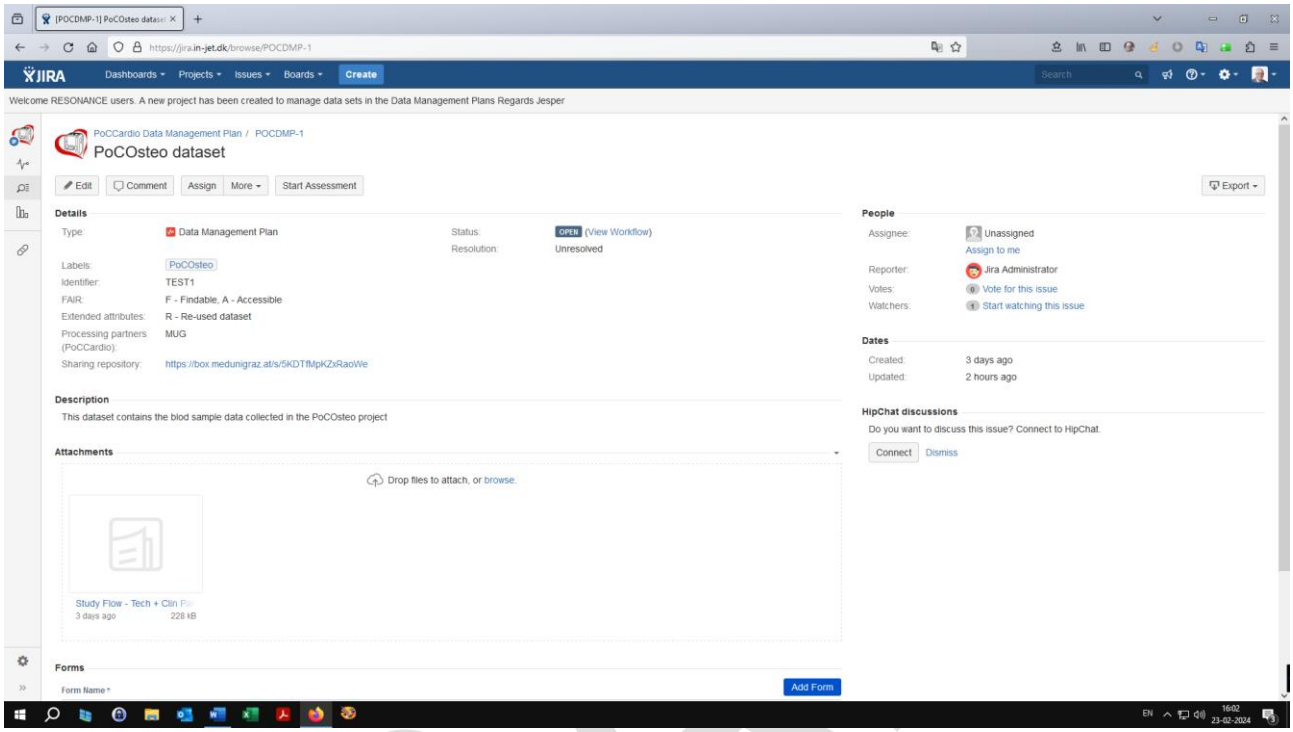


Figure 2 JIRA template for DMP description

provides an overview of the JIRA DMP template with its various features, which will be described in the following sections.

