DOI: 10.4126/FRL01-006483994
Version: V1_0
Release Date: 22.08.2024

Research Data Management Plan — Epidemiological Studies

Inhaltsverzeichnis

This RDMP was created by		on
The RDMP was updated on	by [Name, Department]	
The RDMP was updated on	by [Name, Department]	
The RDMP was updated on	by [Name, Department]	
The RDMP was last updated on	by [Name, Department]	

No.	Question	Answer	
1	General information on research proposal		
1.1	What is the title of the study? (baseline study or further study as part of the baseline study)	The title of the study is …	
1.2	What is the primary research question?	The primary research question is …	
1.3	Which institution or individual is responsible for project management and coordination?	◯ Personal name	◯ Organizational name
		Project manager / sponsor (primary)	
		Name of the affiliation:	
		Address of the affiliation:	
		Webpage of the affiliation:	
		ROR ID of the affiliation:	
1.4	Which project partners are there (institutions, persons)?	◯ Personal name	◯ Organizational name
		Personal name (if applicable)	
		Type of role of partner:	Other
		Name of the affiliation:	

		Address of the affiliation:	
		Webpage of the affiliation:	
		ROR ID of the affiliation:	
1.5	What is the project duration? When does the project start and end?	The project is planned to run / has received funding for [Number of months]. The project will start / has started on and will end on .	
1.6	What DFG discipline does the study belong to?	The study belongs to the following disciplines:	
		◯ 204 Microbiology, Virology und Immunology	
		◯ 205 Medicine	
		◯ 206 Neurosciences	
		Please specify further:	
1.7	Are there rules or guidelines for handling research data and/or data management requirements in these disciplines, at your institution, from the sponsor or for the study? If yes, briefly outline these and refer to further	On institutional level:	
		FDM policy	
		On discipline-level:	
		Good epidemiological practice	
		Good clinical practice	

	information if necessary. To what extent are they binding (e.g. obligation to publish study metadata in a register, sharing of personal data on request)?	ICH	
		Declaration of Helsinki	
		Other:	
1.8	Who is funding the study?	The study / project is founded by:	
		Name of Funder:	
		Funding identifier:	

2 Types of data – content

2.1	What data types are generated?	☐ Administrative databases ☐ Biological samples ☐ Cognitive measurements ☐ Genealogical records ☐ Imaging data ☐ Interview ☐ Medical records ☐ Omics technology ☐ Physiological/Biochemical measurements ☐ Questionnaire ☐ Registries ☐ Other	☐ Biosamples: ☐ Blood ☐ Buccal cells ☐ Cord blood ☐ DNA ☐ Faeces ☐ Hair ☐ Immortalized cell lines ☐ Isolated pathogen ☐ Nail ☐ Plasma ☐ RNA ☐ Saliva ☐ Serum ☐ Tissue (FFPE) ☐ Tissue (frozen) ☐ Urine ☐ Other biological samples	☐ Imaging data: ☐ Computed tomography (CT) ☐ Magnetic resonance imaging (MRI) ☐ Radiography (x-ray) ☐ Ultrasound ☐ Other imaging data ☐ Omics technology: ☐ Biomarkers ☐ Genomics ☐ Metabolomics ☐ Proteomics ☐ Transcriptomics ☐ Other omics technology
2.2	Are data reused?	◯ Yes ◯ No -> skip question 2.2.a + 2.2.b		
2.2.a	Which data are reused?	☐ Administrative databases ☐ Biological samples ☐ Cognitive measurements	☐ Biosamples: ☐ Blood ☐ Buccal cells	☐ Imaging data: ☐ Computed tomography (CT) ☐ Magnetic resonance imaging

☐ Genealogical records ☐ Imaging data ☐ Interview ☐ Medical records ☐ Omics technology ☐ Physiological/Biochemical measurements ☐ Questionnaire ☐ Registries ☐ Other	☐ Cord blood ☐ DNA ☐ Faeces ☐ Hair ☐ Immortalized cell lines ☐ Isolated pathogen ☐ Nail ☐ Plasma ☐ RNA ☐ Saliva ☐ Serum ☐ Tissue (FFPE) ☐ Tissue (frozen) ☐ Urine ☐ Other biological samples	(MRI) ☐ Radiography (x-ray) ☐ Ultrasound ☐ Other imaging data ☐ Omics technology: ☐ Biomarkers ☐ Genomics ☐ Metabolomics ☐ Proteomics ☐ Transcriptomics ☐ Other omics technology

2.2.b Are the re-used data publicly available?

