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	Hilde Brems, 0000-0002-0325-4060
Contributor name(s) (+ ORCID) & roles	Prof. Hilde Brems, Principle investigator, 0000-0002-0325-4060 Seppe Van der Auweraer, PhD student, 0000-0002-4034-038X Rien Blok, consortium member/collaborator Prof. Katharina Wimmer, consortium member/collaborator Magdalena Koczkowska, consortium member/collaborator Prof. Eric Pasmant, consortium member/collaborator, 0000-0002-1881-8762 Elisabeth Castellanos, consortium member/collaborator, 0000-0002-8133-5325
Project number [1] & title	3M230358, Development of new analytic tools and pathways to accelerate diagnosis and facilitate diagnostic monitoring of rare diseases
Funder(s) GrantID [2]	G0L1722N
Affiliation(s)	x 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	Neurofibromatosis type 1 (NF1), NF2 and non-NF2 schwannomatosis (SWN) belong to the NF disease spectrum. Legius syndrome (LGSS) and Constitutional Mismatch Repair Deficiency (CMMRD) are both differential diagnoses of NF1. These five syndromes are rare, genetically and clinically distinct diseases with variable course, however the majority present with severe complications and tumor development. The life expectancy is mainly improved by early diagnosis and personalized medical care in a specialized reference center. A large percentage of these patients currently remain without a genetic diagnosis, although their clinical features strongly suggest the presence of a still unknown genetic cause.
	The main objective of this project is to expand the availability and access of validated diagnostic and functional tools among the different European consortium members and other laboratories to maximize the genetic diagnostic yield for patients with NF-SWN and related disorders. Firstly, we will develop, validate and compare new DNA and RNA-based testing techniques to improve genetic diagnosis, including targeted RNA sequencing for structural and splice variant detection, and duplex sequencing for mosaic variant detection. Secondly, we want to test the pathogenic effect of variants of unknown significance on RNA and DNA level by using functional assays and iPSC models. Furthermore, We will develop a patient-derived hiPSC neuronal model to study the neurodevelopmental difficulties observed in NF1, and Legius syndrome. Thirdly, we will retest NF1-, LGSS-, NF2- and SWN-like patients with negative genetic diagnosis using newly developed RNA and DNA molecular techniques. We expect to share new sensitive and specific diagnostic assays and bioinformatics tools that can be globally used to characterize molecularly patients with NF-SWN-LGSS. A worldwide accurate molecular diagnosis for patients with NFSWN and related disorders will greatly impact the financial and clinical outcome as well as the therapeutic options, quality of life and societal burden.

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].

Dataset Name	Description	New or Reused	Digital or Physical	*ONLY FOR DIGITAL DATA* Digital Data Type	*ONLY FOR DIGITAL DATA* Digital Data Format	*ONLY FOR DIGITAL DATA* Digital Data Volume (MB, GB, TB)	*ONLY FOR PHYSICAL DATA* Physical Volume
Targeted DNA and RNA sequencing data	Raw and processed sequencing data from targeted DNA and RNA sequencing, used pathogenic variant detection in patient samples	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images ☐ Sound ☐ Numerical ☐ Textual ☐ Model ☐ Software ☒ Other: Sequencing data	Sequencing data (FASTQ,…)	☐ < 1 GB ☐ < 100 GB ☐ < 1 TB ☒ < 5 TB ☐ > 5 TB ☐ NA	
Microscope imaging	We will use confocal, fluorescence and brightfield microscope to assess morphology of hiPSCs and derived cell types, and for other cells (including HEK297T cells) used for functional analysis of pathogenic variants	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☒ Images ☐ Sound ☐ Numerical ☐ Textual ☐ Model ☐ Software ☐ Other:	Imaging data (TIF, czi,jpeg,…)	☐ < 1 GB ☐ < 100 GB ☒ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	

---

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

Digital images	Digital images and plots related to gel scans (e.g. western blot), graphs, illustrations, figures	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☒ Images ☐ Sound ☐ Numerical ☐ Textual ☐ Model ☐ Software ☐ Other:	TIFF, PN, JPEG, PDF,AI,…	☐ < 1 GB ☐ < 100 GB ☒ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	
Human iPSC cell lines	Patient-derived human iPSCs will be developed to study the neurodevelopmental difficulties observed in Legius syndrome	☒ Generate new data ☐ Reuse existing data	☐ Digital ☒ Physical				2 patient-derived hiPSCs lines and 2 control lines
DNA and RNA samples from patients	DNA and RNA samples from patients for genetic testing	☐ Generate new data ☒ Reuse existing data	☐ Digital ☒ Physical				60-100 DNA samples from patients (stored in UZ Leuven)
DNA from constructs	Constructs for functional analysis	☒ Generate new data	☒ Physical				60-100 constructs (stored in UZ Leuven)
Functional data	Quantitative data, and corresponding graphs and figures resulting from functional assays to evaluate variants of unknown significance	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images ☐ Sound ☒ Numerical ☐ Textual ☐ Model ☐ Software ☐ Other:	TXT, JPEG, PDF	☐ < 1 GB ☒ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	

