FWO DMP Template - Flemish Standard Data Management Plan

Version KU Leuven

Project supervisors (from application round 2018 onwards) and fellows (from application round 2020 onwards) will, upon being awarded their project or fellowship, be invited to develop their answers to the data management related questions into a DMP. The FWO expects a **completed DMP no later than 6 months after the official start date** of the project or fellowship. The DMP should not be submitted to FWO but to the research co-ordination office of the host institute; FWO may request the DMP in a random check.

At the end of the project, the **final version of the DMP** has to be added to the final report of the project; this should be submitted to FWO by the supervisor-spokesperson through FWO's e-portal. This DMP may of course have been updated since its first version. The DMP is an element in the final evaluation of the project by the relevant expert panel. Both the DMP submitted within the first 6 months after the start date and the final DMP may use this template.

The DMP template used by the Research Foundation Flanders (FWO) corresponds with the Flemish Standard Data Management Plan. This Flemish Standard DMP was developed by the Flemish Research Data Network (FRDN) Task Force DMP which comprises representatives of all Flemish funders and research institutions. This is a standardized DMP template based on the previous FWO template that contains the core requirements for data management planning. To increase understanding and facilitate completion of the DMP, a standardized **glossary** of definitions and abbreviations is available via the following link.

1. General Project Information	
Name Grant Holder & ORCID	**Greet Vanderlinden - ORCID 0000-0002-5673-4121**
Contributor name(s) (+ ORCID) & roles	**Koen Van Laere - promotor**
Project number [1] & title	**Integrative longitudinal investigation of multiparametric PET/MR imaging in mild cognitive impairment**
Funder(s) GrantID [2]	**PDMT2/24/076**
Affiliation(s)	☐ **KU Leuven** ☐ Universiteit Antwerpen ☐ Universiteit Gent ☐ Universiteit Hasselt ☐ Vrije Universiteit Brussel ☐ Other: ROR identifier KU Leuven: 05f950310
Please provide a short project description	Alzheimer's disease (AD) presents a pressing global health concern, particularly given the demographic trend of population aging. Despite the fact that AD's neuropathological hallmarks have long been described, the exact interplay between amyloid-β, tau and neurodegeneration remains unclear. Recent advancements in PET imaging offer enhanced specificity and sensitivity for detecting tau pathology as well as loss of synaptic density in AD. Moreover, state-of-the-art diffusion MRI methods now allow quantification of the brain's white matter fibre populations within an image voxel. However, longitudinal investigations employing these imaging modalities are still scarce. We will integrate PET, MRI, and neuropsychological evaluations to longitudinally study healthy aging and also to follow disease progression in patients with mild cognitive impairment. This will elucidate the interplay between molecular changes and structural/functional brain alterations, and it will help us to investigate how their combination eventually results in cognitive decline.

2. Research Data Summary

[1] "Project number" refers to the institutional project number. This question is optional. Applicants can only provide one project number.

[2] Funder(s) GrantID refers to the number of the DMP at the funder(s), here one can specify multiple GrantIDs if multiple funding sources were used.

List and describe all datasets or research materials that you plan to generate/collect or reuse during your research project. For each dataset or data type (observational, experimental etc.), provide a short name & description (sufficient for yourself to know what data it is about), indicate whether the data are newly generated/collected or reused, digital or physical, also indicate the type of the data (the kind of content), its technical format (file extension), and an estimate of the upper limit of the volume of the data [3].

				ONLY FOR DIGITAL DATA	*ONLY FOR DIGITAL DATA*	*ONLY FOR DIGITAL DATA*	*ONLY FOR PHYSICAL DATA*
Dataset Name	Description	New or Reused	Digital or Physical	Digital Data Type	Digital Data Format	Digital Data Volume (MB, GB, TB)	Physical Volume
ICF	Inform Consent forms	New	Physical	N.A.	N.A.	N.A.	1 ICF per participant, estimated 30 participants (15 patients and 15 healthy controls) to participate in follow-up.
Demographical data	Demographical data on participants: age, sex, educational level,..	Reuse since participants already included at baseline	Digital	Numerical Textual	.xls .xml (redcap)	<1 GB	N.A.
Data on medical history healthy controls	Data on medical history from baseline inclusion to new follow-up assessment	New	Physical	N.A.	N.A.	N.A.	5 pages form/healthy control (estimated 15 subjects)
Neuropsychological test data	Data of neuropsychological tests (MMSE, RAVLT, TMT A and B, AVF, RCPM, BNT, BDI, GDS)	New	Physical	N.A.	N.A.	N.A.	12 pages/subject
Raw imaging data	Raw, unpreprocessed PET and MRI images	New	Digital	Images	.dcm .nii.gz	<5 TB	N.A.

