

## D1.3 Data Management Plan

115966 – PREFER

**Patient Preferences in  
Benefit-Risk Assessments  
during the Drug Life Cycle**

**[WP1 – Project Management]**

<b>Lead contributor</b>	Monika Brand (Actelion – 22) <a href="mailto:monika.brand@actelion.com">monika.brand@actelion.com</a>
<b>Other contributors</b>	Isabelle Huys (University of Leuven, KUL – 4) <a href="mailto:Isabelle.huys@kuleuven.be">Isabelle.huys@kuleuven.be</a>
	Eline van Overbeeke (University of Leuven, KUL – 4) <a href="mailto:eline.vanoverbeeke@kuleuven.be">eline.vanoverbeeke@kuleuven.be</a>
	Josepine Fernow (Uppsala University, UU - 1) <a href="mailto:josepine.fernow@crb.uu.se">josepine.fernow@crb.uu.se</a>
	Susan Bhatti (Merck - 27) <a href="mailto:susan.bhatti@merckgroup.com">susan.bhatti@merckgroup.com</a>
	Conny Berlin (Novartis - 15) <a href="mailto:conny.berlin@novartis.com">conny.berlin@novartis.com</a>

<b>Due date</b>	31 March 2017
<b>Delivery date</b>	20 April 2017
<b>Deliverable type</b>	R <sup>1</sup>
<b>Dissemination level</b>	CO <sup>2</sup>

<sup>1</sup> R: Document, report (excluding the periodic and final reports)

<sup>2</sup> CO = Confidential, restricted under conditions set out in Model Grant Agreement

Description of Work	Version	Date
	V1.2	20 April 2017

## Document History

Version	Date	Description
V0.7	09 March 2017	First Draft – Initial Data Management Plan (DMP)
V0.8	16 March 2017	Comments – Initial DMP
V0.9	30 March 2017	Draft – Initial DMP
V1.0	31 March 2017	Pre-final Version – Initial DMP
V1.1	19 April 2017	Review of the Steering Committee – Initial DMP
V1.2	20 April 2017	Final Version – Initial DMP

## Table of Contents

1. Introduction and aim.....	4
2. General principles .....	5
3. Overview of data managers, data repositories and access rules .....	6
3.1 Projectplace .....	6
3.2 The KU Leuven (KUL) repository for personal data .....	7
3.2.1 Specifications and costs of the KUL repository.....	7
3.2.2 Procedures/tools for data accessibility / security .....	7
3.2.3 Duration of accessibility .....	8
3.2.4 Data transfer.....	8
3.2.5 Back-up process.....	8
3.2.6 Disaster Recovery .....	8
3.3 The Uppsala repository for long-term storage.....	8
3.3.1 Specifications of ALLVIS.....	8
3.3.2 Archiving.....	9
3.3.3 Procedures/tools for data accessibility/security .....	9
3.3.4 Back-up process.....	9
3.3.5 Disaster Recovery .....	9
4. Overview of data types generated and collected in PREFER .....	10
5. Operational data management requirements for PREFER research projects .....	12
5.1 Requirements for the short dataset specific DMP .....	12
5.2 Responsibilities of the data owner .....	14
6. Sharing and secondary use of PREFER generated or collected data .....	15
6.1 Procedures for making data findable.....	15
6.2 Re-use within the PREFER consortium .....	15
6.3 Re-use of PREFER results by third parties.....	15
7. Protection of personal data .....	16
8. Ethical aspects .....	18
8.1 General ethical aspects .....	18
8.2 Interviews and patient preference studies .....	18
9. References .....	20
10. List of abbreviations.....	21
11. GLOSSARY .....	22

## 1. Introduction and aim

The main objective of the PREFER project is to strengthen patient-centric decision making throughout the life cycle of medical products (a term which, in the context of this project, includes medicinal products(1) and medical devices(2)) by developing evidence-based recommendations to guide industry, Regulatory Authorities, Health technology assessment (HTA) bodies and reimbursement agencies on how and when patient-preference studies could be performed, and how the results can be used to support and inform decision making.

The PREFER Consortium Agreement indicates that a specific data management plan (DMP) will be created. More specifically, the Consortium agreement indicated in section 7.5.4 the following:

*'As an exception to Clause 29.3 of the Grant Agreement, as provided for in its last paragraph, certain Beneficiaries have indicated that their main objective in the Action would be jeopardized by making all or specific parts of the research data of the Action openly accessible. Beneficiaries have therefore agreed to a data management plan, which describes how data will be handled instead of open access, and which plan details the reasons for not giving open access. Such data management plan is a deliverable of Work Package 1 and shall be added as Appendix 7 to this Consortium Agreement.'* (3)

The DMP is an evolving document with the final DMP forming the Appendix 7 of the Consortium Agreement, describing all aspects of how the data generated within PREFER were managed.

The Description of Action (DoA) of PREFER (p.19-20) provides the general framework regarding data management, data protection, data sharing, data ownership, accessibility, and sustainability requirements.(4)

In this initial DMP the management of generated and collected individual-level data is described, not the management of analyses and reports containing aggregated data. These issues will be covered in the final DMP.

Overall, the DMP provides a description of the data management that will be applied in the PREFER project including:

- a description of the data repositories, who is able to access the data, and who owns the data.
- the main DMP elements for each of the research projects (interviews, literature review, case study, etc.) contributing to PREFER, to be defined and provided to PREFER (Chapter 5).
- the time period for which data must be stored.
- the standards for data collection and evaluation.
- the possibilities of and conditions for sharing data.
- the implementation of data protection requirements.

As the DMP is an evolving document, some of the aspects may be described in a later version of the DMP.

In summary, the PREFER DMP gives guidance and provides an oversight of general data management, while each research project needs to provide specific data management information including, but not limited to, data capture systems, data analysis systems, data protection and data privacy measures, including description of de-identification of data sets and access rules. And in cases where the research results are not open access a justification needs to be provided.

## 2. General principles

This is the Initial DMP for PREFER. The DMP is a working document, that will evolve during the PREFER project, and will be updated to reflect project progress. Table 1 lists the deliverable version updates of the DMP for PREFER. Additional updates will be done whenever important changes occur e.g. due to the creation of new data sets.

Processes relating to the different data management plan aspects will be worked out between M6 and M18 and explained further in the next version of the DMP due in M18.

**Table 1** PREFER Data Management Plan (DMP) deliverables

Del. no.*	Deliverable name	WP no.	Short name of lead participant	Type	Dissemination level	Delivery Date**
1.3	Initial DMP	1	Actelion	R	PU	6 (March 2017)
1.6	Update DMP	1	Actelion	R	CO	18 (March 2018)
1.9	Final DMP	1	Actelion	R	PU	60 (September 2020)

DMP= Data Management Plan; WP= Work packages; R = Document, report (excluding the periodic and final reports); DEC = Websites, patents filing, press & media actions, videos, etc.; PU = Public, fully open, e.g. web; CO = Confidential, restricted under conditions set out in Model Grant Agreement

\* According to the Table 3.1c: List of deliverables of the PREFER Description of Action(4)

\*\* Measured in months from the project start date (October 2016, Month 1)

The DMP provides practical instructions with respect to any requirements for local exceptions to data management.