Clinical data of patients	Clinical data of patients will be used for interpretation of pathogenic variants. These will be either from local clinical data at UZ Leuven, and data available in databases (Clinvar, LOVD,…)	☐ Generate new data ☒ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☒ Images ☐ Sound ☐ Numerical ☒ Textual ☐ Model ☐ Software ☐ Other:	TXT, PDF, webpages	☒ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ 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.	Genetic variant and anonymized clinical data: https://www.ncbi.nlm.nih.gov/clinvar/; https://www.lovd.nl/, https://www.hgmd.cf.ac.uk/ac/index.php & clinical data from UZ Leuven and collaborating European partners (pseudonymized data)
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: S68494 ☐ Yes, animal data; provide ECD reference number: ☐ Yes, dual use; provide approval number: ☐ No Additional information: This project is approved by ethical committee under nr. S68494.

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: Our research will use and process personal patient information, and personal genetic and genomic data, and human tissue samples (hiPSCs derived from patients). These data will be pseudonymized and handled according to the data handling guidelines stated in the protocol approved by the EC under S68494, and approved PRET-form: G-2024-7833.
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:
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: This project is part of a European consortium, where data dissemination is restricted to the agreements stated in the consortium agreement: "The Parties agree to comply with Article 89 GDPR for the Shared Data. If technically possible the parties will apply pseudonimization, however in any case both parties shall have in place appropriate technical and organisational measures to protect the personal data against accidental or unlawful destruction or accidental loss, alteration, unauthorised disclosure or access, and which provide a level of security appropriate to the risk represented by the processing and the nature of the data to be protected and work towards a technical systems that allows for pseudonimization of personal 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 to what data they relate and which restrictions will be asserted.	☐ Yes ☒ No If yes, please explain:

---

[4] See Glossary Flemish Standard Data Management Plan

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.*	Electronic and physical notebooks will be used to keep track of wet-lab experiments, and metadata. Electronic data will be stored and saved in a secured local drive. All datasets will be accompanied by a README.txt file containing all the associated metadata.
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: When results are published, metadata will be deposited in a trusted data repository, where metadata standards will apply. Controlled vocabulary will be applied. 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: local UZ Leuven drive (M Drive), only accessible by members of the lab of neurofibromatosis research.
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): Data stored at the local UZ Leuven drive (M drive) are backed up by the standard back-up provided by UZ Leuven ICTS.
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: / Yes, storages systems and back up are guaranteed by ICTS of UZ Leuven.

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 local M-drive is accessible only by laboratory members. This is a secure server is maintained by UZ Leuven.
What are the expected costs for data storage and backup during the research project? How will these costs be covered?	No costs are expected. Costs of UZ Leuven M-drive are covered by UZ Leuven - Department of Human Genetics.

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): Local UZ Leuven M-drive Published data will be stored in the KU Leuven research data repository. Data not suitable for data storage will be stored on a local server (sensitive personal data, non-published data).
What are the expected costs for data preservation during the expected retention period? How will these costs be covered?	No costs

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: Whenever possible, datasets and the appropriate metadata will be made publicly available through repositories that support data sharing.
If access is restricted, please specify who will be able to access the data and under what conditions.	Personal data will be restricted, and can only be accessed by members of the lab.

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: Personal data (such genetic information of patients) can not be shared according to GDPR guidelines.
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) ☒ Other (specify) Data of completed work will be published in academic peer reviewed journals and will as such be available in existing and relevant repositories
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 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?	Principle investigator Hilde Brems, and PhD researcher Seppe Van der Auweraer
Who will manage data storage and backup during the research project?	Principle investigator Hilde Brems
Who will manage data preservation and sharing?	The principle investigator Hilde Brems will handle data preservation and sharing after the project has ended.
Who will update and implement this DMP?	Principle investigator Hilde Brems, and PhD researcher Seppe Van der Auweraer