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	**Sandra Marisa Oliveira Tomé (0000-0001-5253-0730)**
Contributor name(s) (+ ORCID) & roles	**Dietmar R. Thal; promoter (0000-0002-1036-1075)** **Rik Vandenberghe; co-promoter (0000-0001-6237-2502)** **Joost Schymkowitz; co-promoter (0000-0003-2020-0168)**
Project number [1] & title	ZKE5788 PDO/24; Deciphering TDP-43 molecular signatures in LATE: towards a multitarget diagnosis
Funder(s) GrantID [2]	1225725N
Affiliation(s)	X KU Leuven ☐ Universiteit Antwerpen ☐ Universiteit Gent ☐ Universiteit Hasselt ☐ Vrije Universiteit Brussel X Other: VIB-KU Leuven 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	Dementia in elderly individuals results from a group of devastating diseases that currently affects more than 55 million people worldwide, with Alzheimer's Disease (AD) being the most common cause of dementia. Recently, LATE was defined as a new disease entity caused by TDP-43 pathology. Importantly, TDP-43 accumulates in up to 75% of AD patients. AD patients with TDP-43 have smaller hippocampal volumes and worsened cognition, suggesting that TDP-43 plays an important role in AD. Additionally, TDP-43 is also a major player in frontotemporal lobar degeneration (FTLD-TDP). Thus, the question arises whether the TDP-43 proteinopathy differs among these diseases. We recently observed distinct patterns of TDP-43 species in AD versus FTLD-TDP cases, however this is not yet known in pure LATE. Therefore, in this project we aim to investigate the differences in TDP-43 aggregate composition, aggregation properties and binding partners across dementing diseases. To address this, I will use human brain samples and state of the art techniques such as spatial proteomics and atomic force microscopy. These findings will have an impact in the clinical stratification of demented patients.

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
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	
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	
Quantificatio	Excel files	☐ Generate new	☒ Digital	☐ Audiovisual	.xlsx	☐ < 1 GB	

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

n files	containing neuropathological quantifications	data ☐ Reuse existing data	☐ Physical	☐ Images ☐ Sound ☒ Numerical ☒ Textual ☐ Model ☐ Software ☐ Other:		☒ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	
Statistical files	Files resulting from various statistical and image processing softwares (Graphpad, R, Anasys, FIJI, ZEN, Python)	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images ☐ Sound ☒ Numerical ☒ Textual ☒ Model ☒ Software ☐ Other:	.Rmd .R .prism .tif .czi .axz .irb .py	☐ < 1 GB ☐ < 100 GB ☐ < 1 TB ☒ < 5 TB ☐ > 5 TB ☐ NA	
Mass spectrometry files	Spectronaut	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images ☐ Sound ☐ Numerical ☐ Textual ☐ Model ☒ Software ☐ Other:	.sne .tsv .csv	☐ < 1 GB ☐ < 100 GB ☒ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	
Histological slides	5ug-thick formalin-fixed, paraffin-embedded tissue slides of post-mortem	☒ Generate new data ☐ Reuse existing data	☐ Digital ☒ Physical	☐ Audiovisual ☐ Images ☐ Sound ☐ Numerical ☐ Textual ☐ Model		☐ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	Estimation: 5kg (tissue slide boxes)

	human brain and spinal cord tissue. These slides will be used for immunohistochemistry and atomic-force microscopy.			☐ Software ☐ Other:			
Brain extracts	Brain homogenates resulting from protein extraction protocols. These homogenates will be sued for: mass spectrometry, ThT assays, immunoprecipitation, western blotm ELISA assays.	☒ Generate new data ☐ Reuse existing data	☐ Digital ☒ Physical	☐ Audiovisual ☐ Images ☐ Sound ☐ Numerical ☐ Textual ☐ Model ☐ Software ☐ Other:		☐ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	Estimation: 1kg (cryoboxes with microtubes)
Standard Operating procedures (SOP)	Protocols used for experiments	☒ Generate new data ☒ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images ☐ Sound ☐ Numerical ☒ Textual ☐ Model ☐ Software	.docx	☒ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	

