PERMIT – PERsonalised MedicIne Trials

Data Management Plan

Work Package 1 - Deliverable 1.4

Deliverable no	1.4
Deliverable Title	Data Management Plan
Contractual delivery month	6 (June 2020)
Responsible Partner	ECRIN
Author(s)	Paula Garcia

Executive summary

The PERMIT project aims to develop recommendations on methodologies for robust and reproducible personalised medicine research. This will be achieved through an in depth analysis of existing literature on methodologies applied at each stage of the personalised medicine research pipeline, followed by dedicated workshops to discuss and define recommendations with field experts and key stakeholders. Recommendations will then be published as a consolidated report, but also disseminated through scientific publications. This document presents the Data Management Plan (DMP) for the PERMIT project, and describes the data that will be collected through project activities, how it will be managed, stored and archived. The DMP has been drafted following the Horizon 2020 template. All questions that did not apply to the data that will be collected and managed within the PERMIT project, were removed.

PERMIT has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N. 874825

DISCLAIMER

This document contains information which is the proprietary to the PERMIT Consortium. Neither this document nor the information contained herein shall be used, duplicated or communicated by any means to any third party, in whole or in parts, except with prior written consent of the PERMIT Coordinator.

The information in this document is provided "as is", and no guarantee or warranty is given that the information is fit for any particular purpose. The document reflects only the authors' view. The PERMIT Consortium is not responsible for any use that may be made of the information it contains. The user uses the information at its solo risk and liability.

Document log

Issue	Date (yyyy-mm-dd)	Comment	Author/partner
1	2020-08-21		Paula Garcia -ECRIN

PERMIT has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N. 874825

Table of Contents

PERMIT has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N. 874825

1. Data Summary

What is the purpose of the data collection/generation and its relation to the objectives of the project? What types and formats of data will the project generate/collect?

The PERMIT project has the objective of producing recommendations on Personalised Medicine (PM) research methodologies. All data that will be collected, analysed and produced will serve this purpose. The project will not produce any datasets through direct observation. It will analyse existing published information on PM research methodologies. This information will be extracted from scientific publications, and grey literature. It will be further processed through analysis and discussions, to produce concise recommendations on research methodologies.

Through the tasks of the project, personal data will be collected on individuals participating in the project workshops. This data will be collected, processed and used solely for the organization of the project workshops, in full respect of the GDPR regulation. Data will include name, email address, telephone number, occupation and current position, and institution. No sensitive personal data will be collected and all individuals beyond the PERMIT Consortium will be asked for written consent, before their personal data is collected. This data will not be shared, transferred, or made publicly available in any way, at any point within the lifecycle of the project. It will be destroyed once the project ends.

Will you re-use any existing data and how?

The first stage of the PERMIT project will focus on a literature review, more specifically a scoping review, to identify and analyse the existing methodologies for PM research. It will search for and select for analysis scientific publications and grey literature that describe methods and their application in PM research. Some of the articles analysed through the scoping review will contain non-personal data, as aggregate statistics. The PERMIT project will not specifically re-use this data, but will focus on which methodologies and how they were applied to collect, manage and process the data.

None of the personal data collected through the project will be re-used in any way, nor will it be made publicly available. The recommendations that the project will produce, will be made publicly available. These recommendations will reference all of the literature that was consulted and analysed to produce them, guarantying transparency and access to referenced information.

What is the expected size of the data?

The volume of data that will be collected will be low. Personal data of workshop participants is expected to be stored in a single Excel file of no more than 25 Kb. Files condensing the extracted information from literature reviews, and the written recommendations will also represent a low volume of data, mainly in files produced via word processing software, not surpassing 1 Gb.

To whom might it be useful ('data utility')?

The information that will be extracted from the existing literature, already publicly available, can be of interest to all PM research stakeholders. Furthermore, the recommendations that the project will produce will be of use to all stakeholders as it will allow them to produce and/or assess PM research in light of concise and clear methodological standards for robust and reliable results.

2. FAIR data

2. 1. Making data findable, including provisions for metadata

Are the data produced and/or used in the project discoverable with metadata, identifiable and locatable by means of a standard identification mechanism (e.g. persistent and unique identifiers such as Digital Object Identifiers)?

The data extracted from scientific publications is currently findable by DOI and/or PMID. The results of the scoping review will be published not only as a project deliverable, but also as a report and as scientific publications. The recommendations will be published as a concise and comprehensible document, publicly available via de PERMIT project website, but will also be described and presented in scientific publications. All publications will be findable by DOI or PMID and will be indexed by major medical databases.

What naming conventions do you follow?

No datasets will be made publicly available, as all information extracted from literature is already publicly available. The results of the scoping review will list full references of all publications include in the review as an appendix. If a suitable repository is identified, this list will also be made publicly available through a repository. In terms of scientific publications on the PERMIT recommendations, all publications will reference the PERMIT consortium.

Will search keywords be provided that optimize possibilities for re-use?

When possible, articles on the recommendations will include key words to boost visibility and make publications easily findable.

2.2. Making data openly accessible

Which data produced and/or used in the project will be made openly available as the default? If certain datasets cannot be shared (or need to be shared under restrictions), explain why, clearly separating legal and contractual reasons from voluntary restrictions.