The DMP follows the principles that research data are findable, accessible, interoperable and reusable (FAIR)(5) as well as being attributable, legible, contemporaneous, original and accurate (ALCOA)(6).

The terminology used in this DMP is explained in the glossary (Chapter 11 of this DMP).

The general principles on access rules are defined in the consortium agreement (section 8) (3).

For research data generated as part of an ongoing medicinal product development program within industry, there may be proprietary and privacy concerns that will be acknowledged and agreements made with the respective partners on data accessibility and data storage. To acknowledge potential differences for industry or academic case studies the DMP will refer to “data generated in industry-led studies” and “data generated in academic-led studies”.

### 3. Overview of data managers, data repositories and access rules

Three repositories / platforms are used in the PREFER project. The responsible contacts are listed in table 2.

- The platform “Projectplace” is used as an interaction platform for PREFER members to **store and exchange reports and anonymous data**.
- The data repository at KU Leuven (Digital Vault for Private Data) is used to **store and exchange sensitive personal data** in a secure and protected environment during the conduct of PREFER.
- The data repository at Uppsala University (ALLVIS) will be used for **long-term storage** of reports and anonymized data particularly after the end of the PREFER project.

The use of the KU Leuven repository is preferred for the storage of interviews and academic patient preference studies. Local national laws and requirements need to be applied and can result in deviations. For example in the UK, the UK Sponsor and the Research Ethics committee will determine where it is allowed to store data.

Data sets containing personal data can also be stored by the data owners in their own repository for a fixed period of time, as defined in the applicable laws or regulations, but this should be a secure repository. Copies of datasets containing personal data in the possession of partners other than the research data owner (see 5.2) must be destroyed at the end of the PREFER project. Other non-public and public datasets not containing personal data will be stored for at least 10 years from the end of the PREFER project in the Uppsala data repository, to ensure their long-term availability to future researchers.

**Table 2** Main contacts for data management aspects

Responsibilities	Name	E-mail address
Data Management compliance contact	Monika Brand	<a href="mailto:monika.brand@actelion.com">monika.brand@actelion.com</a>
Deputy Data Management Compliance contact	Eline van Overbeeke	<a href="mailto:eline.vanoverbeeke@kuleuven.be">eline.vanoverbeeke@kuleuven.be</a>
ProjectPlace contact	Carl Steinbeisser	<a href="mailto:carl@steinbeisser-management.com">carl@steinbeisser-management.com</a>
KU Leuven repository contact	Isabelle Huys	<a href="mailto:Isabelle.huys@kuleuven.be">Isabelle.huys@kuleuven.be</a>
KU Leuven deputy repository contact	Eline van Overbeeke	<a href="mailto:eline.vanoverbeeke@kuleuven.be">eline.vanoverbeeke@kuleuven.be</a>
Uppsala repository contact	Mats Hansson	<a href="mailto:mats.hansson@crb.uu.se">mats.hansson@crb.uu.se</a>
Uppsala deputy repository contact	Head of the Department of Public Health and Caring Sciences	

The WP1 data management team has the responsibility to update the names related to the responsibilities, as people might change position.

All questions related to data management such as rules for uploading data sets, request for access rights should be sent to the Data Management Compliance contact and the Deputy. The processes and the role description of the Data management compliance contact and its deputy will be worked out in the next period between M6 and M18 and explained further in the next version of the DMP due in M18.

WP leads are responsible for informing the Data Management Compliance contacts about all generated data sets, in their research projects.

#### 3.1 Projectplace

Projectplace is the platform used by PREFER to facilitate collaboration between PREFER members, to plan deliverables, to track progress of all tasks, and to store meeting minutes and task reports. All PREFER members have an account so they can access Projectplace.

## 3.2 The KU Leuven (KUL) repository for personal data

A secured repository to store and to exchange sensitive personal data will be provided by KUL and is known as a “digital vault for private data”. Within this digital vault, researchers can keep personal data safe and apply strict rules for data access. In addition, they can also anonymise information and process it outside the digital vault without causing any data privacy risk. The digital vault is a highly secure environment within a secure network. Several vaults can be set up within this secure network, each for a different project. Each vault consists of a protected server (Windows or Linux) and can only be accessed by a well-defined user group.

The KUL repository will function as the virtual workplace to share and assess the individual-level data as needed to fulfil the PREFER objectives. Processes relating to the use of the KU Leuven repository will be worked out after M6 and explained further in the next version of the DMP due in M18.

### 3.2.1 Specifications and costs of the KUL repository

The repository consists of:

- A **secure server and operating system** in the special, secure environment for private data:
  - A virtual Windows server (1 CPU and 4 GB RAM) or a virtual Linux server (1 CPU and 2 GB RAM).
  - An IP address, DNS entry and name for the virtual server.
  - An ICTS-guaranteed licence for the operating system (Windows server or Linux CentOS).
  - Installation of the Windows or Linux operating system (including latest upgrades and security patches) on the virtual server.
  - Monthly maintenance of the Windows or Linux operating system, i.e. regular application of upgrades and security patches.
  - Access to the virtual server via an RDP client (remote desktop protocol) for Windows or an SSH client for Linux. Preliminary VPN connection is required.
- **Application software on the server:**
  - Installation of SAS and SPSS on the virtual server.
  - An ICTS-guaranteed SAS and SPSS software license.
- **Storage capacity for data:**
  - 50 GB storage space for data (server back-end storage, type 1, with mirror).
- **Cost of the repository:** € 1.291,79 per year

### 3.2.2 Procedures/tools for data accessibility / security

Details of all users of the digital vault must be registered with KUL. External users must have minimal details registered. The user/requestor is responsible for ensuring their registration. Access to identifiable personal data on the secure ICTS server is restricted to a minimum number of people, i.e. people whose task it is to decrypt or anonymise information. Anonymised information is sufficient for the majority of researchers involved in a project. These data can be processed outside the digital vault; therefore, access to the digital vault is not necessary or even desirable for these researchers. One person (the data owner, see chapter 5) per task will get access to the data repository. If additional people require access to the digital vault after its initial set up, this access must be requested by the person responsible for the digital vault. This can be done by e-mailing [servicebeheer@icts.kuleuven.be](mailto:servicebeheer@icts.kuleuven.be). For PREFER the KUL repository manager is listed in table 2.

