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	**Pieter Vrancaert** 0000-0001-6757-9350
Contributor name(s) (+ ORCID) & roles	**Inge Fourneau** (0000-0002-3794-4107), provider of human tissues **Aernout Luttun** (0000-0001-7902-9524), PhD supervisor **Mandy Grootaert** (0000-0003-1163-7499), PhD co-supervisor
Project number [1] & title	3M210531 - Unraveling the role of transcription factor Prdm16 in regulating the cellular and genetic landscape of atherosclerosis at single-cell resolution.
Funder(s) GrantID [2]	11P9W24N
Affiliation(s)	☒ KU Leuven ☐ Universiteit Antwerpen ☐ Universiteit Gent ☐ Universiteit Hasselt ☐ Vrije Universiteit Brussel ☒ Other: UZ 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	Our vascular system consists of an arterial and venous arm, lined by an endothelial cell (EC) layer. ECs exposed to different environments also have different features (known as EC heterogeneity) making them susceptible for vascular bed-specific diseases. Atherosclerosis is a highly prevalent artery-restricted condition during which lipid-laden lesions gradually obstruct blood flow and cause tissue loss. Our arterial system is designed in a way that arterial supply can be rerouted through collateral arteries to allow for flow recovery of the ischemic area at risk. The host lab found that lack of a single allele of the arterial-restricted transcription factor Prdm16 impaired this downstream recovery process, and Prdm16 in arterial ECs was required to preserve their function during flow recovery. It remains unknown whether and how endothelial Prdm16 plays a role in the upstream atherogenesis process. Therefore, in this project I will first document the expression pattern of Prdm16 in healthy and atherosclerotic arteries and study the phenotypic repercussions of EC-specific Prdm16 deficiency on atherosclerosis, using atheroprone low-density lipoprotein receptor (Ldlr)-/- and endothelial-specific Prdm16 knock-out mice. Next, I aim to determine the cellular and genetic mechanisms underlying the role of endothelial Prdm16 during atherosclerosis at single-cell resolution in these mice.

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	Digital Data Type	Digital Data Format	Digital Data Volume (MB, GB,	Physical Volume

---

[3]

						TB)	
Single-cell RNA sequencing raw data sets	Acquired from single-cell core facility	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images ☐ Sound ☐ Numerical ☒ Textual ☐ Model ☐ Software ☐ Other:	.fastq files, .gz files, .bam files	☐ < 1 GB ☒ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	
Single-cell RNA sequencing processed datasets	Processed data allowing downstream analysis of single-cell results	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images ☐ Sound ☐ Numerical ☐ Textual ☐ Model ☐ Software ☐ Other:	.xls files, .jpeg files, .pdf files, .txt files, .cvs files	☐ < 1 GB ☐ < 100 GB ☒ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	
Microscopic images of immunofluorescence and histochemic stainings	Images from both mouse and human tissues	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☒ Images ☐ Sound ☐ Numerical ☐ Textual ☐ Model ☐ Software ☐ Other:	.jpeg files, Zen .zvi files, .czi files, .tif files	☐ < 1 GB ☒ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	
Macroscopic images of atherosclerotic aorta segments	Images for plaque burden analysis	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☒ Images ☐ Sound ☐ Numerical	.jpeg files, .tiff files	☐ < 1 GB ☒ < 100 GB ☐ < 1 TB ☐ < 5 TB	

				☐ Textual ☐ Model ☐ Software ☐ Other:		☐ > 5 TB ☐ NA	
Blood chemistry reports	Measurements of plasma lipid levels	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images ☐ Sound ☒ Numerical ☐ Textual ☐ Model ☐ Software ☐ Other:	.xls files	☒ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	
Mouse genotyping gel pictures	Genotyping results from gel electroforesis	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☒ Images ☐ Sound ☒ Numerical ☐ Textual ☐ Model ☐ Software ☐ Other:	.jpeg files	☒ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	
Morphometric analysis data		☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☒ Images ☐ Sound ☒ Numerical ☐ Textual ☐ Model ☐ Software ☐ Other:	.xls files	☒ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	
Pseudonymized	Demographic	☒ Generate new	☒ Digital	☐ Audiovisual	.docx, .xls files	☒ < 1 GB	