○ No

○ Yes, some of the re-used data are publicly available here:

○ Yes, all re-used data are publicly available here:

DOI:	
URL:	
arXiv:	
EAN13:	
EISSN:	
Handle:	
ISBN:	
ISSN:	
ISTC:	
LISSN:	
LSID:	
PMID:	
PURL:	
URN:	
w3id:	
Other:	

2.3 Which data are not digitally available?

☐ Administrative databases ☐ Biological samples ☐ Cognitive measurements ☐ Genealogical records ☐ Imaging data ☐ Interview ☐ Medical records ☐ Omics technology ☐ Physiological/Biochemical measurements ☐ Questionnaire ☐ Registries ☐ Other	☐ Biosamples: ☐ Blood ☐ Buccal cells ☐ Cord blood ☐ DNA ☐ Faeces ☐ Hair ☐ Immortalized cell lines ☐ Isolated pathogen ☐ Nail ☐ Plasma ☐ RNA ☐ Saliva ☐ Serum ☐ Tissue (FFPE) ☐ Tissue (frozen) ☐ Urine ☐ Other biological samples	☐ Imaging data: ☐ Computed tomography (CT) ☐ Magnetic resonance imaging (MRI) ☐ Radiography (x-ray) ☐ Ultrasound ☐ Other imaging data ☐ Omics technology: ☐ Biomarkers ☐ Genomics ☐ Metabolomics ☐ Proteomics ☐ Transcriptomics ☐ Other omics technology

3 Data Formats

3.1 In which file formats and volumes are the data types created?

☐ Text documents ☐ .pdf ☐ .docx ☐ .rtf ☐ .odt Data Volume:	☐ Plain text ☐ .txt Data Volume:

☐ Markup language ☐ .xml ☐ .html ☐ .css Data Volume:	☐ Audio ☐ .bwf ☐ .mxf ☐ .mka ☐ .flac ☐ .wav ☐ .mp3 ☐ .m4a Data Volume:
☐ Programming languages ☐ Python ☐ Java ☐ MATLAB ☐ NetCDF Data Volume:	☐ Video ☐ .mxf ☐ .mkv ☐ .mp4 ☐ .m4a ☐ .mpg ☐ .avi Data Volume:
☐ Spreadsheets ☐ .ods ☐ .csv ☐ .xlsx ☐ .pdf Data Volume:	☐ Computer Aided Design (CAD) ☐ .dxf ☐ .svg ☐ .dxf ☐ .dwg Data Volume:

☐ Databases ☐ .sql ☐ .siard ☐ .csv Data Volume:	☐ Geographical Information Systems (GIS) ☐ .gml ☐ .mif/.mid ☐ .json Data Volume:
☐ Statistical data ☐ .dat/.sps ☐ .dat/.DO ☐ .csv/.html ☐ R ☐ .sav Data Volume:	☐ Georeferenced images ☐ .tif/.tiff Data Volume:
☐ Raster images ☐ .jpg/.jpeg ☐ .tif/.tiff ☐ .png ☐ .dcm Data Volume:	☐ 3D ☐ .obj ☐ .ply ☐ .x3d ☐ .dae ☐ .pdf Data Volume:

		☐ Vector images ☐ .svg ☐ .ai ☐ .eps ☐ .cdr Data Volume:	☐ RDF ☐ .rdf ☐ .trig ☐ .ttl ☐ .nt Data Volume:
3.2	What is the data volume of all the digital data types to be collected?	[Number of] KB / MB	
3.3	What is the expected data volume per year?	[Number of] KB / MB	
3.4	Which instruments, software, tools are used to generate data?	The following tools / software / instruments are used to generate the data …	
3.5	What software, processes or technologies are required to analyse the data?	The following software / tools are needed to analyse the data …	
3.6	What software or analysis scripts will be necessary in order to be able to reuse data?	The following software / tools are needed to reuse the data …	
		Scripts:	

3.7 Where is the data saved during the research project?	Software:	
	Documentation:	
	local server	
	cloud service	
	if applicable, insert URL:	
3.8 Are there internal project guidelines for the uniform data organisation?	◯ There are no project-internal guidelines for the uniform organisation of data.	
	◯ There are project-internal guidelines for the uniform organisation of data:	
	Storage space	
	Folder structure	
	File naming	
	If applicable: insert LINK to guidelines	

3.9	Who is authorised to access which data during the research project?	Authorised person/group:	
		Type of access:	
		Type of accessible data:	
3.10	Who backs up the data and how often?	IT	How often:
		Researchers	How often:
		Other	How often:
3.11	What measures are taken to ensure data security?	Protection against unauthorized access by:	

Transfer of sensitive data is protected by:
Data recovery by means of:
Other:

4 Interoperability/Metadata

4.1 Which metadata standards, ontologies, classifications etc. are used to describe the data?

The following metadata standards / ontologies / classification are used to describe the data:	
☐ MeSH	
☐ ICD-10	
☐ ICD-11	
☐ MedDRA	
☐ SNOMED CT	
☐ Other vocabulary:	

		☐ None	
4.2	Which data are not interoperable? For which data types/content are no (common) standards or project-specific (or rare) ontologies, metadata schemas, vocabularies etc. used?	Describe the type of data:	
4.3	Will mappings be created for common metadata standards?	If applicable, name the standard:	
4.4	Are metadata collected automatically (e.g. from devices)? Which ones?	◯ Yes	
		Instruments:	
		Software:	
		Variables:	
		◯ No	

