FWO DMP Template - Flemish Standard Data Management Plan

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	**Eline Bernaerts 0000-0002-7998-4553**
Contributor name(s) (+ ORCID) & roles	**Patrick Matthys 0000-0002-9685-6836 Supervisor** **Lien De Somer 0000-0002-8488-5090 Co-supervisor** **Jennifer Vandooren 0000-0002-7157-3370 Co-supervisor**
Project number[1] & title	Microglia-associated metalloproteinases in demyelinating disorders of the central nervous system
Funder(s) GrantID[2]	**11H9123N**
Affiliation(s)	☒KU Leuven ☐ Universiteit Antwerpen ☐ Universiteit Gent ☐ Universiteit Hasselt ☐ Vrije Universiteit Brussel ☐ Other: Provide ROR[3] identifier when possible:
Please provide a short project description	In multiple sclerosis (MS) an autoimmune reaction triggers neuroinflammation, resulting in the breakdown of the myelin sheet (demyelination) surrounding axons, thereby causing neurological deficits. Recently, it became evident that microglia are the principal effector cells in CNS pathologies. However, the exact contribution and role of their produced metalloproteinases (MPs) remains unclarified. We hypothesize that the cellular localization of MPs produced by microglia can shape central nervous system (CNS) pathology and contribute to demyelinating and remyelinating processes. This study will provide the first systematic analysis of MP localization and activity in disease-associated microglia *ex vivo* and *in vivo.* We will validate our obtained results in samples from patients with clinically isolated syndromes (the initial phase of MS). The proposed research will provide new insight in the contribution of microglial MPs in CNS pathology.

[1] "Project number" refers to the institutional project number. This question is optional since not every institution has an internal project number different from the GrantID. 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.

[3] Research Organization Registry Community. https://ror.org/

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[4].

				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
Patient data	- Data CIS, SMA and HC - Samples CIS, SMA and HC	☒ Generate new data ☐ Reuse existing data	☒ Digital ☒ Physical	☒ Observational ☒ Experimental ☐ Compiled/ aggregated data ☐ Simulation data ☐ Software ☐ Other ☐ NA	☐ .por ☐ .xml ☐ .tab ☐ .csv ☒ .pdf ☐ .txt ☐ .rtf ☐ .dwg ☐ .tab ☐ .gml ☒ other: .xlsx ☐ NA	☐ < 100 MB ☐ < 1 GB ☒ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	- Informed consent of the patients will be stored on paper (600 pages) - Samples of patients will be stored in the KU/UZ Leuven Biobank.
Instrument data files	-Flow cytometric analysis of leukocytes - Data of in vivo	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☒ Observational ☐ Experimental ☐ Compiled/ aggregated data ☐ Simulation	☐ .por ☐ .xml ☐ .tab ☐ .csv ☒ .pdf	☐ < 100 MB ☐ < 1 GB ☒ < 100 GB ☐ < 1 TB ☐ < 5 TB	

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

	experiments, including recorded disease parameters of mice - Data from ELISA - Data from proteolytic activity			data ☐ Software ☐ Other ☐ NA	☐ .txt ☐ .rtf ☐ .dwg ☐ .tab ☐ .gml ☒ other: .xlsx, .fcs, .wps., .pzfx ☐ NA	☐ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	
Single cell RNA sequencing	- Sequencing data files - R scripts - List of gene counts - Figures of output	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Observational ☒ Experimental ☐ Compiled/ aggregated data ☐ Simulation data ☐ Software ☐ Other ☐ NA	☐ .por ☐ .xml ☐ .tab ☐ .csv ☒ .pdf ☐ .txt ☐ .rtf ☐ .dwg ☐ .tab ☐ .gml ☒ other: .xlslx, .png, fastq., .bam, .fa, .mtx, .tsv, .R, .RData, .rds ☐ NA	☐ < 100 MB ☐ < 1 GB ☒ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	
Documentation	- Experimental protocols - Breeding programs	☒ Generate new data ☐ Reuse existing data	☒ Digital ☒ Physical	☒ Observational ☒ Experimental ☐ Compiled/ aggregated data	☐ .por ☐ .xml ☐ .tab ☐ .csv	☐ < 100 MB ☐ < 1 GB ☒ < 100 GB ☐ < 1 TB	- Experimental protocols will also be stored on paper (5000 pages)

