Plan Overview

A Data Management Plan created using DMPonline.be

Title: UNRAVELING RESILIENCE MECHANISMS TO ADDRESS NEURONAL VULNERABILITY IN ALZHEIMER'S DISEASE

Creator: Patrick Vandormael

Principal Investigator: n.n.

Project Administrator: Patrick Vandormael

Affiliation: KU Leuven (KUL)

Template: KU Leuven BOF-IOF

Principal Investigator: n.n. n.n.

Project abstract:

Alzheimer's disease (AD) is a highly prevalent neurodegenerative disorder that, in absence of effective therapies, is expected to affect ~139 million people worldwide by 2050. While current data point to amyloid-β (Aβ) as the pathogenic trigger, it is still not known how Aβ drives pathogenic cascades that ultimately lead to neurodegeneration. We believe that studying amyloid-positive, cognitively unimpaired (AP-CU) cases is instrumental to decipher AD-linked pathogenic and resilience mechanisms. Here, we propose a novel hypothesis where Aβ is essential but disease only results from the complex interplay between Aβ, other protein stressors, lipid metabolism and inflammation. Our research aims at unraveling the bases of neuron vulnerability and resilience by applying an interdisciplinary approach and cutting-edge approaches. The gained insights may guide novel strategies to fight AD.

ID: 213706

Start date: 01-10-2024

End date: 30-09-2028

Last modified: 27-03-2025

UNRAVELING RESILIENCE MECHANISMS TO ADDRESS NEURONAL VULNERABILITY IN ALZHEIMER'S DISEASE

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.

1. Human Brain Tissue Samples
- Work Package: WP1.1, WP1.2, WP1.3
- Description: Post-mortem brain samples from Alzheimer's disease (AD) patients, including early-onset familial AD (FAD), late-onset sporadic AD (SAD), and amyloid-positive cognitively unimpaired (AP-CU) individuals.
- Data Type: Observational
- Newly Generated/Collected or Reused: Collected and reused
- Digital or Physical: Physical
- Content Type: Biological tissue

2. Human Peripheral Blood Mononuclear Cells (PBMCs)
- Work Package: WP3.1
- Description: Cryopreserved PBMCs from AP-CU and AD patients, including early-onset and late-onset groups.
- Data Type: Observational
- Newly Generated/Collected or Reused: Collected and reused
- Digital or Physical: Physical
- Content Type: Biological cells

3. iPSC-Derived Neurons
- Work Package: WP2.1, WP2.2, WP2.3
- Description: Induced pluripotent stem cell (iPSC)-derived neurons from familial AD (FAD) patients with specific PSEN1 mutations.
- Data Type: Experimental
- Newly Generated/Collected or Reused: reused
- Digital or Physical: Physical
- Content Type: Biological cells

4. Mass Spectrometry Imaging (MSI) Data
- Work Package: WP1.1, WP1.2
- Description: Spatially-resolved mass spectrometry imaging data of amyloid plaques in human brain samples.
- Data Type: Experimental
- Newly Generated/Collected or Reused: Generated
- Digital or Physical: Digital
- Content Type: Imaging data
- Technical Format: MSI file format
- Volume Estimate: Up to 1 TB

5. Immunohistochemistry (IHC) and Fluorescence Microscopy , and Super-Resolution Microscopy Data
- Work Package: WP1.2, WP2.1, WP2.2,
- Description: Data from immunohistochemistry and fluorescence microscopy analyses of brain tissue samples and cells cultured in vitro.
- Data Type: Experimental
- Newly Generated/Collected or Reused: Generated
- Digital or Physical: Digital
- Content Type: Imaging data
- Technical Format: TIFF, JPEG
- Volume Estimate: Up to 500 GB

6. Spatial Transcriptomics Data
- Work Package: WP1.2
- Description: Spatial transcriptomics data from brain tissue samples using Multiplexed Error Robust Fluorescence In Situ Hybridization (MERFISH).
- Data Type: Experimental
- Newly Generated/Collected or Reused: Generated
- Digital or Physical: Digital
- Content Type: Gene expression data
- Technical Format: FASTQ, CSV
- Volume Estimate: Up to 2 TB