5 Quality assurance

5.1	Are metadata and contextual information checked for correctness and completeness?	◯ Metadata and contextual information are not checked for completeness and correctness.	
		◯ Metadata and contextual information are checked for completeness and correctness.	
5.2	Who is responsible for documenting and checking the metadata and context information for accuracy and completeness?	Metadata and context information are checked for completeness and correctness by:	
		Responsible Person / Department:	
5.3	What quality assurance measures are taken for the generated data?	☐ **Data completeness**	The following person/s is/are responsible for ensuring data quality during the indicated phases:
		☐ Check whether all information is included	
		☐ Check whether any files are missing	
		☐ **Data consistency**	
		☐ Check whether structure and formatting are consistent throughout the dataset	
		☐ **Data integrity**	
		☐ Check whether the data content is correct and complete	
		☐ There is a procedure in case of missing or incorrect data	

Procedure:	
☐ Check for duplicates	
☐ Data is supplemented or enriched	
Supplementation / Enriching:	
☐ **Data collection**	
Responsible person / position:	
Measures taken:	
☐ **Data compiling**	
Responsible person / position:	
Measures taken:	

☐ **Data processing**	
Responsible person / position:	
Measures taken:	

6 Persistent Identifiers (PIDs)

6.1 Which PID system(s) is/are used?

- ☐ DOI
- ☐ URL
- ☐ arXiv
- ☐ EAN13
- ☐ EISSN
- ☐ Handle
- ☐ ISBN
- ☐ ISSN
- ☐ ISTC
- ☐ LISSN
- ☐ LSID
- ☐ PMID
- ☐ PURL
- ☐ URN
- ☐ w3id
- ☐ Other

6.2 Which data are given its own identifier?

Raw data:	
Cleaned data:	

Analysed data:	
Processed data:	
Published data:	

7 Laws and ethics

7.1 Which law is relevant for the project with regard to personal data protection issues?

7.2 Does the data set contain sensitive data (BDSG §3, para.9))?

- ◯ No
- ◯ Yes, information on:
  - ☐ racial and ethnic origin
  - ☐ political opinions
  - ☐ religious or philosophical beliefs
  - ☐ trade union membership
  - ☐ health
  - ☐ sex life
- ☐ other non-personal sensitive data:

7.3	Will the sensitive data anonymised be pseudonymised?	◯ Personal data will be anonymized
		◯ Personal data will be pseudonymised
		◯ Undecided
7.4	Is the "informed consent" of the data subjects obtained? What research purposes does informed consent cover?	Research purposes covered by informed consent:
		Where are the informed consent documents stored:
7.5	Are data used that are protected by copyright?	◯ Yes
		◯ No
7.6	Is there a license or usage agreement for the re-used data specified in 2.2.a?	◯ No
		◯ Yes, the following license / usage agreement applies:

7.7	Who is the ethics committee?	The Ethics Committee of:
7.8	Who decides if data access is granted in response to requests?	◯ There is no data access committee. ◯ There is a data access committee that decides whether access is granted for access requests. The committee consists of [Person / Department]:

8 Reuse and publication of data

8.1	Where will the metadata and/or study documents be published?	Institutional repository:	
		Specialised repository:	
		Generic repository:	

8.2 Which data will be published?

- Study metadata
- Study documents, which ones?
  - - Instruments (templates), which ones?

Data Journal:	
Other:	
Study metadata will be made publicly available on:	
Repository:	
Link:	
The following study documents/instruments will be made available on: Repository, Link	
Substudy/Data collection event	
Biobank	
Case report form	
Code book	
Data dictionary	

Data management plan	
Dataset	
Discussion guide	
Informed consent form	
Interview scheme and themes	
Manual of operations (SOPs)	
Observation guide	
Participant tasks	
Patient information sheet	
Questionnaire	
Registry	
Secondary data source	

Statistical analysis plan	
Study protocol	
Other data collection instrument	
Other study document	
Other	

8.3 Under which terms of use / which licence should the data be published?

◯ Creative Commons 4.0
Please specify:
◯ Individual data usage contracts

8.4 Are there embargo periods (for political/commercial/patent law reasons)? If yes, please provide details.

◯ Yes:
◯ No

8.5	Will individual patient data (IPD) be made available on request?	◯ Yes ◯ No ◯ Undecided
8.5.a	Where will IPD be made available on request?	The IPD will be made available upon request here: [LINK]
8.5.b	Which individual patient data (IPD) will be made available (on request)?	The following IPD will be made available:

9 Storage and long-term archiving

9.1	Where are the data (including metadata, documentation and any relevant code or software) stored or archived?	◯ The data will be archived in a public repository	
		Name of repository:	

URL:	
◯ The data will be archived in a non-public repository (i.e., the data's existence is verifiable, but data is only accessible on request)	
Name of repository:	
URL:	
◯ The data will be archived in a storage service, integrity and accessibility of the data are guaranteed	
Name of repository:	
URL:	

9.2 Which data will be archived for the long term?

Study metadata:	
Study documents:	
Instruments (templates):	

IPDs:	
Software/Scripts:	

Basis: RDMO catalogue of questions (generic catalogue) of 19.11.2019