personal data humans in different formats	data of included patients	data ☐ Reuse existing data	☐ Physical	☐ Images ☐ Sound ☐ Numerical ☒ Textual ☐ Model ☐ Software ☐ Other:		☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	
Statistical analysis		☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☒ Images ☐ Sound ☒ Numerical ☒ Textual ☐ Model ☐ Software ☐ Other:	Prism .pzfx files, R package .r files	☐ < 1 GB ☒ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	
Composition figures, digital images	Figures for abstracts, posters and publications	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☒ Images ☐ Sound ☐ Numerical ☐ Textual ☐ Model ☐ Software ☐ Other:	.eps files, Acrobat .pdf files, Adobe Indesign .indd files, Adobe Illustrator .ai files, .TIFF files, .jpeg files, .PNG files	☐ < 1 GB ☒ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	
SOPs	Staining, qPCR, Genotyping, nuclei isolation	☒ Generate new data ☒ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images ☐ Sound ☐ Numerical ☒ Textual ☐ Model	Word .doc files	☒ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	

				□ Software □ Other:			
quantitative (q)RT-PCR data		☒ Generate new data □ Reuse existing data	☒ Digital □ Physical	□ Audiovisual □ Images □ Sound ☒ Numerical ☒ Textual □ Model □ Software □ Other:	.cvs files	☒ < 1 GB □ < 100 GB □ < 1 TB □ < 5 TB □ > 5 TB □ NA	
Tissues	Paraffin embedded (human and animal origin)	☒ 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	Approximately 300 tissue blocks
Paraffin sections	Human and animal origin	☒ 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	30 drawers of 100 slides
Snap frozen tissues	Human and animal origin, stored at -80	☒ Generate new data □ Reuse existing	□ Digital ☒ Physical	□ Audiovisual □ Images □ Sound		□ < 1 GB □ < 100 GB □ < 1 TB	2 boxes of samples stored in -80 freezer

	degrees Celsius	data		☐ Numerical ☐ Textual ☐ Model ☐ Software ☐ Other:		☐ < 5 TB ☐ > 5 TB ☐ NA	
Cryo embedded tissues	Human and animal origin	☒ 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	5 boxes of samples stored in -80 freezer
Cryo sections	Human and animal origin	☒ 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	10 boxes of slides stored in -20 freezer

3. Research Data Summary

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.	/
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: S65580 ☒ Yes, animal data; provide ECD reference number: ECDN°P061/2021 ☐ 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: G-2021-3397/S65580 • Privacy Registry Reference: UZ Leuven patient registry • Short description of the kind of personal data that will be used: Data (health status, age, gender, medication and treatment history, body weight, body length) will be collected from patients with peripheral arterial disease who come to the Vascular Surgery unit at UZ Leuven (headed by Prof. Dr. I. Fourneau) for above-knee level limb amputation. Amputated tissue will be processed for histological analysis.
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:

---

[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: Dissemination and exploitation of patient data to 3rd parties in collaboration with I. Fourneau is restricted, and we will ensure privacy of the donors by pseudonymisation (a link between the patient pseudonym and their unique 'eenmalig administratief dossiernummer' or EAD number in the UZ Leuven patient registry will be secured with access right management only to the principal investigator and the health care professionals supervised by Dr. I. Fourneau).
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. 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.*	Main results and methods will be published in peer-reviewed journals (open access as required by FWO) and all publications will be archived in Lirias, the digital KU Leuven document repository. All **digital data** generated in the project for each WP and associated metadata will be archived digitally and a searchable database format (Excel or Access) will be implemented. Electronic lab note books will be used and templates have been designed for writing protocols/SOPs, for excel spreadsheets for raw data and (statistical) analysis. When raw data are uploaded on repositories, keywords will be affixed along with readme files containing the needed information for reuse. In the final stage of the project, a master index file with the combined metadata for each WP will be generated and archived on a non-editable drive of the host- institution KU Leuven ('K drive'). All **physical data** collected during the course of the project will be stored at designated storage places (human-derived material where possible will be registered and stored at the UZ Leuven Biobank) and location and preservation method of the biological samples (tissues, tissue sections, blood plasma, genetic material) will be documented digitally (.xls files).
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: Metadata will be a combination of machine-generated and manually generated metadata. Metadata of raw images (file size, pixel number, acquisition date, settings, etc.) and qRT-PCR are captured automatically and saved on the server together with the corresponding data files. Other metadata (on quantification procedures, biochemical analysis, etc.) are mostly captured manually and logged in lab notebooks or in searchable Excel/Access databases. For these metadata, we will progressively design own metadata standards using http://dublincore.rog/. We will also consider archiving our data using general data repositories (https://figshare.com/ and https://zenodo.org/). RNAseq data will be uploaded to the GEO repository which uses the MIAME standard.