Access to the digital vault is only possible through a Luna account (KU Leuven user ID and -password). The digital vault is only accessible through the KU Leuven VPN solution. The user must authenticate when setting up the VPN connection. A vault-specific VPN profile ensures that access is possible only to the corresponding vault in the secure network. Access to the secure network environment that houses all the vaults is strictly protected. The secure network environment is protected from the outside by a firewall, which only allows traffic:

- from the VPN solution (through a specific profile) to the servers and information in the corresponding vault;
- from the KUL ICTS management network (for system administration) from a central system console.

The server in the vault is managed by KUL ICTS and only KUL ICTS personnel have administrator/root rights. KUL ICTS personnel are bound by the KUL ICT code of conduct for staff.

### 3.2.3 Duration of accessibility

Users with access to the digital vault only have user rights for access to the data in their own vault. A service agreement for a “Digital vault for private data” has a duration of 1 year, after which it tacitly renews each year unless the IT manager responsible gives notice on the agreement by e-mail to [servicebeheer@icts.kuleuven.be](mailto:servicebeheer@icts.kuleuven.be), at the latest 3 months before the end of the agreement. If notice is given on the digital vault agreement after the project ends, the information will be irrevocably deleted and will become irrecoverable. An agreement will be set up with KUL to guarantee access for 5 years, namely during the duration of the IMI PREFER project. Long term storage after the end of the PREFER project are described in section 3.3.

### 3.2.4 Data transfer

Data transfer files to be generated and uploaded to the digital vault can directly be uploaded by the data owner in the secure environment. For this the data owner needs to have access to the digital vault (see chapter 5). If the data is not directly available to the data owner, the data can be transferred to the data owner through a secure FTP (SFTP) or can be delivered to the data owner via a physical medium (DVD/CD/USB).

### 3.2.5 Back-up process

Stored data is backed up using “snapshot” technology, where all incremental changes in respect of the previous version are kept online on a different server at the KU Leuven. As standard, 10% of the requested storage is reserved for backups using the following backup regime:

- An hourly backup (at 8 a.m., 12 p.m., 4 p.m. and 8 p.m.), the last 6 of which are kept.
- A daily backup (every day) at midnight, the last 6 of which are kept.
- A weekly backup (every week) at midnight between Saturday and Sunday, the last 2 of which are kept.

### 3.2.6 Disaster Recovery

The repository has 50 GB storage space for data (server back-end storage, type 1), and a mirror storage system at a different building of the KU Leuven in another part of the city is provided to enable disaster recovery.

## 3.3 The Uppsala repository for long-term storage

The data repository ALLVIS at Uppsala university will be used to archive the PREFER anonymized data used for publications as well as the PREFER recommendation documents and all content from ProjectPlace. Mats G. Hansson is the owner and responsible for the ALLVIS repository, listed in table 2. He is deputized by the Head of the Department of Public Health and Caring Sciences, if applicable.

### 3.3.1 Specifications of ALLVIS

ALLVIS is a storage platform and the respective research data owner is responsible for transferring anonymized data to ALLVIS for storage. The process and timing for such storage will be further worked out after M6 and detailed in the next version of the DMP due in M18. If stored data need to be transferred to platform for processing again the research data owner is responsible for the data transfer. ALLVIS will not release any data without the agreement between the repository owner and the research data owner (see section 5.2 for definition). However, the research data owner has to comply with the principle of public access to official records.

The Principle of Public Access (Offentlighetsprincipen) in Sweden means that activities of public authorities are open to the public and research activities are no exception. Universities in Sweden are legally considered



as public authorities. Records of data and research results created in the research process are subject to implementation of the Principle of Public Access, regardless of the kind of research or source of funding.

Public access can either be 1) public without restrictions, 2) public but with restricted access regulated by Secrecy Law. However, there might be working documents that do not fall under the public access rules.

### **3.3.2 Archiving**

Administrative records (e.g. Ethics approval) are stored by public authorities with reference to Archive Law. During the course of the PREFER project, administrative records and documents are stored in ProjectPlace. These documents and records will be archived in the ALLVIS repository at Uppsala University for 10 years after the end of the project. Once archived, records are subject to the principle of public access. Uppsala University shall draw up a description of this archive and a systematic archival inventory.

### **3.3.3 Procedures/tools for data accessibility/security**

Access to the file repository is granted via a Windows file share using SMB v3. Outside Uppsala University, access is granted only through a secure VPN-connection. Authentication against the VPN and authentication against the file share is granted using a personal/identifiable user account from Uppsala University. Authentication at Uppsala University is handled by a central user database and is used by the VPN and file share. Access to the project area is limited to the research data owner (e.g. Principal Investigator (PI)) and users granted access. Data is stored by an enterprise-grade NAS-system, which has been installed and configured in accordance with the supplier's guidelines and is hosted in an on-campus server hall.

### **3.3.4 Back-up process**

Backups are incrementally saved every night using an enterprise-grade backup system at a University-affiliated off-campus site.

### **3.3.5 Disaster Recovery**

Disaster recovery is in place and is possible. Disaster recovery is handled on a per-case base. Requests can be made either by phone or e-mail, contacting Uppsala University's Servicedesk. The Servicedesk can be contacted weekday's 08:00-21:00 and on weekends 14:00-17:30. E-mail: <http://uadm.uu.se/it/om/servicedesk>.

## 4. Overview of data types generated and collected in PREFER

The data generated and collected during the PREFER project can be divided into two categories of decreasing confidentiality:

1. datasets containing personal data
2. datasets containing non-personal data

The data generated within the PREFER project are (a) primary data (original research) produced by different stakeholder e.g. interviews and case studies, and (b) secondary data (reuse of existing data) such as database studies and literature reviews. Primary data are data sets more likely to contain personal data, while secondary data sets are more likely to containing non-personal data.

Patient data will be generated and processed during the activities planned in WP 2, WP 3 and WP 4 (table 3).

- **WP 2** will generate datasets containing literature reviews, recorded interviews, transcriptions of interviews, and review of reports in preference research
- **WP 3** will create datasets containing both aggregated and patient-level identified or de-identified data. These data can be created from historical case studies, prospective industry-led and academic-led case studies, surveys, as well as from simulation studies
- **WP 4** will generate datasets containing literature reviews and data resulting from expert panel discussions and consultation rounds.

Appropriate strategies have to be put in place by the individual research project owners, to ensure (personal) data protection/privacy, and individual studies are asked to provide a small DMP as described in this DMP (Chapter 5). The processes will be worked out and implemented between M6 and M18, in collaboration with task WP 3 task 3.1 to align the templates and to use synergies for research project descriptions.

The WP1 Data Management Team will generate a meta-data repository of all research projects in a format as outlined in table 3 and with the support from the WP leads, or research project owners, respectively. This meta-data repository will be updated regularly (at least on a quarterly basis) and is the master file for more detailed information of each research project as described in table 3.

The WP1 Data Management Compliance Contacts (table 2) together with the WP leads will establish a process to ensure that all generated data sets, or research projects, respectively, will be gathered as described in this DMP.