7. Proteomics Data
- Work Package: WP1.2, WP2.1
- Description: Quantitative proteomics data from human neurons and brain tissue samples.
- Data Type: Experimental
- Newly Generated/Collected or Reused: Generated
- Digital or Physical: Digital
- Content Type: Protein expression data
- Technical Format: RAW, CSV
- Volume Estimate: Up to 1 TB

8. Lipidomics Data
- Work Package: WP1.1, WP1.2, WP2.3, WP3.2
- Description: Data on lipid composition in amyloid plaques and brain tissue samples.
- Data Type: Experimental
- Newly Generated/Collected or Reused: Generated
- Digital or Physical: Digital
- Content Type: Lipid profiles
- Technical Format: RAW, CSV
- Volume Estimate: Up to 500 GB

9. Cytokine and Chemokine Release Data
- Work Package: WP3.1
- Description: Data on cytokine and chemokine release from microglia and astrocytes treated with Aβ profiles and TDP-43.
- Data Type: Experimental
- Newly Generated/Collected or Reused: Generated
- Digital or Physical: Digital
- Content Type: Cytokine profiles
- Technical Format: CSV
- Volume Estimate: Up to 100 GB

10. Calcium Imaging Data
- Work Package: WP3.1
- Description: Data from calcium imaging studies on neuronal-astrocytic cultures.
- Data Type: Experimental
- Newly Generated/Collected or Reused: Generated
- Digital or Physical: Digital
- Content Type: Imaging data
- Technical Format: TIFF, JPEG
- Volume Estimate: Up to 200 GB

11. Coherent Raman Scattering (CARS) Microscopy Data
- Work Package: WP3.1, WP3.2
- Description: Data from CARS microscopy to investigate lipid changes in cells.
- Data Type: Experimental
- Newly Generated/Collected or Reused: Generated

- Digital or Physical: Digital
- Content Type: Imaging data
- Technical Format: TIFF, JPEG
- Volume Estimate: Up to 300 GB

12. Fluorescence Activated Cell Sorting (FACS- Data
- Work Package: WP3.1, WP3.2
- Description: Data from FACS to investigate cell markers and phagocytosis in cells.
- Data Type: Experimental
- Newly Generated/Collected or Reused: Generated
- Digital or Physical: Digital

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:

Own datasets/tools

WP/Task	dataset/tool	source
WP1.2	immunostaining with spatial transcriptomics, and immunostaining with spatial lipidomics	unpublished, generated in-house

external datasets/tools

WP/Task	dataset/tool	source
WP1	human Brain samples	- UCL-Queen Sqaire Brain Bank - UZ Leuven Brain Bank (Dietmar Thal)
WP1	PBMCs	- Flemish Prevent AD Cohort KU Leuven and Alzheimer Research Cohort KU Leuven (Rik Vandenberghe)
WP1	iPSC	- iNDI repository

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 below)

Our study involves the creation and re-use of data derived from human bodily materials (HBM). We are committed to ensuring all ethical considerations are addressed.
This study will utilise human postmortem brain samples, patient PBMCs, and patient induced pluripotent stem cells and differentiated cell types. Our labs have approval from the Ethics Committee Research UZ/KU Leuven for the use of the human lines (study S65730), as well as for the SAD and AP-CU postmortem tissues (S60470).
In case additional samples need to be added, we will seek approval via amendment before initiation of the described experiments.

Will you process personal data? 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)

personal data types:

- demographic data: such as age, gender
- medical data: such as diagnosis, age at death, co-morbidities
- genetic data: Genetic information obtained from tissue samples, including DNA sequences

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

We do not exclude that the proposed work could result in research data with potential for tech transfer and valorization. Ownership of the data generated belongs to KU Leuven and VIB in accordance with the framework agreement of both institutes. VIB has a policy to actively monitor research data for such potential. If there is substantial potential, the invention will be thoroughly assessed, and in a number of cases the invention will be IP protected (mostly patent protection or copyright protection). As such the IP protection does not withhold the research data from being made public. In the case a decision is taken to file a patent application it will be planned so that publications need not be delayed.