5. 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: During the project **digital data** will be stored at different locations, depending on the type of data and accessibility. Non-personal data will be stored on the researchers' computers, on the KU Leuven network editable drives where the principal investigator has reserved dedicated space (the J drive for data that needs to be accessible daily and is exchangeable between research lab members or (later during the project) the L drive for longer-term storage of large data files that do not need to be frequently accessed). Personal patient- related data will be stored on the UZ Leuven central server. Both UZ Leuven and KU Leuven servers are compatible with GDPR regulations. All **physical data** collected during the course of the project will be stored at designated storage places (human-derived material where possible will be registered and stored at the UZ Leuven Biobank).
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): All of the data and documents on the researcher's computer are automatically synchronized to the KU Leuven Onedrive cloud with a capacity of up to 2TB per user. ☐ 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 ☐ 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*	Both UZ Leuven and KU Leuven network servers are compatible with GDPR regulations and allow for secure storage of personal data. KU Leuven ICTS services provide the option to control data access for authorized persons only (in this case, research lab members involved in this project). As mentioned above for personal data stored on the UZ Leuven central server, access will be restricted with access right management only to the principal investigator and the health care professionals involved in this project supervised by Dr. I. Fourneau. Secure storage of **physical data** at the Biobank is guaranteed by controlling access to the storage location. Access to this location will be limited to one person and one back-up person per research group.
What are the expected costs for data storage and backup during the research project? How will these costs be covered?	Costs for **digital data** storage and back-up during the project have been included in the research budget of the project. The current cost rate for the KU Leuven network drives are: 52€/y/100Gb block (J) and 6.4€/y/100Gb block (K), 869€/y/5Tb block. Costs for **physical data** are only applying for biological samples of human origin. The UZ Leuven Biobank is currently still in the process of calculating the yearly storage cost per sample.

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) All **digital data** and metadata will be retained for 5 years after the project (per the requirements of FWO); The same term will be applied to **physical data**. Long-term storage of personal data additionally requires GDPR clearance, which has been obtained upon approval from the Ethics Committee of UZ Leuven.
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): Digital data will be archived on the KU Leuven K drive for storage of read-only data.
What are the expected costs for data preservation during the expected retention period? How will these costs be covered?	Cost rate for storage on the K drive is 6.4€/year/100Gb. To store a total of 5 Tb for 5 years, the estimated cost hence is 1,600 €. Costs will be allocated to the project budget.

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/#INF OEUREPO-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: Main findings of the research with all supporting processed data will be made available through publication in peer-reviewed journals with open access policies (as required by FWO). All manuscripts will also be deposited in the KU Leuven Lirias digital repository. Raw RNAseq data will be made available publicly upon acceptance of the manuscript. Other raw data related to published manuscripts may be available upon specific request as will be stated in a data availability statement included in the published manuscripts.
If access is restricted, please specify who will be able to access the data and under what conditions.	/
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 will be shared only with certain third parties (as will be specified in the GDPR addendum to the informed consent form) if needed, thereby always ensuring the privacy of the donors.

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): Open Access repository ☒ Other (specify): Raw RNAseq data will be made available publicly through the GEO repository upon acceptance of the manuscript.
When will the data be made available?	☒ Upon publication of research results ☐ Specific date (specify) ☒ Other (specify): Other data will be made available upon request, where considered appropriate, following 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.* 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?	For sharing **digital data**, no sharing costs are foreseen. For sharing **physical data**, Material Transfer Agreements will have to be put in place which will be mutually signed. Shipping costs would be covered by either party (through the FWO budget in case of the provider) as long as the costs are low, however, significant sharing costs will be expected to be borne by the requestor.

7. Responsibilities	
Who will manage data documentation and metadata during the research project?	All researchers involved in this project.
Who will manage data storage and backup during the research project?	All researchers involved in this project.
Who will manage data preservation and sharing?	The supervisor (A. Luttun).
Who will update and implement this DMP?	The supervisor (A. Luttun)/the grant holder (P. Vrancaert).