[3] Add rows for each dataset you want to describe.

					.json		
Processed imaging data	Processed PET and MRI images from which output can be generated	New	Digital	Images	.nii.gz .nii	>5 TB	N.A.
Scripts	Scripts used for image processing and for statistical analysis	Reuse scripts for PET processing+ New scripts for statistics and MRI processing	Digital	Numerical Software	.py .sh .mat .R	<1 GB	N.A.
Results	Results/output of the processed image data	New	Digital	Numerical Textual	.txt .xls	<1 GB	N.A.
Reports	Papers and presentations of the results	New	Digital	Numerical Textual	.pptx .docx .pdf	<1 GB	N.A.

GUIDANCE:
The data description forms the basis of your entire DMP, so make sure it is detailed and complete. It includes digital and physical data and encompasses the whole spectrum ranging from raw data to processed and analysed data including analysis scripts and code. Physical data are all materials that need proper management because they are valuable, difficult to replace and/or ethical issues are associated. Materials that are not considered data in an RDM context include your own manuscripts, theses and presentations; documentation is an integral part of your datasets and should described under documentation/metadata.
RDM Guidance on data

If you reuse existing data, please specify the source, preferably by using a persistent identifier (e.g. DOI, Handle, URL etc.) per dataset or data type.	The baseline processed imaging, demographical and neuropsychological data of 26 healthy controls and 30 MCI patients, as well as the 2-year follow-up data of 21 MCI patients will be reused to investigate the longitudinal changes in the new 5-year follow-up segment. Baseline and 2-year pilot data were published as Vanderlinden et al. (2022) Molecular Psychiatry PMID 35794185 and full study data are accepted for publication in Alzheimer's & Dementia. Data from these baseline and 2-year follow-up segments are stored on a hard drive and on the UZ Leuven servers which are password-protected and access is restricted.

Are there any ethical issues concerning the creation and/or use of the data (e.g. experiments on humans or animals, dual use)? If so, refer to specific datasets or data types when appropriate and provide the relevant ethical approval number.	☒ Yes, human subject data; provide SMEC or EC approval number: ☐ Yes, animal data; provide ECD reference number: ☐ Yes, dual use; provide approval number: ☐ No Additional information: Baseline and 2-year follow-up data were acquired under ethical approval (S60721). The request for an amendment to perform a 5-year follow-up segment will be submitted in the next month.
Will you process personal data[4]? If so, please refer to specific datasets or data types when appropriate and provide the KU Leuven or UZ Leuven privacy register number (G or S number).	☒ Yes (provide PRET G-number or EC S-number below) ☐ No Additional information: Several types of personal data will be processed: - Personal data for the organization of the study visits: phone number, e-mail adres, home address, bank account number. These data will not be included in analyses. - Personal data for research purpose: ICF, demographical data (age, sex, educational level), medical history, dominant hand, medication intake, neuropsychological test data, image data. These personal data will be pseudonymized. The file where the pseudonyms are linked to the personal data and identifiers will be stored separately and secured, with access only for study staff.
Does your work have potential for commercial valorization (e.g. tech transfer, for example spin-offs, commercial exploitation, …)? If so, please comment per dataset or data type where appropriate.	☐ Yes ☒ No If yes, please comment:
Do existing 3rd party agreements restrict exploitation or dissemination of the data you (re)use (e.g. Material/Data transfer agreements, research collaboration agreements)? If so, please explain to what data they relate and what restrictions are in place.	☐ Yes ☒ No If yes, please explain:

[4] See Glossary Flemish Standard Data Management Plan

Are there any other legal issues, such as intellectual property rights and ownership, to be managed related to the data you (re)use? If so, please explain to what data they relate and which restrictions will be asserted.	☐ Yes ☒ No If yes, please explain:

3. Documentation and Metadata	
Clearly describe what approach will be followed to capture the accompanying information necessary to keep **data understandable and usable**, for yourself and others, now and in the future (e.g. in terms of documentation levels and types required, procedures used, Electronic Lab Notebooks, README.txt files, Codebook.tsv etc. where this information is recorded). *RDM guidance on documentation and metadata*.	All physical files will be stored per subject in a standardized case report form (CRF). Neuropsychological and demographical data will also be stored in RedCap (eCRF). Raw images will be saved in the international BIDS (brain imaging data structure) format. A user guide on the image processing pipeline that was used, will be saved as a READme.txt file according to KU Leuven's template.
Will a metadata standard be used to make it easier to **find and reuse the data**? If so, please specify which metadata standard will be used. If not, please specify which metadata will be created to make the data easier to find and reuse. *REPOSITORIES COULD ASK TO DELIVER METADATA IN A CERTAIN FORMAT, WITH SPECIFIED ONTOLOGIES AND VOCABULARIES, I.E. STANDARD LISTS WITH UNIQUE IDENTIFIERS.*	☒ Yes ☐ No If yes, please specify (where appropriate per dataset or data type) which metadata standard will be used: - Imaging data will be saved in the BIDS standard. - Pseudonymized information on demographics and neuropsychological test data will be stored in RedCap. RedCap offers the possibility to download a .xml file of the metadata. If no, please specify (where appropriate per dataset or data type) which metadata will be created:

4. Data Storage & Back-up during the Research Project	
Where will the data be stored? *Consult the interactive KU Leuven storage guide to find the most suitable storage solution for your data.*	☐ Shared network drive (J-drive) ☐ Personal network drive (I-drive) ☐ OneDrive (KU Leuven) ☐ Sharepoint online ☐ Sharepoint on-premis ☐ Large Volume Storage ☐ Digital Vault ☒ Other: Image data will be stored on external hard drives as well as on UZ Leuven MIM and PACS servers, neuropsychological and demographical data will be stored in RedCap, paper ICFs and paper neuropsychological tests will be stored in a CRF folder in a locked cabinet in an environment with restricted access.
How will the data be backed up? *WHAT STORAGE AND BACKUP PROCEDURES WILL BE IN PLACE TO PREVENT DATA LOSS?*	☐ Standard back-up provided by KU Leuven ICTS for my storage solution ☒ Personal back-ups I make (specify): I will make a back-up of the data on an external hard drive which will be stored in a locked cabinet in an access-controlled environment. Physical data files will be scanned and stored on the UZ Leuven server which is backed-up daily. ☐ Other (specify)
Is there currently sufficient storage & backup capacity during the project? If yes, specify concisely. If no or insufficient storage or backup capacities are available, then explain how this will be taken care of.	☒ Yes: RedCap provides unlimited capacity. Volume storage on the UZ Leuven server is also sufficient. ☐ No If no, please specify:

How will you ensure that the data are securely stored and not accessed or modified by unauthorized persons? *CLEARLY DESCRIBE THE MEASURES (IN TERMS OF PHYSICAL SECURITY, NETWORK SECURITY, AND SECURITY OF COMPUTER SYSTEMS AND FILES) THAT WILL BE TAKEN TO ENSURE THAT STORED AND TRANSFERRED DATA ARE SAFE.* *Guidance on security for research data*	The CRF (including ICF and all information collected on paper such as neuropsychological and demographical data) will be stored in a locked cabinet in an access-controlled environment. Digital data will be stored in RedCap. This platform has the possibility of allowing detailed access control on file and folder level, in this way we can prevent access to data and modification of data by unauthorized persons. The digital data will additionally be stored on the UZ Leuven server. Data back-up hard drives will be stored in a locked cabinet in an access-controlled environment. Image data will be stored on the MIM server.
What are the expected costs for data storage and backup during the research project? How will these costs be covered?	The price to set-up a RedCap projects is € 80 per year. There is no charge for paper storage. Costs for MIM server storage and UZ shared data drive IT storage (currently about 7500 Euro/year for the whole division) will be covered by the general research budget of the division of nuclear medicine, as has been done for the past 10 years for all KU Leuven researchers of Nuclear Medicine and Molecular Imaging.

5. Data Preservation after the end of the Research Project	
Which data will be retained for at least five years (or longer, in agreement with other retention policies that are applicable) after the end of the project? In case some data cannot be preserved, clearly state the reasons for this (e.g. legal or contractual restrictions, storage/budget issues, institutional policies...). *Guidance on data preservation*	☐ All data will be preserved for 10 years according to KU Leuven RDM policy ☒ All data will be preserved for 25 years according to CTC recommendations for clinical trials with medicinal products for human use and for clinical experiments on humans ☐ Certain data cannot be kept for 10 years (explain)

Where will these data be archived (stored and curated for the long-term)? *Dedicated data repositories are often the best place to preserve your data. Data not suitable for preservation in a repository can be stored using a KU Leuven storage solution, consult the interactive KU Leuven storage guide.*	☐ KU Leuven RDR ☐ Large Volume Storage (longterm for large volumes) ☐ Shared network drive (J-drive) ☒ Other (specifiy): The CRF (including ICF and all information collected on paper such as neuropsychological and demographical data) will be stored in a locked cabinet in an access-controlled environment. The digital data will be stored on the UZ Leuven server. Data back-up hard drives will be stored in the PI's office. Image data will be stored on the MIM server.
What are the expected costs for data preservation during the expected retention period? How will these costs be covered?	There is no charge for paper archiving. Costs for MIM server storage and UZ shared data drive IT storage (currently about 7500 Euro/year for the whole division) will be covered by the general research budget of the division of nuclear medicine, as has been done for the past 10 years for all KU Leuven researchers of Nuclear Medicine and Molecular Imaging.