Do existing 3rd party agreements restrict exploitation or dissemination of the data you (re)use (e.g. Material or Data transfer agreements, Research collaboration agreements)? If so, please explain in the comment section to what data they relate and what restrictions are in place.

- No

No third-party agreement restricts dissemination or exploitation of the data or strains generated from this project. Existing agreements between VIB and KU Leuven do not restrict publication of data.

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 in the comment section to what data they relate and which restrictions will be asserted.

- No

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).

The organisation of research data follows a hierarchical **folder structure** based on model organism (eg. iPSCs, human biopsies, PBMCs). At the second level, these folders contain subfolders per experiment type (eg. western blotting, immunofluorescence, spatial transcriptomics,...). At the third level, these folders contain subfolders per individual experiment (with date indication, see below). Each of these experiment folders contain subfolders for raw data, processed data, and analysed data.

Digital files and folders of individual experiments will be **named** following a standard naming convention. All file names in a given dataset will follow the same format: they will start with the date, followed by the project acronym, a concise descriptive name, and an optional version number. Whenever possible, names will be kept under 32 characters, containing only letters, numbers, and underscores. Dots will only be used for version control indicators (minor revisions indicated by decimal numbers, and major revisions by whole numbers): YYYYMMDD_Project_Experiment_ version.format. Similarly, folder naming for individual experiments follows the following structure: YYYYMMDD(_Project)_Experiment

Data will be accompanied by **documentation** containing all contextual and descriptive features of the research data, which allow to understand and (re)use the data. Detailed procedures of each experiment are logged in the electronic laboratory notebook (**E-notebook)**, with reference of each experiment to the directory path of the folder containing raw and analysed data files. This also includes data collection methods, protocols, and code explanation. Documentation is stored at the study- and the data-level, providing data provenance from the original source data to specific datasets linked to publications. Data will be generated following standardized **protocols**. Clear and detailed descriptions of these protocols will be stored in our lab protocol database and E-notebook.

Metadata will be documented by the research and technical staff at the time of data collection and analysis, by taking careful notes in the E-notebook and/or in hard copy lab notebooks that refer to specific datasets.

To allow long term access and use of research data will be stored or converted to **open file formats** as much as possible.

• Containers: TAR, ZIP
• databases: XML(xlsx), CSV, JSON
• Statistics: DTA, POR, SAS, SAV
• Images: TIFF, JPEG 2000, PNG, GIF
• Tabular data: CSV, TXT
• Text: XML(docx), PDF/A, HTML, JSON, TXT, RTF
• Sequencing data: FASTA, FASTQ

We use **controlled vocabularies** or ontologies when applicable to provide unambiguous meaning, for example:
• Gene Onotology: molecular function, cellular component, and biological role of RNA seq
• ENSEMBL or NBCI identifiers: gene identity
• HUGO Gene Nomenclature Committee: names and symbol of human genes
• UniProt protein accessions: protein identity

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.

- Yes

Metadata will be documented by the research and technical staff at the time of data collection and analysis, by taking careful notes in the electronic laboratory notebook (E-notebook) and/or in hard copy lab notebooks that refer to specific datasets. All datasets will be accompanied by a README.txt file containing all the associated metadata, which will include the following elements:
• Title: free text
• Creator: Last name, first name, organization
• Date and time reference
• Subject: Choice of keywords and classifications
• Structure: internal structure of the dataset, or the meaning of abbreviations (not necessary when it is clear from the in-file documentation).
• Description: Text explaining the content of the data set and other contextual information needed for the correct interpretation of the data, the software(s) (including version number) used to produce and to read the data, the purpose of the experiment, etc.
• Format: Details of the file format,
• Resource Type: data set, image, audio, etc.
• Identifier: DOI (when applicable)
• Access rights: closed access, embargoed access, restricted access, open access.
Additionally, we will closely monitor MIBBI (Minimum Information for Biological and Biomedical Investigations) for metadata standards more specific to our data type. For specific datasets, additional metadata will be associated with the data file as appropriate.