	- Manuscripts			☐ Simulation data ☐ Software ☐ Other ☐ NA	☒ .pdf ☐ .txt ☐ .rtf ☐ .dwg ☐ .tab ☐ .gml ☒ other: .xlsx, .docx ☐ NA	☐ < 5 TB ☐ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	
Images and movies	- Images generated from Western Blot analysis - Images generated from whole slide imaging scanning - Images and movies generated from confocal)microscopy - Images generated from Incucyte analysis	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Observational ☒ Experimental ☐ Compiled/ aggregated data ☐ Simulation data ☐ Software ☐ Other ☐ NA	☐ .por ☐ .xml ☐ .tab ☐ .csv ☐ .pdf ☐ .txt ☐ .rtf ☐ .dwg ☐ .tab ☐ .gml ☐ other: .tif, .png, .mp4, .qptiff, .ims ☐ NA	☐ < 100 MB ☐ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☒ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	

Guidance: *Data can be digital or physical (for example biobank, biological samples, …). Data type: Data are often grouped by type (observational, experimental etc.), format and/or collection/generation method.* *Examples of data types: observational (e.g. survey results, sensor readings, sensory observations); experimental (e.g. microscopy, spectroscopy, chromatograms, gene sequences); compiled/aggregated data[5] (e.g. text & data mining, derived variables, 3D modelling); simulation data (e.g. climate models); software, etc.* *Examples of data formats: tabular data (.por,. spss, structured text or mark-up file XML, .tab, .csv), textual data (.rtf, .xml, .txt), geospatial data (.dwg,. GML, ..), image data, audio data, video data, documentation & computational script.* *digital data volume: Please estimate the upper limit of the volume of the data per dataset or data type.* *physical volume: Please estimate the physical volume of the research materials (for example the number of relevant biological samples that need to be stored and preserved during the project and/or after).*	
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.	NA
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, please describe these issues further and refer to specific datasets or data types when appropriate.	☐ Yes, human subject data ☐ Yes, animal data ☒ Yes, dual use ☐ No If yes, please describe: ***Dataset "Patient Data"*** The Ethics Committee Research UZ Leuven/KU Leuven approved the use of samples of patients and healthy controls for the purposes of this research project (S66508/2022 ECD). ***Dataset "Experimental data"*** Ethical approval of animal studies has also been granted by the Committee for Animal Experimentation at KU Leuven (project 046/2021: Regulation of extracellular proteolysis by microglia and brain-infiltrating macrophages).

[5] These data are generated by combining multiple existing datasets.

Will you process personal data[6]? If so, briefly describe the kind of personal data you will use. Please refer to specific datasets or data types when appropriate. If available, add the reference to your file in your host institution's privacy register.	☒ Yes ☐ No If yes: - Short description of the kind of personal data that will be used: ***Dataset "Patient Data"*** We will use personal data of patients and healthy controls included in the study (S66508). The collected patient data consist of general characteristics (i.e. age, sex) as well as disease-specific information (e.g. plasma levels of inflammatory markers, type of treatment, etc.) and the results of functional assays and immunophenotyping of leukocytes. The data of healthy controls consist of general characteristics and the results of functional assays and immunophenotyping of leukocytes. All patient data will be collected by Prof. Dr. Lien De Somer, Prof. Dr. Katrien Jansen and Prof. Dr. Liesbeth De Waele after gaining informed consent. When the researchers receive the patient data, this data is always pseudonymized to make patient name tracing impossible. - Privacy Registry Reference: ***G-2022-5212 PRET and* S66508/2022 CTC**
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:

[6] 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: An MTA agreement is in place for the use of HexB mice. This agreement states that that the RECIPIENT (Prof. Dr. Patrick Matthys) shall provide the PROVIDER (Prof. Dr. Macro Prinz) with an advance copy of any proposed publication or disclosure for its review at least thirty (30) days prior to the scheduled disclosure of the RESULTS. The PROVIDER may request that the RECIPIENT deletes any reference to the PROVIDER's confidential information.
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).	Documentation including experimental protocols, breeding schemes, and observational numeric data will be recorded in physical lab books and stored into Word or Excel files, which automatically imprint the metadata (user, date, time, equipment parameters) from these experiments. Data folders containing raw and processed data will be hierarchically organized and labelled based on the date of data generation, the number of the experiment and the source of the data. Imaging data will be created by default with metadata imprinted by the image acquisition software automatically. This includes information on user, data and time, duration of experiments, equipment parameters and imaging configurations. The metadata are saved and transferred with the original imaging file. The created data files will be organized in folders named by the data of the experiment (DDMMYYYY) followed by the research who performed it and the title of the experiment. Overall, all files will be stored in the KU Leven shared Storage space (J-Drive), with sharing possibilities via Box Sync and One Drive (managed by the KU Leuven IT department).