The data are expected to be useful for the PREFER project, especially for the specific tasks that generate or collect or re-use the data, and the analyses and reports will be useful to all stakeholders.

Table 3 will be updated with the unique identification numbers as described in Chapter 5.

**Table 3** Summary of the PREFER-generated data

Task*	Objective	Design	Type	Format	Re-use**	Origin	Size	Ca***
2.1	Identifying desires, expectations, concerns and requirements of stakeholders about methodologies for PP elicitation and their use in decision making	Literature review	Born digital, reference	Textual	2.3	Secondary	TBA	2
		Interviews	Born digital, observational	Multimedia + textual	2.3	Primary	TBA	1
2.2	Determine processes, conditions, contextual factors that influence the utility and role of PP studies	Literature review	Born digital, reference	Textual	2.3	Secondary	TBA	2
		Interviews	Born digital, observational	Multimedia + textual	2.3	Primary	TBA	1
2.3	Identification of assessment criteria used at decision points throughout the DLC	Literature review	Born digital, reference	Textual	/	Secondary	TBA	2
		Interviews	Born digital, observational	Multimedia + textual	/	Primary	TBA	1
2.4	Identification of preference elicitation methods	Literature review	Born digital, reference	Textual	2.6	Secondary	TBA	2
		Interviews	Born digital, observational	Multimedia + textual	2.6	Primary	TBA	1
2.5	1. Identification of educational/gamified tools	Literature review	Born digital, reference	Textual	/	Secondary	TBA	2
	2. Identification of psychological tools	Literature review	Born digital, reference	Textual	/	Secondary	TBA	2
	3. Presentation of risks	Literature review	Born digital, reference	Textual	/	Secondary	TBA	2
2.7	Identification of candidate methodologies and criteria to assess empirical case and simulation studies	Interviews	Born digital, observational	Multimedia + textual	/	Primary	TBA	1
3.3	Identifying and assessing historical case studies from industry partners	Review of historical case studies	Born digital, reference	Textual	/	Secondary	TBA	2
3.3	Lessons learned survey of PREFER members with preference research experience.	Survey	Born digital, reference	Textual	/	Primary	TBA	1
3.4	Identifying and supporting prospective case studies from industry partners	PP case study	Origin TBD, observational	Textual, numerical, multimedia	/	Primary	TBA	1
3.5-3.7	Empirical case studies and simulation studies	PP case study	Origin TBD, observational + simulation	Textual, numerical, multimedia, models	/	Primary	TBA	1
3.8	Additional case studies	PP case study	Origin TBD, observational	Textual, numerical, multimedia	/	Primary	TBA	1
4.3	Expert panels on recommendations	Interviews	Born digital, observational	Multimedia + textual	/	Primary	TBA	1
4.4	Consultation rounds on recommendations	Interviews	Born digital, observational	Multimedia + textual	/	Primary	TBA	1

PP= Patient preferences; DLC=Drug Life Cycle; Ca= Category; TBA= To be announced

\* According to the description of the tasks and different work packages in the PREFER DoA document of 16/07/17.

\*\* Displays in which other tasks of WP2 and WP3 the data are used.

\*\*\* The data produced and used during the PREFER project can be divided into two categories (Ca):

1. datasets containing (sensitive) personal data
2. datasets containing non-personal data

## 5. Operational data management requirements for PREFER research projects

Each research project (interviews, literature review, surveys, case studies, etc.) needs to provide a short dataset specific DMP, including but not limited to data capture systems, data analysis systems, data protection and data privacy measures, including description of de-identification of data sets and access rules. If the research results cannot be open access a justification needs to be provided.

### 5.1 Requirements for the short dataset specific DMP

All data owners need to fill in **Table 4** (available on ProjectPlace as a template) containing the meta data and describing the data management of data sets. Metadata are specifications for data that provide the contextual information required to understand those data. Such specifications describe the structure, data elements, interrelationships and other characteristics of data, the data repository used, and need to be securely stored with the database.

These tables will be reviewed by the WP1 data management team for completeness, compliance with the DMP and compliance with the Consortium Agreement. The text in *blue and italic* gives guidance on what information should be provided and should be replaced.

As part of the DMP an evolving data governance document of the different study types will be maintained (WP 1, Deliverables 1.3, 1.6 and 1.9, M6, M18, M60). This data governance document (based on table 4) will be kept and maintained in Projectplace and attached to the DMP at the given deliverables times.

Table 4 Metadata requested per dataset (adapted from the Data Management General Guidance of the DMP Tool)(7)

*This table will be made available on Projectplace as a template to fill in for every dataset, research project by the data owner. The text in blue and italic gives guidance on what information should be provided and should be replaced.*

General Overview	
<b>Title</b>	<i>Name of the dataset</i>
<b>PREFER task</b>	<i>Mention to which (sub)task in PREFER this dataset belongs</i>
<b>Identifier</b>	<i>An identifier will be given to all datasets. Format: PREFER_#.#_L/I/P_YYYY-mm-dd. (L, I, or P is chosen according to the design of the study: L= literature review, I= interviews, P= Patient Preference study (whatever design it takes). Example for the interviews of task 2.2: PREFER_2.2_I_2016-12-10)</i>
<b>Research Data owner</b>	<i>Names and addresses of the responsible person and deputy of the organizations who created the data; preferred format for personal names is surname first (Format: Organization; Surname, First name).</i>
<b>E-mail address of the data owner</b>	<i>Please provide the e-mail address of the data owner</i>
<b>Start and end date</b>	<i>Project start and end date. Format: YYYY.mm.dd-YYYY.mm.dd.</i>
<b>Method</b>	<i>How the data were generated, listing equipment and software used (including model and version numbers), formulae, algorithms, experimental protocols, and other things one might include in a lab notebook</i>
<b>Standards</b>	<i>Reference to existing suitable standards of the discipline can be made. If these do not exist, an outline on how and what metadata will be created. Depending of type of data, different standards for collection exist, including but not limited to:</i> <ol style="list-style-type: none"> <li><i>Systematic literature review: Cochrane and Joana Bridge institute standards</i></li> <li><i>Interviews: QUAGOL</i></li> <li><i>Focus group discussion: AMEE 91 guide</i></li> <li><i>Patient preference studies: depending on the type of method, e.g. ISPOR guide for DCE</i></li> </ol>
<b>Type of data</b>	<ul style="list-style-type: none"> <li><i>datasets containing personal data</i></li> <li><i>datasets containing non-personal data</i></li> </ul>
<b>Processing</b>	<i>How the data have been altered or processed</i>