Data Storage & Back-up during the Research Project

Where will the data be stored?

- ManGO
- Shared network drive (J-drive)
- Other (specify below)
- Large Volume Storage

Digital data

- Primary storage for **active digital files** will be on KU Leuven servers. KU Leuven offers fast ("J-drive) and slower ("L-drive") storage that allows reading/writing/modification of non-confidential, confidential, and strictly confidential data.
- KU Leuven further offers the ManGO platform for storage and management of **large volumes of active research data** . This platform allows secure storage, manual and automated metadata coupling, data workflows, and file sharing.

- Data that is no longer active, can be **archived** on the KU Leuven K-drive, which allows reading of non-confidential, confidential, and strictly confidential data.

Physical samples

- Tissue samples: Tissues will be stored locally in the laboratory at the appropriate temperature. All human tissue samples will be registered with a Belgian biobank, in compliance with the Belgian law on human body material (dd 19-12-2008).
- Cell lines: Human cell lines will be stored locally in the laboratory in liquid nitrogen storage and will be deposited in the UZ Leuven-KU Leuven Biobank. Animal cell lines will be stored in liquid nitrogen of the laboratory.
- Vectors: As a general rule at least two independently obtained clones will be preserved for each vector, both under the form of purified DNA (in -20°C freezer) and as a bacterial glycerol stock (-80°C). All published vectors and the associated sequences will be sent to the non-profit plasmid repository Addgene, which will take care of vector storage and shipping upon request.
- Bacterial strains will be stored in a -80°C freezer.

How will the data be backed up?

- Standard back-up provided by KU Leuven ICTS for my storage solution

KU Leuven drives are backed-up according to the following scheme:
- data stored on the "**L-drive**" is backed up daily using snapshot technology, where all incremental changes in respect of the previous version are kept online; the last 14 backups are kept.
- data stored on the "**J-drive**" is backed up hourly, daily (every day at midnight) and weekly (at midnight between Saturday and Sunday); in each case the last 6 backups are kept.
- data stored in **ManGO**: Snapshots are made at regular intervals (hourly, daily and monthly) in case data needs to be recovered. The data itself is synchronized on two separate hardware storage systems, each 6 PB large, located at Leuven and at Heverlee (ICTS). The data is protected against calamities at either site by synchronizing it in real-time at hardware level.

Is there currently sufficient storage & backup capacity during the project?

If no or insufficient storage or backup capacities are available, explain how this will be taken care of.

- Yes

KU Leuven servers offer sufficient storage for active data (J/L-drive, ManGO) and archived data (K-drive). Required data-storage volumes can be easily scaled up.

How will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?

The buildings on our campus are restricted by badge system so only employees are allowed in and visitors are allowed under supervision after registration.
Access to KU Leuven servers is possible only through using a KU Leuven user-id and password, and user rights only grant access to their own data, or data that was shared to them. Data in these drives are mirrored in the second ICTS datacenter for business continuity and disaster recovery so that a copy of the data can be recovered within an hour.

What are the expected costs for data storage and backup during the research project? How will these costs be covered?

-The costs of digital data storage are as follows:

- 519€/TB/Year for the "J-drive" (fast storage, active data),
- 569,2€/5TB/Year for the "L-drive" (medium speed storage, active data),
- 35€/TB/Year for the ManGO platform (large volume storage, active data),

Data storage and backup costs are included in general lab costs.

Data Preservation after the end of the Research Project

Which data will be retained for 10 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...).

- All data will be preserved for 10 years according to KU Leuven RDM policy

According to KU Leuven RDM policy, relevant research data will be preserved on the university's servers for a minimum of 10 years. Such data include data that are at the basis of a publication, that can only be generated or collected once, that are generated as a result of a substantial financial or personal effort, or are likely to be reused within the research unit or in wider contexts.

Where will these data be archived (stored and curated for the long-term)?