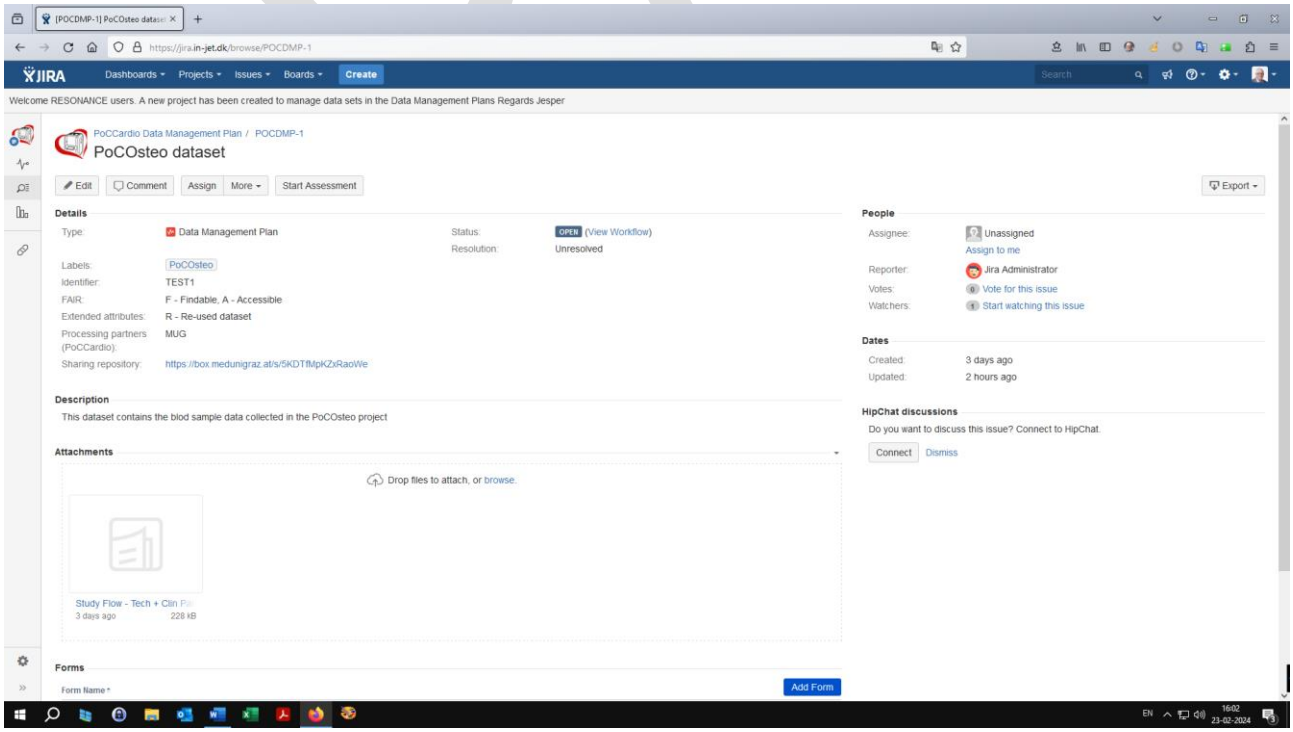


Figure 2 JIRA template for DMP description

6.4 DMP Database

The fields are defined in respect to the nature of the datasets covered by the DMP. The main fields are listed in Table 1 below:

Table 1 Dataset Template in JIRA

Field	Dataset Attributes
ID	Unique identifier of the dataset preceded by “POCDMP-“
Title	The title of the dataset
Label	Users can create tags for DMPs to be used with filters and statistics
Identifier	The identifier of the dataset
FAIR	Adherence to the FAIR principles represented by the letters F A I R if the attribute is achieved
Extended attributes (additional information)	Additional attributes of the dataset: • S: Security of the exported dataset provided • P: Data privacy noted • A: Dataset is anonymised • C: Co-Created dataset • O: Other issues exist • R: Re-used dataset
Processing partners	The PoCCardio partners responsible for processing the dataset
Sharing repository	The url of the repository where the dataset will be hosted at the end of the project
Description	A full description of the dataset
Attachments	Additional files relevant to the dataset can be uploaded here
Forms	Not used
Activity	A collection of comments, history and activities related to the dataset
Comments	Comments (which may be used to elaborate on the status and to describe the next steps to be taken, if relevant).
Status	The stage (which can be Open, In Progress, Under Review, Published). When a dataset has been cleared, its status is Published
Comments	Comments (which may be used to elaborate on the status and to describe the next steps to be taken, if relevant).
Resolution	The resolution as to how the dataset arrived at a certain stage (status)
Assignee	Name of user responsible for assessing and reviewing the dataset
Reporter	Name of person responsible for managing that the DMP process is properly handled and datasets are reviewed and approved by the Clinical Trial Lead
Due	Due date, which can be an upcoming date when the DMP shall be reviewed, approved, or cleared (archived)

6.5 DMP Screens

Screens provide front-end access to the datasets. The screens are specific to showing a particular dataset in the DMP scope (Note that the JIRA term used for these screens is “Create Issue” which cannot be changed)

There are four different screens provided. The screens will automatically appear in the context of the action which the user is planning to do.

Create DMP: This screen (Figure 3) allows the user to enter the basic information about the dataset as defined in Table 1: Identifier, labels, FAIR assessment, Extended attribute assessment, processing partners, description etc. Fields marked with * are mandatory.

Edit DMP: This screen is similar to the "Create DMP" screen. Only fields that contains information will be shown. To edit blank fields, click on the Configure Fields in the top right corner to add blank fields to the screen.

View DMP: This screen is used to show all aspects of a particular dataset as is shown in Table 1. The screen also contains the history and change audit trail as well as the comments provided by users having worked on the dataset.