PERMIT has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N. 874825

All project results (i.e. the recommendations) will be made publicly available. They will include the full references of all literature that was analysed for their production. The protocol for the scoping review of the literature will be made publicly available via Zenodo. Furthermore, all project deliverables, including a report on the scoping review that will detail the methodology applied for the selection and analysis of publications within the review, will be made available.

The extraction forms from the literature analysis, will contain only data that is currently publicly available, therefore the extraction forms will not be published as such.

How will the data be made accessible (e.g. by deposition in a repository)?

The protocol for the scoping review of the literature will be made publicly available via Zenodo. All recommendations will be made available through the PERMIT project website, as will the approved project deliverables. The Consortium will favour scientific publications in Open Access journals, but will not be limited to such.

What methods or software tools are needed to access the data?

No specific methods or software will be required to access the project's data.

2.3. Making data interoperable

Are the data produced in the project interoperable, that is allowing data exchange and re-use between researchers, institutions, organisations, countries, etc. (i.e. adhering to standards for formats, as much as possible compliant with available (open) software applications, and in particular facilitating re-combinations with different datasets from different origins)?

What data and metadata vocabularies, standards or methodologies will you follow to make your data interoperable?

Will you be using standard vocabularies for all data types present in your data set, to allow inter-disciplinary interoperability?

In case it is unavoidable that you use uncommon or generate project specific ontologies or vocabularies, will you provide mappings to more commonly used ontologies?

As no raw datasets will be produced by the PERMIT project through direct data collection or observation, and all analyses for the drafting of the recommendations will be built on existing, publicly available information, the notion of data interoperability is not relevant. The data extracted through the scoping review will constitute the basis for the production of the recommendations. As mentioned above, the results of the scoping review will be published and publicly available, and will include the full list of all publications included in the scoping review.

The recommendations that the PERMIT project will produce will be drafted in collaboration with all relevant PM research stakeholders, this will not only allow all perspectives to be taken into consideration; it will also allow the recommendations to clearly address how stakeholders can pragmatically interpret and implement the recommendations.

2.4. Increase data re-use (through clarifying licences)

How will the data be licensed to permit the widest re-use possible?

When will the data be made available for re-use? If an embargo is sought to give time to publish or seek patents, specify why and how long this will apply, bearing in mind that research data should be made available as soon as possible.

Are the data produced and/or used in the project useable by third parties, in particular after the end of the project? If the re-use of some data is restricted, explain why.

How long is it intended that the data remains re-usable?

Are data quality assurance processes described?

Further to the FAIR principles, DMPs should also address:

The PERMIT recommendations will outlive the project, becoming a guidance document for all stakeholders in the PM research ecosystem. They will remain publicly available through the PERMIT project website that will be hosted by ECRIN, beyond the end of the project. Publications on the project will also be available as open access publications.

3. Allocation of resources

What are the costs for making data FAIR in your project?

How will these be covered? Note that costs related to open access to research data are eligible as part of the Horizon 2020 grant (if compliant with the Grant Agreement conditions).

The PERMIT project will not produce any raw datasets; therefore, no costs are associated to making these openly accessible. In terms of scientific publications, 10 000 € of the budget from the WP focused on communication, dissemination and implementation, have been set aside for publications. This budget will be used to finance scientific publications as open access publications.

Who will be responsible for data management in your project?

PERMIT has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N. 874825

Management of project data will be ensured by all consortium beneficiaries, as all beneficiaries will participate in the scoping review, and subsequent analysis and discussion of findings for the production of recommendations. As stipulated in the Consortium Agreement, all beneficiaries will uphold EU and applicable national regulations when handling any and all data.

Are the resources for long term preservation discussed (costs and potential value, who decides and how what data will be kept and for how long)?

All personal data collected for the project workshops will be used only for this purpose during the life of the project, and will be deleted at the end of the project.

The recommendations and approved deliverables will remain publicly available through the PERMIT project website that will be hosted by ECRIN, beyond the end of the project. The site will be hosted as a micro-site within the ECRIN website.

4. Data security

What provisions are in place for data security (including data recovery as well as secure storage and transfer of sensitive data)?

No sensitive data is collected through the PERMIT project. Nevertheless, all personal data collected for the workshops is stored by beneficiaries in a secure archiving system, as all beneficiaries (as stipulated in the Consortium Agreement), have agreed to uphold EU regulation, including GDPR. Therefore, all beneficiaries commit to storing personal data safely and ensuring that any correction or deletion of personal data request can be executed.

Is the data safely stored in certified repositories for long term preservation and curation?

No project data will be stored in certified repositories for long term preservation and curation. Approved project deliverables and recommendations, publicly available through the project website will remain available beyond the end of the project, as the website will continue to be hosted by ECRIN after the end of the project.

5. Ethical aspects

Are there any ethical or legal issues that can have an impact on data sharing? These can also be discussed in the context of the ethics review. If relevant, include references to ethics deliverables and ethics chapter in the Description of the Action (DoA).

Is informed consent for data sharing and long term preservation included in questionnaires dealing with personal data?

As no sensitive data will be collected through the project, no ethical or legal issues are foreseen. Deliverable 8.1 – POPD Requirement n°2 will describe the precise procedures for collection, processing and storage of personal data during the project lifecycle for the project workshops.

6. Other issues

Do you make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones?

No other specific procedures for data management are implemented in the PERMIT project.

PERMIT has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement N. 874825