General Overview	
<b>Source</b>	<i>Citations to data derived from other sources, including details of where the source data is held and how it was accessed</i>
<b>Funded by</b>	<i>Provide information regarding financial support such as research grants, or indicate that the data owner funds the study</i>
Content Description	
<b>Data description</b>	<i>Keywords or phrases describing the dataset or content of the data. Indicate version number if applicable. Describe the nature and origin of the data.</i>
<b>Language</b>	<i>All languages used in the dataset</i>
<b>Variable list</b>	<i>Description with variable name, length, type, etc. and code lists. Example: SEX, length of field (1 or more characters), values: F for female; M for male. DOB (Date of birth), length of field (1 or more characters), values: yyyy.mm.dd</i>
<b>Data quality</b>	<i>Data quality: This section should include description of data quality standards, procedures to assure data quality</i>
<b>Code list</b>	<i>Explanation of codes or abbreviations used in either the file names or the variables in the data files (e.g. '999 indicates a missing value in the data')</i>
Technical Description	
<b>Repository</b>	<i>Mention where the data is stored</i>
<b>File inventory</b>	<i>All files associated with the project, including extensions (e.g. 'NWPalaceTR.WRL', 'stone.mov')</i>
<b>File Formats</b>	<i>Formats of the data, e.g., FITS, SPSS, HTML, JPEG, etc. No data standards are used in general in PREFER to enable interoperability of data, but the PREFER consortium is striving to use file formats that are interoperable, such as .txt, .csv, or .rtf files.</i>
<b>File structure</b>	<i>Organization of the data file(s) and layout of the variables, where applicable</i>
<b>Checksum</b>	<i>A digest value computed for each file that can be used to detect changes; if a recomputed digest differs from the stored digest, the file must have changed</i>
<b>Necessary software</b>	<i>Names of any special-purpose software packages required to create, view, analyse, or otherwise use the data</i>
Access	
<b>Rights</b>	<i>The data owner should indicate which access rights are applicable. Any known intellectual property rights, statutory rights, licenses, or restrictions on use of the data</i>
<b>Access information</b>	<i>Where and how your data can be accessed by other researchers</i>
<b>Sharing</b>	<i>Description of how data will be shared, including access procedures, embargo periods (if any), outlines of technical mechanisms for dissemination and necessary software and other tools for enabling re-use, and definition of whether access will be widely open or restricted to specific groups. Identification of the repository where data will be stored, if already existing and identified, indicating in particular the type of repository (institutional, standard repository for the discipline, etc.). In case the dataset cannot be shared or made open access, the reasons for this should be mentioned (e.g. ethical, rules of personal data, intellectual property, commercial, privacy-related, security-related).</i>
<b>Archiving and preservation (including storage and backup)</b>	<i>Archiving and preservation (including storage and backup): Description of the procedures that will be put in place for long-term preservation of the data. Indication of how long the data should be preserved, what is its approximated end volume, what the associated costs are and how these are planned to be covered. This information should include the archiving procedure of the research project at the data owner's site and also if the data can be archived at the UU repository ALLVIS - for a detailed description see chapter 6 of the DMP.</i>

## 5.2 Responsibilities of the data owner

Data owners per task will be identified and described in table 5, which will be maintained. The data owner of the respective research projects must ensure and is responsible to comply with all legal and ethical requirements for data collection, handling, protection and storage. This includes adherence to regulations, guidelines such as (but not limited to) the EU clinical trial directive 2001/20/EC, Good clinical practice (GCP), Good Pharmacoepidemiology Practice (GPP), as applicable. Only the research data owner will be granted access to the secure data repository of KU Leuven. The process of granting access to deputies will be worked out between M6 and M18.

All data protection rules described in chapter 7 of the DMP apply to the archiving of the results underlying PREFER publications and recommendation documents. Data generated in academic-led studies which cannot be fully anonymized, e.g. interviews and personal data, may only be stored at the KUL repository described in chapter 3.

**Table 5** Overview of data owners and data repository used per task

This table will be further employed after M6 to update the research data owner including additions of research owner deputies, as people might change position. The updated table will be displayed in the next version of the DMP, due in M18.

Data owners				
Task	Design	Data repository	Research Data Owner	E-mail address
2.1	Literature review	KU Leuven	Rosanne Janssens	<a href="mailto:Rosanne.janssens@kuleuven.be">Rosanne.janssens@kuleuven.be</a>
	Interviews	KU Leuven	Rosanne Janssens	<a href="mailto:Rosanne.janssens@kuleuven.be">Rosanne.janssens@kuleuven.be</a>
2.2	Literature review	KU Leuven	Eline van Overbeeke	<a href="mailto:Eline.vanoverbeeke@kuleuven.be">Eline.vanoverbeeke@kuleuven.be</a>
	Interviews	KU Leuven	Eline van Overbeeke	<a href="mailto:Eline.vanoverbeeke@kuleuven.be">Eline.vanoverbeeke@kuleuven.be</a>
2.3	Literature review	EUR	Chiara Whichello	<a href="mailto:whichello@bmg.eur.nl">whichello@bmg.eur.nl</a>
	Interviews	KU Leuven	Chiara Whichello	<a href="mailto:whichello@bmg.eur.nl">whichello@bmg.eur.nl</a>
2.4	Literature review	EUR	Vikas Soekhai	<a href="mailto:v.soekhai@erasmusmc.nl">v.soekhai@erasmusmc.nl</a>
	Interviews	KU Leuven	Vikas Soekhai	<a href="mailto:v.soekhai@erasmusmc.nl">v.soekhai@erasmusmc.nl</a>
2.5	Literature review 1	TBD	Sarah Verschueren	<a href="mailto:Sarah.verschueren@mindbytes.be">Sarah.verschueren@mindbytes.be</a>
	Literature review 2	TBD	Selena Russo	<a href="mailto:Selena.Russo@ieo.it">Selena.Russo@ieo.it</a>
	Literature review 3	TBD	Elisabeth Furberg	<a href="mailto:Elisabeth.furberg@crb.uu.se">Elisabeth.furberg@crb.uu.se</a>
2.7	Interviews	KU Leuven	TBD	
3.3	Review of historical case studies	KU Leuven	Leo Russo	<a href="mailto:leo.russo@pfizer.com">leo.russo@pfizer.com</a>
3.3	Lessons Learned Survey		Rachel DiSantosstefano or Jorien Veldwijk	
3.4	PP case study	Industry*	TBD	
3.5	PP case study	KU Leuven	TBD	
3.6	PP case study	KU Leuven	TBD	
3.7	PP case study	KU Leuven	TBD	
3.8	PP case study	KU Leuven	TBD	
4.3	Interviews	KU Leuven	TBD	
4.4	Interviews	KU Leuven	TBD	

TBD= To be discussed; EUR= Erasmus University Rotterdam

\* The datasets containing survey data and/or recorded and transcribed interviews generated by the industry-led case studies are by definition to be regarded as personal data and require safe storage and handling in accordance with national and European regulatory frameworks. The industry partner responsible for conducting the case study will be responsible for the secure storage of the personal data.

## 6. Sharing and secondary use of PREFER generated or collected data

### 6.1 Procedures for making data findable

With the unique identifier of the individual dataset of PREFER and the overview of data owners and data repository used per task (table 5) available on Projectplace, the data owner can be identified and contacted.