Publish DMP: This screen is used after the dataset has been approved. The content of the DMP entry cannot be changed at this stage and will be published in the validated form that it was approved. A comment field is provided.

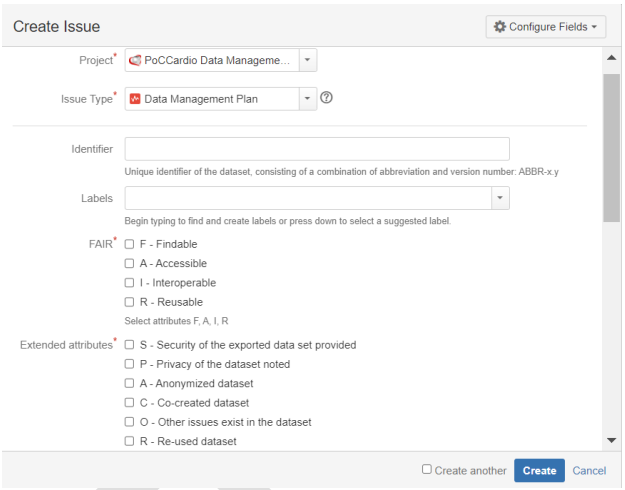


Figure 3 "Create DMP" screen in JIRA (partial screen shot)

6.6 DMP Workflow

The life cycle of the dataset in the DMP is managed through the JIRA workflow feature. This feature allows for the definition of various stages and the transition between stages. Once the workflow has been designed, the user can perform the work needed in on stage and transition the dataset to the following stage. At each transition various actions are performed automatically, such as setting database fields and sending messages to users following the DMP.

The workflow is presented as the diagram shown in Figure 4:

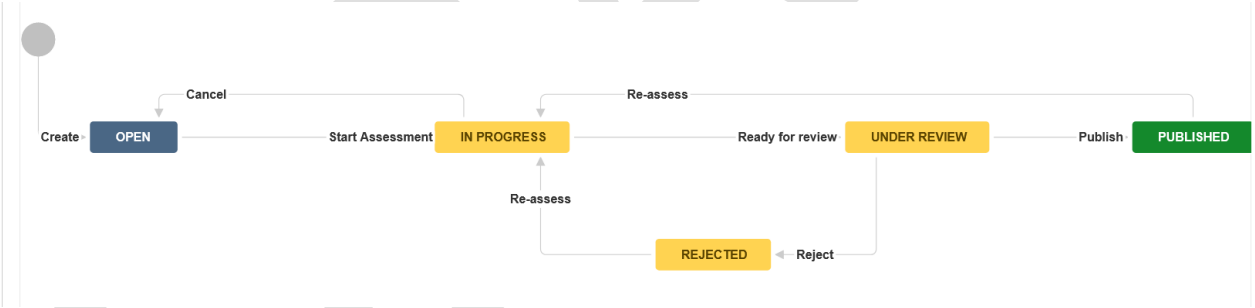


Figure 4 DMP Workflow

The workflow is built to carry a dataset through the whole process from identification and creation to assessment, review by relevant project experts and to the final publishing of the dataset DMP plan. After the approval, the dataset is published as a validated DMP plan. The Clinical Trial Lead oversees the workflow.

The workflow has five stages. The yellow stages are *Work-in-Progress* stages, the blue stage is a *To-Do* stage, and the green stage is the *Done* stage. In between stages, there are several possible transitions.

When a dataset is first identified, it is entered in the *Open* stage.

When finished, it is transitioned to the *In Progress* stage, where expert team members work on the refinement and classification of the DMP. When done, they transition the DMP to the *Under Review* stage for review by the Clinical Trial Lead. If the review is negative, it can be transitioned to the *Rejected* stage, from where it can be taken back to the *In Progress stage* for re-assessment.

When the review has been completed, the dataset is transitioned to the formal *Published* stage.

For each transition, several automatic actions take place. Most of them are of technical nature to keep the JIRA synchronised, but some transitions involve human or automatic changes to the DMP. A message is simultaneously sent to all listeners and actual assignees of the DMP.

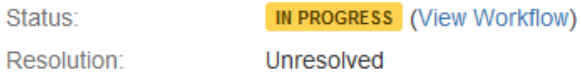
When the *Open* stage is entered, the user is presented with the *Create DMP* screen where the basic information is entered. In the transition to the *In Progress* stages, the *Edit DMP* screen is opened allowing the user to finalise the information and assign it to a specific expert (the assignee).