				☐ Other:			
Database and tissue-Netherlands Brain Bank	Data and tissue received from a collaboration with the Netherlands Brain Bank	☐ Generate new data ☒ Reuse existing data	☒ Digital ☒ Physical	☐ Audiovisual ☐ Images ☐ Sound ☐ Numerical ☒ Textual ☐ Model ☐ Software ☐ Other:	.docx .xlsx	☒ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	500g of tissue slides
Database and tissue-U. Kentucky (USA)	Data and tissue received from a collaboration with Prof. Peter Nelson (USA)	☒ Generate new data ☒ Reuse existing data	☒ Digital ☒ Physical	☐ Audiovisual ☐ Images ☐ Sound ☐ Numerical ☒ Textual ☐ Model ☐ Software ☐ Other:	.docx .xlsx	☒ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	5kg of tissue slides 500g of brain homogenates
Data files and antibodies-ADx neurosciences	Data received from a collaboration with ADx Neurosciences	☐ Generate new data ☒ Reuse existing data	☒ Digital ☒ Physical	☐ Audiovisual ☐ Images ☐ Sound ☒ Numerical ☒ Textual ☐ Model ☐ Software ☐ Other:	.docx .xlsx	☒ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	200g of antibodies provided

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 and SOPs are a common database generated by our research group. These are not open access, but are found in our shared J drive (J: \GBW-0352_Neuropathology\NeuroPatho\BIOBANK\Database; J: \GBW-0352_Neuropathology\NeuroPatho\Protocols). The reused database with neuropathological information is stored in a pseudonymized file for the UZ/KU Leuven biobank (doi:10.1007/s00401-024-02815-w; doi: 10.1093/braincomms/fcae442). For research purposes the database is shared with the institutional teams hub.
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 (approval for collecting tissue) ☐ 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:

[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: There is a very low probability for valorization ('Mass spectrometry files' or 'Data files-ADx neurosciences')
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: - MTA with ADx Neurosciences (Ghent, Belgium) We provided them with brain homogenates, and they provided us with both antibodies and an excel file containing analyzed data). As per the MTA, we own all the generated data, and the agreement is that we will include ADx personnel as co-authors if manuscripts arise from this collaboration. - MTA with Prof. Pete Nelson (U. Kentucky, USA). It was provided to us post-mortem human tissue in the form of snap-frozen tissue and formalin-fixed tissue slides. We own the material and data generated by this collaboration. Prof. Nelson will be listed as co-author in any manuscript resulting from this collaboration. - Agreement with VIB-KU Leuven (one co-promoter). Part of the project will be conducted in VIB-KU Leuven facilities. Data is shared. Co-authorships will be generated from this collaboration. An MTA will be set up. - MTA with Netherlands Brain Bank (this MTA is done between KU Leuven and the Netherlands Brain Bank). The tissue and data provided by the NBB is owned by the grant holder.
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 files 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: All the remaining datasets will be stored as clearly identified excel files.

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: 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 Petabyte storage) ☐ 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?	This is covered by Prof. Dietmar Thal's research budget (approximately 1100 euro/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?	This will be covered by Prof. Dietmar Thal's research budget (approximately 1100 euro/year).

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. Part of the data will be published as open data (manuscripts). ☐ Yes, as embargoed data (temporary restriction) ☒ Yes, as restricted data (upon approval, or institutional access only). Other parts of the data will be restricted data (such as non-open access manuscripts), accessible upon reasonable request. ☐ No (closed access) ☐ Other, please specify:
If access is restricted, please specify who will be able to access the data and under what conditions.	For non-open access articles and for distinct datasets, the data will be restricted and accessible only 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: For privacy reasons, if additional clinical/personal information is requested, this information will not be shared as the researcher (grand holder) has no access.

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, GAIN database (Alzheimer's Association) ☒ Other (specify) As supplementary data for published manuscripts
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?	Currently, no extra costs will be expected. Only publication costs for open access articles need to be paid (ca. 3000-4000 €/article). Costs will be covered by Prof. Thal's research budget.

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