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	Anne Wiedmer; 0009-0001-4297-962X
Contributor name(s) (+ ORCID) & roles	Dietmar Thal; 0000-0002-1036-1075 Patrik Verstreken; 0000-0002-5073-5393 Jolien Schaeverbeke; 0000-0003-2257-0568
Project number [1] & title	11E1225N - The role of co-morbid TDP-43 pathology in Alzheimer's disease and its impact on synaptic integrity
Funder(s) GrantID [2]	
Affiliation(s)	X KU Leuven ☐ Universiteit Antwerpen ☐ Universiteit Gent ☐ Universiteit Hasselt ☐ Vrije Universiteit Brussel X Other: CBD-VIB 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) constitutes a devastating form of neurodegeneration and is the leading cause of dementia worldwide. Primary neuropathological characteristics of AD include the formation of amyloid β senile plaques and the accumulation of phosphorylated tau into neurofibrillary tangles. The development of these pathologies occurs with progressive loss of synaptic function, which is one of the strongest neuropathological correlates of cognitive decline. Human post-mortem studies of AD brains revealed that co-morbid TDP-43 pathology, also known as limbic-predominant age-related TDP-43 encephalopathy neuropathologic changes (LATE-NC), enhances synaptic loss in AD. However, it is still unknown which molecular mechanisms are involved in this process. To address this, I will use mass spectrometry-based approaches to assess the impact of TDP-43 pathology on synaptic protein expression and its relation to other proteome changes. I will further determine whether specific post-translational modifications of TDP-43 are required to induce synaptic loss. For this, I will use a cohort of human postmortem AD cases covering cases with Braak stages ranging from 0-VI, with and without TDP-43 co-morbidity. The functional impact of differentially abundant proteins observed in AD cases with and without TDP-43 pathology will be determined in Drosophila after up- or downregulating of the respective genes.

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
Mass spectrometry files	Spectronaut	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images ☐ Sound ☐ Numerical ☐ Textual ☐ Model ☒ Software ☐ Other:	.sne .tsv	☐ < 1 GB ☐ < 100 GB ☐ < 1 TB ☒ < 5 TB ☐ > 5 TB ☐ NA	NA
Histological slides	5ug-thick formalin-fixed, paraffin-embedded tissue slides of post-mortem human brain and spinal cord tissue. These slides will be used for immunohistoch	☒ Generate new data ☐ Reuse existing data	☐ Digital ☒ Physical	☐ Audiovisual ☐ Images ☐ Sound ☐ Numerical ☐ Textual ☐ Model ☐ Software ☐ Other:	NA	☐ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☒ NA	Estimation: 1kg (tissue slide boxes)

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

	emistry and atomic-force microscopy.						
Database: human autopsy cases	Excel file containing demograpgics and neuropathological data regarding human autopsy cases previously collected.	☐ Generate new data ☒ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images ☐ Sound ☒ Numerical ☒ Textual ☐ Model ☐ Software ☐ Other:	.xlsx	☒ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	NA
Microscopic images	Brightfield, fluorescence and atomic-force microscopy images	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☒ Images ☐ Sound ☐ Numerical ☐ Textual ☐ Model ☐ Software ☐ Other:	.Tif; .jpg; .czi	☐ < 1 GB ☒ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	NA
Quantification files	Excel files containing neuropathological quantifications	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images ☐ Sound ☒ Numerical ☒ Textual ☐ Model ☐ Software ☐ Other:	.xlsx	☐ < 1 GB ☒ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	NA

Processed scientific data	Files resulting from statistical analysis and processing softwares (GraphpadPrism, R, Python Files, FIJI, ZEN)	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images ☐ Sound ☒ Numerical ☒ Textual ☐ Model ☒ Software ☒ Other: Data files	.Rmd .R .prism .tif .czi .xlsx .py	☐ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☒ > 5 TB ☐ NA	NA
Standard operating procedures (SOPs)	Protocols for experiments conducted for project	☒ Generate new data ☒ Reuse existing data		☐ Audiovisual ☐ Images ☐ Sound ☒ Numerical ☒ Textual ☐ Model ☐ Software ☐ Other:	.xlsx .docx	☒ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	NA