The assignee (and other experts) can now enter the JIRA DMP and add or change information via the *Edit DMP* screen. Once the assessment has been completed, the experts assign the dataset to the Reporter (e.g. The Clinical Trial Lead) and transitions the dataset to the *Under Review* stage.

If the Clinical Trial Lead is satisfied with the assessment, the dataset is approved and transitioned to the *Published* stage. In the transition, a comment can be added. In this transition, the *Resolution* is automatically changed to Published, indicating that no further changes are allowed.

The transition between different stages is done from the main *View DMP* screen using the buttons at the top of the screen.

The group second from the right automatically updates with buttons that show the possible transitions from the present stage, which is shown below it together with the actual *Resolution*.



In the case shown, the possible transitions from the In Progress stage are: *Ready for Review* and *Cancel.*

The (View Workflow) is a convenient way to display an overview of the full workflow in a popup window.

6.7 DMP Analytics

The JIRA platform has several tools available for managing the DMP.

Filters

Filters can be designed and stored to show only subsections of the total list of datasets in the DMP, such as list of all datasets, list of reviewed datasets, list of published datasets, etc. More complicated filters can also be constructed such as all datasets in progress and assigned to someone, all dataset entries rejected and the reason for rejections.

All dataset entries on the list can be bulk changed in a single operation. The filtered list can also be exported in many different formats for analysis or inclusion in reports.

Filters can be designed for private use, or they can be shared with the full group of PoCCardio users.

A typical list is shown in Figure 5 below.

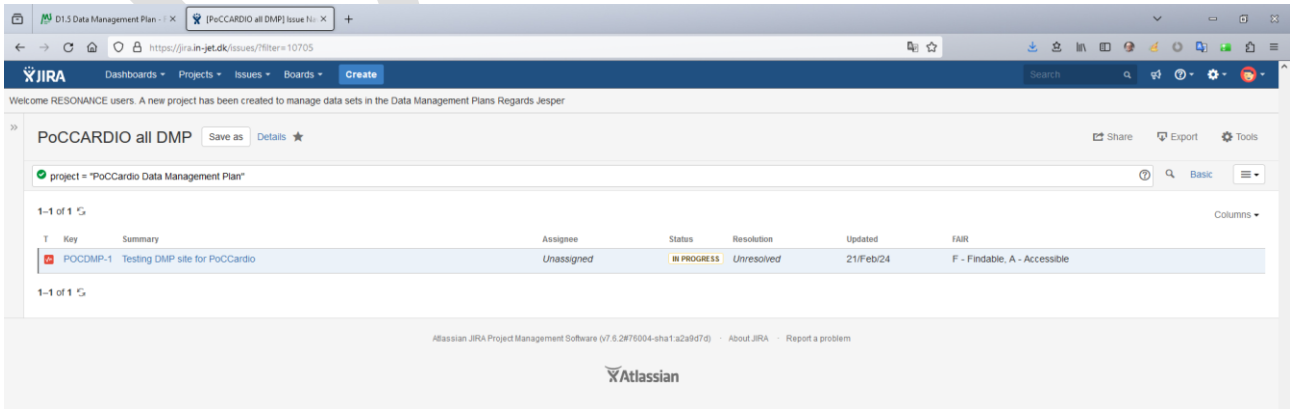


Figure 5 List of all DMP entries

Analytics

A large variety of reports can be generated and exported. They can be either based on the whole set of DMP datasets or on a filtered list. The most useful types of reports are pie charts and time lines.

The pie charts can visualise the distribution of datasets according to various parameters such as probability, assignees, stages, resolution, etc.

The time resolved reports show for example created vs. resolved datasets. It is a useful tool to see the evolvement of the total number of outstanding datasets in the DMP.

7 Conclusion

This DMP will be at the core of the research data management in PoCCardio. It provides an analysis of the elements of the PoCCardio data management policy about all the data that will be generated within the project, including how the FAIR principles will be implemented in practice.

The DMP covers the complete data life-cycle and the use of the online repository JIRA that will be used internally will ensure that all datasets are properly documented, managed, and tracked through the data life-cycle workflow. The description of the datasets in JIRA will include an assessment of FAIR principles and additional attributes such as privacy, anonymised dataset, re-used dataset, security, and other.

This first version of the PoCCardio DMP is a live document and will be revised and updated as the project progresses. The next DMP report will be submitted in M36 (November 2026).

8 List of figures and tables

8.1 Figures

8.2 Tables

9 References

(EC, 2022) European Commission (2022). Horizon Europe. EU Grants: Data management plan (HE):V1.1 – 01.04.2022.