### 6.2 Re-use within the PREFER consortium

To achieve the objectives of PREFER, it is imperative to follow the collaborative approach the partners agreed on when signing the consortium agreement. This includes the necessity to share data from the individual research projects while respecting data protection and intellectual property of the partners' work.

For those individual research projects within PREFER that need to use data generated in another PREFER task, table 5 contains the data owner contact details to whom a requester can reach out if they need to access the results.

### 6.3 Re-use of PREFER results by third parties

Scientific organizations all over the world are promoting a principle of open science and sharing of research data. By making data public, duplicate research can be prevented and there is a possibility to combine data. Also, money and time can be saved. The PREFER-generated data will be a valuable asset for further research.

For those external individual research projects wanting to use PREFER generated or collected data during the course of PREFER, the Data Management Compliance contact should be contacted (table 2). For those external individual research projects wanting to use PREFER generated or collected data when PREFER is completed, the Uppsala repository manager should be contacted (table 2). Giving access to external parties will be considered by the Steering Committee on a case by case basis. Access rules for the time after PREFER termination will be worked out and described in the final DMP.

Only when participants of e.g. patient preference studies or PREFER surveys agreed via informed consent that their study results may be used for secondary research and the data are anonymous, the data can be shared. To obtain the agreement of participants to use their data for secondary, research the following lines can be included in the consent form:

- *I understand the information collected about me will be stored in secure database, which will be used for future research.*
- *I authorise the research to use my anonymised study data for additional medical and/or scientific research projects.*

## 7. Protection of personal data

The collection of personal data will be conducted under the applicable international, IMI, and national laws and regulations and requires previous written informed consent by the individual, i.e., with public and commercial entities and if applicable outside the EU in countries with lower data protection standards. To obtain the agreement of participants of e.g. patient preference studies or PREFER surveys to use their data for secondary, research the following lines can be included in the consent form:

- *I understand the information collected about me will be stored in secure database, which will be used for future research.*
- *I authorise the research to use my anonymised study data for additional medical and/or scientific research projects.*

PREFER researchers commit to the highest standards of data security and protection in order to preserve the personal rights and interests of study participants. They will adhere to the provisions set out in the:

- General data protection regulation (GDPR), foreseen coming into effect in 2018(8)
- Directive 2006/24/EC of 15 March 2006 on the retention of data generated or processed in connection with the provision of publicly available electronic communication services or of public communications networks(9)
- Directive 2002/58/EC of the European Parliament and of the Council of 12 July 2002 concerning the processing of personal data and the protection of privacy in the electronic communications sector (Directive on privacy and electronic communications)(10)
- Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data(11)

Prior to collecting, storing, and processing sensitive personal data, the consortium will seek approval of the applicable local and/ or national data protection authorities and work within the processes recommended in the e-Health Task Force Report “Redesigning Health in Europe for 2020.”

Consent forms will contain information on how personal data will be managed.

To secure the confidentiality, accuracy, and security of data and data management, the following measures will be taken:

- All personal data obtained within the academic-led case studies will be transmitted to partners within the consortium only after anonymization or pseudonymization. Keys to identification numbers will be held confidentially within the respective research units. In situations where re-identification of study participants becomes necessary, for example the collection of additional data, this will only be possible through the research unit and in cases where informed consent for such cases has been given.
- Personal data are entered to secure websites. Data are processed only for the purposes outlined in the patient information and informed consent forms of the respective case studies. Use for other purposes will require explicit patient approval. Also, data are not transferred to any places outside the consortium without patient consent.
- Access to experimental data will be granted to partners in non-EU countries for restricted use within the PREFER project. Data handling in non-EU countries will be fully conforming to national laws and regulations and the European Directive 95/46/EC. In cases of contradiction, the tighter regulation shall prevail. The necessary and legally adequate measures will be taken to ensure that the data protection standards of the EU shall be complied with (see below). Transfer and subsequent use of PREFER data by partners in US will be governed in accordance with federal and state laws.
- None of the personal data will be used for commercial purposes, but the knowledge derived from the research using the personal data may be brought forward to such use as appropriate, and this



process will be regulated by the Grant Agreement and the Consortium Agreement, in accordance with any generally valid legislation and regulations.

The following points to consider will guide the protection of data within the PREFER project:

(i) The entity providing personal data to the project shall verify that:

- the initial collection of these data has been compliant with the requirements of the original purpose
- the collection and the provision of the data to the project meets all legal requirements to which the entity is subject
- further storage and processing of the data after completion of the research project is in compliance with applicable law

(ii) The entity which provides personal data to the project shall document any restriction of use or obligation applicable to these data (e.g., the limited scope of purpose imposed by the consent form)

The entity which uses personal data in the project shall be responsible to ensure that it has the right under the applicable data protection and other laws to perform the activities contemplated in the project.

Personal data shall always be collected, stored, and exchanged in a secure manner, through secure channels.

## 8. Ethical aspects

### 8.1 General ethical aspects

The participants of PREFER are requested to adhering to all relevant international, IMI, and national legislation and guidelines relating to the conduct of prospective case studies as detailed below.

All research activities within PREFER requiring approval on ethical and legal grounds through responsible local or national Ethics Committees and Regulatory Authorities will be conducted only after obtaining such approval. All ethics approvals will be submitted to IMI before commencement of any prospective case study. A report by the Ethics Advisory Board will be submitted to IMI within the periodic reports.

The proposed research will comply with the highest ethical standards, including those outlined in the Grant Agreement (Article 34 of the Model Grant Agreement) and the European Code of Conduct for Research integrity. The balance between the research objectives and the means used to achieve them will be given special attention. To ensure this, PREFER is supported by its Ethical Advisory Board. The Ethical Advisory Board will consist of four experts on ethics, law, and drug development representing the key areas of the project, including a patient representative. The Ethical Advisory Board will monitor the progress of the project and ensure a high standard of research by taking part in the annual General Assembly meetings. In addition, it will:

- provide expert support to the consortium in all relevant ethical questions
- ensure compliance with legislation and guidelines
- conduct regular project reviews
- issue recommendations to the consortium when appropriate

Researchers are requested to have appropriate training regarding Good Scientific, Good Clinical, Good Pharmacoepidemiology Practice Guidelines and the legal and regulatory framework described in the following sections.

### 8.2 Interviews and patient preference studies

The methodologies for eliciting patient preferences will be tested in prospective case studies. At this stage, it is not yet fully decided which patient populations will be involved in the case studies, but we foresee the possibility of approaching vulnerable patient populations, children, parents, care givers, and healthy volunteers. Each patient preference study requires approval from the relevant ethical review boards with adherence to requirements related to informed consent and protection of privacy.

