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	**Yasmin Dahdouh-Guebas; https://orcid.org/0000-0002-6578-4462**
Contributor name(s) (+ ORCID) & roles	**Supervisor:** **Lies De Groef; https://orcid.org/0000-0002-3329-3474** **Co-supervisors:** **Ingeborg Stalmans; https://orcid.org/0000-0001-7507-4512** **Joost Schymkowitz; https://orcid.org/0000-0003-2020-0168** **Xavier Hadoux; https://orcid.org/0000-0002-4524-3706**
Project number [1] & title	**1SHC824N** **Hyperspectral retinal imaging as a novel biomarker for Alzheimer's disease: towards a better understanding of the underlying pathological changes**
Funder(s) GrantID [2]	/
Affiliation(s)	☒ KU Leuven ☐ Universiteit Antwerpen ☐ Universiteit Gent ☐ Universiteit Hasselt ☐ Vrije Universiteit Brussel ☐ Other: ROR identifier KU Leuven: 05f950310

[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.

Please provide a short project description	Alzheimer's disease (AD) is a neurodegenerative disorder and leading cause of dementia, affecting more than 50 million people worldwide. In spite of decades of AD research, no effective treatment strategy has been found and adequate screening and diagnostic tools are lacking due to high-cost, invasiveness, and non-scalability. The retina is the only part of the central nervous system that is easily accessible for high-resolution, non-invasive imaging and offers unique opportunities for diagnosis and screening of neuropathies. Indeed, one of the AD hallmarks, amyloid beta (Aβ) deposition, can be detected using a novel imaging technique, namely hyperspectral retinal imaging (HSRI). Thus, in this research project, we aim to elucidate the molecular basis and pathological correlates of HSRI signals in AD, in order to use this approach as an AD biomarker. This will be achieved by studying HSRI signals *in vitro* and *in vivo*, starting with defining the signature of Aβ conformational forms; followed by characterizing the HSRI signature in different cell and mouse models of neurodegenerative proteinopathies; and correlating these data with biomolecular assays. The information on the specificity, molecular basis and pathological correlates of HSRI in AD gained from this project, will be critical for its rational implementation in the clinic.

2. Research Data Summary

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
Hyperspectral data-set work package 1	Hyperspectral images of different amyloid conformations	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☒ Images ☐ Sound ☐ Numerical ☐ Textual ☐ Model ☐ Software ☐ Other:	.hdr and .raw	☐ < 1 GB ☐ < 100 GB ☐ < 1 TB ☒ < 5 TB ☐ > 5 TB ☐ NA	/
Atomic force microscopy-based infrared spectroscopy data work package 1	Collection of AFM-IR, confocal, transmission electron and fluorescence microscopy images	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☒ Images ☐ Sound ☐ Numerical ☐ Textual ☐ Model ☐ Software ☐ Other:	.tiff, .czi, .oib, .jpg, .png	☐ < 1 GB ☐ < 100 GB ☒ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	/
Hyperspectral data-set work package 2	Hyperspectral images of biosensor cell lines	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☒ Images ☐ Sound ☐ Numerical ☐ Textual ☐ Model	.hdr and .raw	☐ < 1 GB ☐ < 100 GB ☐ < 1 TB ☒ < 5 TB ☐ > 5 TB ☐ NA	/

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

				☐ Software ☐ Other:			
Automated high-content images work package 2	Automated high content images of biosensor cell line amyloid, tau, a-syn aggregation	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☒ Images ☐ Sound ☐ Numerical ☐ Textual ☐ Model ☐ Software ☐ Other:	.r3d, .cif, .oir, .tif, .jpg, .png	☐ < 1 GB ☐ < 100 GB ☐ < 1 TB ☒ < 5 TB ☐ > 5 TB ☐ NA	/
Hyperspectral data-set work package 3	*In vivo/ ex vivo* hyperspectral images of mouse models	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☒ Images ☐ Sound ☐ Numerical ☐ Textual ☐ Model ☐ Software ☐ Other:	.hdr and .raw	☐ < 1 GB ☐ < 100 GB ☐ < 1 TB ☒ < 5 TB ☐ > 5 TB ☐ NA	/
Hyperspectral image analysis pipline	Python script to analyse hyperspectral images	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images ☐ Sound ☐ Numerical ☐ Textual ☐ Model ☐ Software ☒ Other: script	.py	☐ < 1 GB ☒ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	/
Biomolecular /histological data work package 3	Data from ELISA/MSD, Western blots and histological	☒ Generate new data	☒ Digital ☒ Physical	☐ Audiovisual ☒ Images ☐ Sound	.scn (WB images), .tiff, .oib, .txt, .csv	☐ < 1 GB ☐ < 100 GB ☒ < 1 TB	1x -80 freezer shelf 1x 4-8°C fridge shelf

	stainings of mouse brain/retina samples	☐ Reuse existing data		☒ Numerical ☒ Textual ☐ Model ☐ Software ☐ Other: script		☐ < 5 TB ☐ > 5 TB ☐ NA	(depending on the type of material)
Manuscripts/ reports	Manuscripts and reports concerning the project	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images ☐ Sound ☐ Numerical ☒ Textual ☐ Model ☐ Software ☐ Other:	.docx	☒ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	/
Brain and retina samples work package 3	Collected brain/retina samples from mice either for immunohistochemistry or molecular assays	☒ Generate new data ☐ Reuse existing data	☐ Digital ☒ Physical	☐ Audiovisual ☐ Images ☐ Sound ☐ Numerical ☐ Textual ☐ Model ☐ Software ☒ Other: Tissue	/	☐ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☒ NA	1x -80 freezer shelf 1x -20 freezer shelf 1x 4-8°C fridge shelf (depending on the type of material)

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.	/
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: We will conduct animal experiments, more specifically on mice, in accordance with the standard laboratory safety rules. All necessary safety measures for both laboratory and animal experiments will be strictly observed. Our methods follow the guidelines and rules set by the HSE Department (Health, Safety and Environment) and the Animal Ethics Committee at KU Leuven. Ethical permission for animal work will be obtained during year 2 of the FWO project.
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: .