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: 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?	Electronical data will be stored in conformity with KU Leuven and FWO RDM policy. All datatypes, including protocols, raw data and analysed data, will be stored at a shared Rega Drive. We will keep a copy of all data on an external hard disc and/or computer, and on the KU Leuven One drive account, except for "the patient dataset". There is sufficient storage and backup available at the Rega Institute. In case additional storage is required, the KU Leuven data centre provides storage on two additional locations, in order to preserve data for a period of more than 20 years. Hard copy notebooks with raw data will be stored physically in our laboratory. Informed consents will be stored by the treating clinicians (UZ Leuven). Physical samples will be stored in the KU Leuven/UZ Leuven Biobank.

How will the data be backed up? *What storage and backup procedures will be in place to prevent data loss? Describe the locations, storage media and procedures that will be used for storing and backing up digital and non-digital data during research.*[7] *Refer to institution-specific policies regarding backup procedures when appropriate.*	We will use the central server storage of KU Leuven (Data centre ICTS Luna storage), which provides a daily automatic back up. Moreover, the data will be backed up on the Rega Institute Virtual Drives (Rega NAS (network adapted storage)) and on external hard-drives kept by the investigators.
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 yes, please specify concisely: Enough storage and back-up capacity is available at the systems of Rega Institute. All data will be stored on the J-drive, which has an unlimited maximum size (shared storage). Backup will be stored on the K-drive which has an unlimited maximum size (archive storage). 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.* [7]	Direct access to research data will be restricted to laboratory members, project members and collaborators. To protect our data, the shared Rega drive is secured with a login connected to your personal KU Leuven account and password.

[7] Source: Ghent University Generic DMP Evaluation Rubric: https://osf.io/2z5g3/

What are the expected costs for data storage and backup during the research project? How will these costs be covered?	Long-term data storage and costs will be managed by the principal investigator working in the project, Prof. Patrick Matthys and our IT-manager (Mr Dieter Devos). The cost for data storage is 519 euro/TB/year, thus the accumulated cost for 4 years is approximately 2000 euro. The costs will be covered by previous funding obtained by the host lab and by the bench fee offered by the FWO PhD fellowship.

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...).	"**Patient data**" will be preserved for 20 years as required by the informed consent form of the ethical committee of UZ Leuven. All other research data will be stored up to 5 years after the end of the project.
Where will these data be archived (stored and curated for the long-term)?	Data will be stored redundantly during and after the research in the KU Leuven data centers (ICTS Luna storage [J:// drive], [K:// drive], and Rega NAS [network adapted storage]).
What are the expected costs for data preservation during the expected retention period? How will these costs be covered?	The cost for data storage is 519 euro/TB/year. Long-term data storage and costs will be managed and evaluated by the principal investigator of this project, i.e. my promotor Prof. Patrick Matthys.

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, in an Open Access repository ☒ Yes, in a restricted access repository (after approval, institutional access only, …) ☐ No (closed access) ☐ Other, please specify: All data will be made available at the end of the project (after publishing the results).
If access is restricted, please specify who will be able to access the data and under what conditions.	Direct access to research data will be restricted to laboratory members, project members and collaborators. External members, who are not directly related to the project, will be given access after contact and evaluation by the principal investigator, Prof. Patrick Matthys.
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.	This is not known yet.

When will the data be made available? *This could be a specific date (dd/mm/yyyy) or an indication such as 'upon publication of research results'.*	Upon publication of the research results, data will become immediately available after publication.
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.* *Example Answer: E.g. "Data from the project that can be shared will be made available under a Creative Commons Attribution license (CC-BY 4.0), so that users have to give credit to the original data creators."* [8]	Data from the project that can be shared will be made available under a creative commons attribution license (cc-by 4.0), so that users have to give credit to the original data creators.
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 ☐ No If yes: Concerning the single cell RNA sequencing data set, we will add a PID to this dataset to identify and retrieve data. Since this dataset does not exist yet, there is no PID available.
What are the expected costs for data sharing? How will these costs be covered?	Local costs are minimal. Data transfer to external partners will be a at the partners cost.

[8] Source: Ghent University Generic DMP Evaluation Rubric: https://osf.io/2z5g3/

7. Responsibilities	
Who will manage data documentation and metadata during the research project?	The principal investigator (Prof. Dr. Patrick Matthys) and the researcher (Eline Bernaerts) bear the responsibility for data documentation.
Who will manage data storage and backup during the research project?	The principal investigator (Prof. Dr. Patrick Matthys) and the researcher (Eline Bernaerts) bear the responsibility for data storage and back up during the project.
Who will manage data preservation and sharing?	The principal investigator (Prof. Dr. Patrick Matthys) and the researcher (Eline Bernaerts) bear the responsibility for data preservation and sharing.
Who will update and implement this DMP?	The principal investigator (Prof. Dr. Patrick Matthys) bears the end responsibility of updating and implementing this DMP.