Our foremost principles for the conduct of any research involving human participants within PREFER are:

- respect for the rights, integrity, and privacy of patients
- protection of vulnerable patients
- continuous monitoring of patients' safety
- generation of meaningful, high-quality data
- timely publication of case study results

All research in PREFER involving human participants will be conducted under the applicable international, IMI, and national laws and regulations and only after obtaining approval by the applicable local or national Ethics Committees and Regulatory Authorities. In particular, the consortium is committed to:

- the Declaration of Helsinki: Ethical Principles for Medical Research Involving Human Subjects (Adopted by the 18th World Medical Association (WMA) General Assembly, Helsinki, Finland, June 1964, and last amended by the 64th WMA General Assembly, Fortaleza, Brazil, October 2013)(12)
- the standards of the International Conference on Harmonisation on Good Clinical Practice(13)
- the Convention for the Protection of Human Rights and Dignity of the Human Being with regard to the Application of Biology and Medicine: Convention on Human Rights and Biomedicine, ETS No. 164, Oviedo, 4 April 1997; and the Additional Protocol on Biomedical Research (CETS No. 195), 2005(14)

- the UNESCO: Universal Declaration on Bioethics and Human Rights (2005)(15)
- Case studies have not yet been defined in detail so at this stage it is unclear which countries will be involved. Research with human participants will be conducted in the applicable countries in accordance with national and international regulations. Preference studies regarding future risks or investigating how to balance benefits and risks may cause psychological distress e.g., for vulnerable patient groups. This implies that all studies conducted within the PREFER project will have to take actions in order to be able to support/counsel patients appropriately. This will be one of the requirements assigned to each leader of a clinical case study.
- As mentioned above PREFER will seek to include a broad selection of patient populations, including vulnerable patients if necessary. For ethical reasons it is important that perspectives from these patient groups are also included, and that patients who may experience certain difficulties to get their voice heard and their preferences taken into account. Vulnerable patient populations may be identified in the field of Neuromuscular disorders where many of the diagnosed diseases are rare and the patients are not adults. This is also why the PREFER project has included a patient organization within this disease area, i.e. Muscular Dystrophy UK. They, as well as the other patient organisations, will be asked to give extra attention to the situation of vulnerable patients and the how they are included in the case studies.

Patient Information and informed consent procedures will be approved by the relevant national or local ethics boards. Data collectors collecting personal data for a prospective collaborative research project will inform the study participants about the project in an appropriate manner, including:

- the identity of the data controller
- the voluntariness of the collection of data
- the purposes of the processing
- the nature of the processed data, including its type (identifiable, coded, anonymised)
- the handling of the data
- the existence of the right of access to, and the right to rectify the data concerning themselves
- if the research project reasonably anticipates the sharing of data across research groups (including academic and commercial entities) and national borders (including information about potentially lower data protection standards outside EU)
- if the project involves collaboration with both academic and commercial partners
- that consent may be withdrawn and how this is done

The research conducted in PREFER does not have the potential for malevolent/criminal/terrorist abuse. There are no other ethics issues currently identified beyond those discussed above. Any potential issues that arise during the project duration will be presented to the Ethics Advisory Board who will ensure they are addressed by taking the appropriate organisational, legal, and regulatory steps.

## 9. References

1. Directive 2001/83/EC of the European Parliament and of the Council on the community code relating to medicinal products for human use; 2001 [cited 8 February 2017]. Available from: [http://ec.europa.eu/health/sites/health/files/files/eudralex/vol-1/dir\\_2001\\_83\\_consol\\_2012/dir\\_2001\\_83\\_cons\\_2012\\_en.pdf](http://ec.europa.eu/health/sites/health/files/files/eudralex/vol-1/dir_2001_83_consol_2012/dir_2001_83_cons_2012_en.pdf).
2. Council directive 93/42/EEC concerning medical devices; 1993 [cited 8 February 2017]. Available from: <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CONSLEG:1993L0042:20071011:en:PDF>.
3. Consortium Agreement for [Patient Preferences In Benefit-Risk Assessments during the Drug Life Cycle (PREFER). Innovative Medicines Initiative 2 Joint Undertaking; 2015.
4. Annex 1 of the Grant Agreement no.: 115966 - Description of Action (DoA), Patient Preferences in Benefit-Risk Assessments during the Drug Life Cycle (PREFER). Innovative Medicines Initiative 2 Joint Undertaking; 2016.
5. The European Commission D-GfRI. Guidelines on FAIR Data Management in Horizon 2020. 2016.
6. Guidance for Industry, Electronic Source Data in Clinical Investigations. U.S. Department of Health and Human Services Food and Drug Administration, Center for Drug Evaluation and Research (CDER), Center for Biologics Evaluation and Research (CBER), Center for Devices and Radiological Health (CDRH); 2013 [cited 15 February 2017]. Available from: <http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM328691.pdf>.
7. [dmptool.org/dm\\_guidance](http://dmptool.org/dm_guidance) [page on the Internet] Data Management General Guidance, DMP Tool; 2017 [cited 20 February 2017]. Available from: [https://dmptool.org/dm\\_guidance](https://dmptool.org/dm_guidance).
8. Regulation (EU) 2016/679 of the European Parliament and of the Council on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation); 2016 [cited 15 February 2017]. Available from: <http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679&from=EN>.
9. Directive 2006/24/EC of the European Parliament and of the Council on the retention of data generated or processed in connection with the provision of publicly available electronic communications services or of public communications networks and amending Directive 2002/58/EC; 2006 [cited 15 February 2017]. Available from: <http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32006L0024&from=en>.
10. Directive 2002/58/EC of the European Parliament and of the Council concerning the processing of personal data and the protection of privacy in the electronic communications sector (Directive on privacy and electronic communications); 2002 [cited 15 February 2017]. Available from: <http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32002L0058&from=en>.
11. Directive 95/46/EC of the European Parliament and of the Council on the protection of individuals with regard to the processing of personal data and on the free movement of such data; 1995 [cited 15 February 2017]. Available from: <http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:31995L0046&from=en>.
12. Declaration of Helsinki: Ethical Principles for Medical Research Involving Human Subjects [Adopted by the 18th World Medical Association (WMA) General Assembly, Helsinki, Finland, June 1964, and last amended by the 64th WMA General Assembly, Fortaleza, Brazil, October 2013]; 1964 [cited 20 February 2017]. Available from: <http://www.wma.net/en/30publications/10policies/b3/>.
13. Guideline For Good Clinical Practice E6(R1). International Conference On Harmonisation Of Technical Requirements For Registration Of Pharmaceuticals For Human Use; 1996 [cited 20 February 2017]. Available from: [http://www.ich.org/fileadmin/Public\\_Web\\_Site/ICH\\_Products/Guidelines/Efficacy/E6/E6\\_R1\\_Guideline.pdf](http://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Efficacy/E6/E6_R1_Guideline.pdf).
14. Convention for the Protection of Human Rights and Dignity of the Human Being with regard to the Application of Biology and Medicine: Convention on Human Rights and Biomedicine. Council of Europe; 1997 [cited 20 February 2017]. Available from: <https://rm.coe.int/CoERMPublicCommonSearchServices/DisplayDCTMContent?documentId=090000168007cf98>.
15. Universal Declaration on Bioethics and Human Rights. UNESCO; 2005 [cited 20 February 2017]. Available from: <http://unesdoc.unesco.org/images/0014/001461/146180e.pdf>.