6. Data Sharing and Reuse	
Will the data (or part of the data) be made available for reuse after/during the project? Please explain per dataset or data type which data will be made available. *NOTE THAT 'AVAILABLE' DOES NOT NECESSARILY MEAN THAT THE DATA SET BECOMES OPENLY AVAILABLE, CONDITIONS FOR ACCESS AND USE MAY APPLY. AVAILABILITY IN THIS QUESTION THUS ENTAILS BOTH OPEN & RESTRICTED ACCESS. FOR MORE INFORMATION: HTTPS://WIKI.SURFNET.NL/DISPLAY/STANDARDS/INFO-EU-REPO/#INFOEUREPO-ACCESSRIGHTS*	☐ Yes, as open data ☐ Yes, as embargoed data (temporary restriction) ☒ Yes, as restricted data (upon approval, or institutional access only) ☐ No (closed access) ☐ Other, please specify: Data will only be shared if the research is approved by the ethical committee.
If access is restricted, please specify who will be able to access the data and under what conditions.	Members of our own research group and everyone who is trained for the study will have access to the data. Data will only be shared if the research is approved by the ethical committee.

Are there any factors that restrict or prevent the sharing of (some of) the data (e.g. as defined in an agreement with a 3rd party, legal restrictions)? Please explain per dataset or data type where appropriate.	☒ Yes, privacy aspects ☐ Yes, intellectual property rights ☒ Yes, ethical aspects ☐ Yes, aspects of dual use ☐ Yes, other ☐ No If yes, please specify: All data originate from patients and healthy controls, they are personal data. Privacy regulations and ethical aspects restrict the sharing of these sensitive data, therefore pseudonymization of the full data set will be provided. Furthermore, the consent form specifies that data will only be shared for research that is approved by the ethical committee.
Where will the data be made available? If already known, please provide a repository per dataset or data type.	☒ KU Leuven RDR ☐ Other data repository (specify) ☐ Other (specify)
When will the data be made available?	☐ Upon publication of research results ☐ Specific date (specify) ☒ Other (specify): only upon reasonable request and after approval of the ethics committee.
Which data usage licenses are you going to provide? If none, please explain why. *A DATA USAGE LICENSE INDICATES WHETHER THE DATA CAN BE REUSED OR NOT AND UNDER WHAT CONDITIONS. IF NO LICENCE IS GRANTED, THE DATA ARE IN A GREY ZONE AND CANNOT BE LEGALLY REUSED. DO NOTE THAT YOU MAY ONLY RELEASE DATA UNDER A LICENCE CHOSEN BY YOURSELF IF IT DOES NOT ALREADY FALL UNDER ANOTHER LICENCE THAT MIGHT PROHIBIT THAT.* *Check the RDR guidance on licences for data and software sources code or consult the License selector tool to help you choose.*	☐ CC-BY 4.0 (data) ☒ Data Transfer Agreement (restricted data) ☐ MIT licence (code) ☐ GNU GPL-3.0 (code) ☐ Other (specify)

Do you intend to add a PID/DOI/accession number to your dataset(s)? If already available, please provide it here. *INDICATE WHETHER YOU INTEND TO ADD A PERSISTENT AND UNIQUE IDENTIFIER IN ORDER TO IDENTIFY AND RETRIEVE THE DATA.*	☐ Yes, a PID will be added upon deposit in a data repository ☐ My dataset already has a PID ☒ No
What are the expected costs for data sharing? How will these costs be covered?	There are no expected costs.

7. Responsibilities	
Who will manage data documentation and metadata during the research project?	**The grant holder, Greet Vanderlinden, in collaboration with the research group's data manager, Marie Cohilis, and the PI of the study, prof. Koen Van Laere.**
Who will manage data storage and backup during the research project?	**The grant holder, Greet Vanderlinden, in collaboration with the research group's data manager, Marie Cohilis, and the PI of the study, prof. Koen Van Laere.**
Who will manage data preservation and sharing?	**The grant holder, Greet Vanderlinden, in collaboration with the research group's data manager, Marie Cohilis, and the PI of the study, prof. Koen Van Laere.**
Who will update and implement this DMP?	**The grant holder, Greet Vanderlinden, in collaboration with the research group's data manager, Marie Cohilis, and the PI of the study, prof. Koen Van Laere.**