- Large Volume Storage (longterm for large volumes)
- Other (specify below)

As a general rule all research outputs (data, documentation, and metadata) related to publications will be made openly accessible, whenever possible via existing platforms that support FAIR data sharing (www.fairsharing.org). We aim at communicating our results in top journals that require full disclosure upon publication of all included data, either in the main text, in supplementary material or in a separate data repository.
Other research data will be archived on KU Leuven servers as described above.

What are the expected costs for data preservation during the expected retention period? How will these costs be covered?

-The costs of digital data storage are as follows: 569,2€/5TB/Year for the "K-drive" and the “L-drive”, 519€/TB/Year for the “J-drive”, and 35€/TB/Year for the ManGO platform. Data storage and backup costs are included in general lab costs.

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.

- Yes, as open data
- Yes, as restricted data (upon approval, or institutional access only)

- Upon publication, datasets and metadata generated from **human (omics) data that do not contain identifiable data** will be stored in general public repositories such as Zenodo or in domain-specific repositories such as the NCBI Gene Expression Omnibus, where they will receive a unique and persistent identifier.
- Datasets and metadata generated from **human omics that contain sensitive data** will be deposited under restricted access on the European Genome Phenome Archive (EGA), where they will be assigned a unique and persistent identifier.
- To ensure data findability, links and references these datasets, workflows and modes will be included in the data availability statements of the associated publication.

If access is restricted, please specify who will be able to access the data and under what conditions.

Access to restricted access dataset (such as human omics datasets) is governed by the Data Access Committees of KULeuven/UZ Leuven or VIB.

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

- Human omics data are considered sensitive personal data, and are are only made available on restricted access repositories such as the European Genome Phenome Archive (EGA). Access to these datasets is under control of a Data Access Committee.
- The researchers involved and the IP team of the VIB TechTransfer Office shall make the necessary arrangements in order to maintain an embargo on the public access of research data, at least until the essential steps in securing intellectual property (e.g. the filing of a patent application) have been taken. As such the IP protection does not withhold the research data from being made public. In the case a decision is taken to file a patent application it will be planned so that publications need not be delayed.

Where will the data be made available?

If already known, please provide a repository per dataset or data type.

- Other data repository (specify below)

- Upon publication, datasets and metadata generated from **human (omics) data that do not contain identifiable data** will be stored in general public repositories such as Zenodo or in domain-specific repositories such as the NCBI Gene Expression Omnibus, where they will receive a unique and persistent identifier.
- Datasets and metadata generated from **human omics that contain sensitive data** will be deposited under restricted access on the European Genome Phenome Archive (EGA), where they will be assigned a unique and persistent identifier.

When will the data be made available?

- Upon publication of research results

Which data usage licenses are you going to provide?

If none, please explain why.

- CC-BY 4.0 (data)
- Data Transfer Agreement (restricted data)

Do you intend to add a persistent identifier (PID) to your dataset(s), e.g. a DOI or accession number? If already available, please provide it here.

- Yes, a PID will be added upon deposit in a data repository

What are the expected costs for data sharing? How will these costs be covered?

It is the intention to minimize data management costs by implementing standard procedures e.g. for metadata collection and file storage and organization from the start of the project, and by using free-to-use data repositories and dissemination facilities whenever possible. Data management costs will be covered by the laboratory budget.

Responsibilities

Who will manage data documentation and metadata during the research project?

The researchers who generate the data are responsible for managing data, documentation, and metadata.

Who will manage data storage and backup during the research project?

The researchers who generate the data are responsible for storage and backup, with support from René Custers and Alexander Botzki for the electronic laboratory notebook (ELN) and from Raf De Coster for the KU Leuven drives.

Who will manage data preservation and sharing?

The PI is responsible for data preservation and sharing, with support from the research and technical staff involved in the project, from René Custers and Alexander Botzki for the electronic laboratory notebook (ELN), and from Raf De Coster for the KU Leuven drives.

Who will update and implement this DMP?

The PI is ultimately responsible for all data management during and after data collection, including implementing and updating the DMP, with support from Patrick Vandormael.