## 10. List of abbreviations

Abbreviation	Full term
ALCOA	Attributable, legible, contemporaneous, original and accurate
Ca	Category
CentOS	Community Enterprise Operating System
CETS	Council of Europe Treaty Series
CO	Confidential, restricted under conditions set out in Model Grant Agreement
CPU	Central processing unit
DEC	Websites, patents filing, press & media actions, videos, etc.
DLC	Drug Life Cycle
DMP	Data Management Plan
DNS	Domain Name System
DoA	Description of Action
EC	European Commission
EU	European Union
EUR	Erasmus University Rotterdam
FAIR	Findable, accessible, interoperable and reusable
FTP	File Transfer Protocol
GB	Gigabyte
GCP	Good Clinical Practice
GDPR	General Data Protection Regulation
GPP	Good Pharmacoepidemiology Practice
HTA	Health Technology Assessment
ICTS	Information and communications technology services
IP address	Internet Protocol address
KUL	University of Leuven (KU Leuven)
PI	Principal Investigator
PP	Patient preferences
PREFER	Patient Preferences in Benefit-Risk Assessments during the Drug Life Cycle
PU	Public, fully open, e.g. web
R	Document, report (excluding the periodic and final reports)
RAM	Random-access memory
RDP	Remote desktop protocol
SAS	Statistical Analysis Software
SFTP	Secure File Transfer Protocol
SMB	Server Message Block
SPSS	Statistical Package for the Social Science
SSH	Secure Shell
TBA	To be announced
TBD	To be discussed
UK	United Kingdom
UNESCO	United Nations Educational, Scientific and Cultural Organization
UU	Uppsala University
VPN	Virtual private network
WMA	World Medical Association
WP	Work packages

## 11. GLOSSARY

Term	Explanation
Access	The general principles on access rules are defined in the consortium agreement (section 8).(3) These rules are different for research results to exploit and research results to disseminate and are further explained in chapter 2 of this DMP.
Anonymization	type of information sanitization whose intent is privacy protection. It is the process of either encrypting or removing personally identifiable information from data sets, so that the people whom the data describe remain anonymous.
Born digital (data type)	The term born-digital refers to materials that originate in a digital form. This is in contrast to digital reformatting, through which analog materials become digital.
CentOS	CentOS (from Community Enterprise Operating System) is a Linux distribution that attempts to provide a free, enterprise-class, community-supported computing platform functionally compatible with its upstream source, Red Hat Enterprise Linux (RHEL).
Data owner	The Consortium Agreement regulates the ownership and access to key knowledge generated in the project.(3) Each dataset that is or will be generated will have an owner responsible for its creation and/or management (registered as creator in the metadata). The responsibilities of the data owner are explained in chapter 5.2.
Domain Name System (DNS)	DNS stands for Domain Name System, which is the largest digital database in the world, containing information about every web site on the internet. Every web site online has an IP address that is its actual internet location, and this number is used to locate the web site within the database. The data that tells the web server how to respond to your input is known as the DNS records, or zone files. These records play a vital role in the functionality of the internet, and any aspiring internet technology expert should learn the following facts about DNS records and how they are used.
IP address	An IP address (abbreviation of Internet Protocol address) is an identifier assigned to each computer and other device (e.g., printer, router, mobile device, etc.) connected to a TCP/IP network that is used to locate and identify the node in communications with other nodes on the network.
Long-term data storage	Storage of PREFER data after the end of the PREFER project.
Model (data format)	Data generated in modelling studies.
Multimedia (data format)	The multimedia data include one or more primary media data types such as text, images, graphic objects (including drawings, sketches and illustrations) animation sequences, audio and video. Multimedia data consists of a variety of media formats or file representations including TIFF, BMP, PPT, IVUE, FPX, JPEG, MPEG, AVI, MID, WAV, DOC, GIF, EPS, PNG, etc. Because of restrictions on the conversion from one format to the other, the use of the data in a specific format has been limited as well. Usually, the data size of multimedia is large such as video; therefore, multimedia data often require a large storage.
Numerical (data format)	Files that contain decimal numbers and integer numbers (including negative values).
Observational (data type)	Data generated in observational research.
Personal data	Any information relating to an identified or identifiable natural person ('Data Subject'); an identifiable person is one who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity.
Pseudonymization	A procedure by which the most identifying fields within a data record are replaced by one or more artificial identifiers, or pseudonyms. There can be a single pseudonym for a collection of replaced fields or a pseudonym per replaced field.
Random-access memory (RAM)	Random-access memory is a form of computer data storage which stores frequently used program instructions to increase the general speed of a system.
Receiver	Non-public datasets may be transferred from the owner to a receiver in order to carry out the activities of the project.
Reference (data type)	In programming language theory, a reference type is a data type that refers to an object in memory. In Java a reference data type is a variable that can contain the reference or

	an address of dynamically created object. In PREFER the data that classify as reference data are the references of the articles used for literature reviews.
Remote desktop protocol (RDP)	Remote Desktop Protocol (RDP) is a proprietary protocol developed by Microsoft, which provides a user with a graphical interface to connect to another computer over a network connection. The user employs RDP client software for this purpose, while the other computer must run RDP server software.
Secure Shell (SSH)	Secure Shell (SSH) is a cryptographic network protocol for operating network services securely over an unsecured network. The best known example application is for remote login to computer systems by users.
Sensitive personal data	Sensitive personal data consists of information relating to the data subject with regard to racial or ethnic origin; political opinions; religious beliefs or other beliefs of a similar nature; trade union membership; physical or mental health or condition; sexual life; the commission or alleged commission by the data subject of any offence; or any proceedings for any offence committed or alleged to have been committed by the data subject, the disposal of such proceedings or the sentence of any court in such proceedings.(8)
Server Message Block (SMB) Protocol	A network file sharing protocol, and as implemented in Microsoft Windows is known as Microsoft SMB Protocol. The set of message packets that defines a particular version of the protocol is called a dialect. The Common Internet File System (CIFS) Protocol is a dialect of SMB.
Simulation (data type)	Data generated in simulation studies.
Textual (data format)	A text file (sometimes spelled "textfile": an old alternative name is "flatfile") is a kind of computer file that is structured as a sequence of lines of electronic text.
Virtual private network (VPN)	A virtual private network (VPN) extends a private network across a public network, such as the Internet. It enables users to send and receive data across shared or public networks as if their computing devices were directly connected to the private network. Applications running across the VPN may therefore benefit from the functionality, security, and management of the private network.