[4] See Glossary Flemish Standard Data Management Plan

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: The project largely contains strategic research that will generate insights that may, in the long term, be further exploited for valorisation. The project consists of two valorization tracks. On the one hand, we aim to gather a solid base of evidence to support further clinical development of HSI as a diagnostic tool to assess retinal Aβ burden. On the other hand, we aim to develop tools, together with AD mouse models, that may serve as a preclinical research platform to assess the efficacy of a therapeutic intervention. While there is full FTO, there is no opportunity to generate IP. However, valorisation opportunities will be considered, in consultation with our IP officer Julien Compagnon at Leuven Research and Development.
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: Material transfer agreements have been signed to obtain the transgenic mouse lines used in this project. Valorisation opportunities, in agreement with these MTAs, will be evaluated case by case in consultation with our IP officer Julien Compagnon at Leuven Research and Development.
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.*	We will keep records for each work package for 10 years (where applicable): **Digital data:** - Experimental design and protocol (.docx file) - List of abbreviations used (.docx file) - Data structure documentation (.docx file) - Data analysis steps and relevant scripts (MATLAB, Python, ImageJ and Imaris scripts) - Raw data (specific file format according to data type) - Analyzed data (.xlsx and .prism) - Index file/readme file (.txt file) for each work package, detailing the names, locations (folder and subfolder structure), and descriptions of the aforementioned files. **Physical data:** Samples collected during experiments will be documented and preserved for the duration of the project. Depending on the sample type, they will be stored in fixatives or freezers. Immunohistologically stained slides will be stored in appropriate containers in a dry place or freezer. Due to the nature of these samples, they cannot be kept for 10 years and they will be discarded when analyses have been concluded.
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: The experimental set-up in this project is unique, and does not have a readily available metadata standard. The data will be standardized according to data-type across experiments, in which metadata will be created starting from the Dublin core (http://www.dcc.ac.uk/resources/metadata-standards/dublin-core). 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:
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) ☐ 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 ☐ No If no, please specify: The research group currently has 15 TB of storage on KU Leuven servers, and this can be expanded at hoc.

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 network storage is located in the KU Leuven ICTS data center, with a duplicate in the second ICTS center. This setup ensures disaster recovery and additional backup capacity, ensuring data availability in the long term. Access to the data is restricted by KU Leuven security groups, and all data will be password protected.
What are the expected costs for data storage and backup during the research project? How will these costs be covered?	Back-up cost per Tb (KU Leuven ICTS): 295€/year Large Volume Storage: 95,14€/Tb/year Expected amount of data (5 Tb). -The costs will be covered by complementary funding (obtained SAO project and lab resources).

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 (specify): Notebooks will be kept in the lab for at least 5 years, conform the KU Leuven RDM policy.
What are the expected costs for data preservation during the expected retention period? How will these costs be covered?	Back-up cost per Tb (KU Leuven ICTS): 295€/year Large Volume Storage: 95,14€/Tb/year Expected amount of data (5 Tb). -The costs will be covered by complementary funding (obtained SAO project and lab resources).

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: Written progress reports will be stored internally. Relevant findings will be disseminated through publication in high profile, peer-reviewed international journals. In addition, data will be presented at (inter)national scientific meetings specific to the field, such as ARVO, AAIC, AD-PD meetings, etc. We intend to make data openly available via data repositories. Requests for non-deposited data will be evaluated on a case-by-case basis and may be provided upon request.
If access is restricted, please specify who will be able to access the data and under what conditions.	/
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:
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): Open Access repository ☒ Other (specify): request by mail

When will the data be made available?	☒ Upon publication of research results ☐ Specific date (specify) ☐ Other (specify)
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/DOI/accession number 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?	We will opt for free repositories. The expected cost for ad hoc data sharing will be low, since the use of OneDrive is free for KU Leuven members up to 1TB. In addition, Belnet will be used to share data up to 6TB. We do not expect to exceed this.

7. Responsibilities	
Who will manage data documentation and metadata during the research project?	Responsibility for ensuring data preservation and sharing, as well as the end responsibility for updating and implementing the DMP is with the supervisors (Lies De Groef, Ingeborg Stalmans, Joost Schymkowitz).

Who will manage data storage and backup during the research project?	Data documentation, data storage & back up during the project is the responsibility of all researchers working on this project, including Yasmin Dahdouh-Guebas.
Who will manage data preservation and sharing?	Lies De Groef, Ingeborg Stalmans, Joost Schymkowitz
Who will update and implement this DMP?	Yasmin Dahdouh-Guebas, Lies De Groef, Ingeborg Stalmans, Joost Schymkowitz