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 reused database is a common database from our research group. The database is not open access, but is found in our shared J drive (J:\GBW-0352_Neuropathology\NeuroPatho\BIOBANK\Database) Protocols were also developed in our research group, and are not open access but are located on out shared server found here: J:\GBW-0352_Neuropathology\NeuroPatho\Protocols

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: S-59292, S-52791 ☐ Yes, animal data; provide ECD reference number: ☐ Yes, dual use; provide approval number: ☐ No Additional information:
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:
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 mass spectrometry files could yield potential disease related biomarkers to be used in the clinical setting.

[4] See Glossary Flemish Standard Data Management Plan

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: -One mass spectrometry dataset will be generated in collaboration with Prof. Judith Steen at Harvard Medical School, and both parties will share the data equally and all publications results from these data will be with co-authorship -Part of the project will be conducted with my promoter Prof. Patrik Verstreken, at CBD-VIB, and data generated will also be equally shared and any resulting publications using these data will have co-authorship between the contributing parties. -MTAs will be established for these collaborations to share results, materials, methodologies and co-authorship.
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*.	**Our approach complies with FAIR:** • **Findable: Data files will be labelled with the date and name of experiment which serve as keywords to find the proper files. Files related to each project work package/experiment, will be organized in folders. Additionally, README.txt files will be created for every experiment type, including explanation of how the data was generated and/or quantified.** • **Accessible: Published data will provide a database in the supplementary material. All abbreviations are defined. Mass spectrometry data are uploaded to ProteomeXchange.** • **Interoperable: Standard and vocabolatories are provided in the related publications. Database upload follows standardized roads.** • **Reusable: With the measures explained above other researcher will be able to reuse our data properly.**
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: Mass spectrometry raw data will be uploaded for public access at Proteome Xchange. If no, please specify (where appropriate per dataset or data type) which metadata will be created: Neuropathological datasets are excel sheets with clearly identified variables. To our knowledge there are no neuropathology metadata standards for neurodegenerative diseases.

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: Institutional Microsoft teams
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 – 4.28 PB available on Large Volume Storage drive ☐ 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*	Data will be exclusively processed using KU Leuven computers and will not be at any moment transferred to personal devices. All data is user-protected. For reused data, there is an agreement in which each user makes a copy of the original file and makes the changes on that new file, preserving all unmodified versions in a separate folder. For transferring data between KU Leuven users, Belnet or UZ liquid files will be used.

What are the expected costs for data storage and backup during the research project? How will these costs be covered?	Covered by the research budget of Prof. Dietmar Thal. The cost of the J drive is <100EUR per year, and the L(large volume storage) drive is 10Tb, and the cost is 950EUR per year.

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):
What are the expected costs for data preservation during the expected retention period? How will these costs be covered?	Covered by the research budget of Prof. Dietmar Thal, approx. 1000 EUR per year to be stored on the J and L drives provided by the lab.

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:
If access is restricted, please specify who will be able to access the data and under what conditions.	Part of the data will be published as open data access /data manuscripts. For non-open access articles, data will be restricted, and accessible upon reasonable request.
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: In case that personal patient information be requested, this would not be possible as these data are not accessible to me.

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) Proteome Xchange ☐ Other (specify)
When will the data be made available?	☒ Upon publication of research results ☐ Specific date (specify) ☒ Other (specify) Part of the data will only be made available upon reasonable request.
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?	**No costs are expected.**

7. Responsibilities	
Who will manage data documentation and metadata during the research project?	**The grant holder (Anne Wiedmer), promoter (Dietmar Thal), and lab manager of the group (Alicja Ronisz).**
Who will manage data storage and backup during the research project?	**The grant holder (Anne Wiedmer) and promoter (Dietmar Thal).**
Who will manage data preservation and sharing?	**The grant holder (Anne Wiedmer) and promoter (Dietmar Thal).**
Who will update and implement this DMP?	**The grant holder (Anne Wiedmer).**