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	Prof Olivier Govaere 0000-0002-4426-6930
Contributor name(s) (+ ORCID) & roles	
Project number [1] & title	206953_ELUCIDATING MACROPHAGE HETEROGENEITY AND IMMUNOMETABOLISM IN NON-ALCOHOLIC FATTY LIVER DISEASE
Funder(s) GrantID [2]	3M230703
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	The global increase in obesity has led to a dramatic rise in the prevalence of non-alcoholic fatty liver disease (NAFLD), affecting approximately 25% of the adult population worldwide. NAFLD is associated with a substantial socioeconomic burden, which, coupled with rising prevalence, is a growing public health challenge. Patients with NAFLD are not only at high cardiovascular risk but show a substantially higher incidence of cancer. Liver macrophage immune cells can eat lipids, cholesterol and apolipoproteins, and even damaged fatty liver cells. This digestion leads to foamy cell formation and facilitates recruitment of more inflammatory cells to the liver. However, how this formation and recruitment of other cells in the liver of obese patients is still unknown. During this project, we aim to investigate the different macrophage populations and their specific locations in different stages of NAFLD by using state-of-the-art spatial transcriptomics screening methods. These findings will be validated functionally and linked with clinical data to establish a non-invasive blood test to identify patients at risk of developing end-stage liver disease.

2. Research Data Summary

List and describe all datasets or research materials that you plan to generate/collect or reuse during your research project. For each dataset or data type (observational, experimental etc.), provide a short name & description (sufficient for yourself to know what data it is about), indicate whether the data are newly generated/collected or reused, digital or physical, also indicate the type of the data (the kind of content), its technical format (file extension), and an estimate of the upper limit of the volume of the data [3].

				ONLY FOR DIGITAL DATA	*ONLY FOR DIGITAL DATA*	*ONLY FOR DIGITAL DATA*	*ONLY FOR PHYSICAL DATA*
Dataset Name	Description	New or Reused	Digital or Physical	Digital Data Type	Digital Data Format	Digital Data Volume (MB, GB, TB)	Physical Volume
GeoMX-CosMX nanoString spatial transcriptomics	RNA sequencing of labelled barcodes in a spatial context using the nanoString and Illumina technology (resulting in reads on image location). Reads will be analysed using R.	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images ☐ Sound ☒ Numerical ☐ Textual ☐ Model ☐ Software ☐ Other:	.pdf/.png/.tif/.xlsx	☐ < 1 GB ☐ < 100 GB ☒ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	
MILAN data	Data generated through proteome multiplexing and scanning with the Zeiss AxioScan2.	New	Digital	Images and numerical	.tif/.rds/.pdf/.zen	<5TB	

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

	Analysed using R.						
Microscopy assay	Imaging Acquisition using AxioScan 2 and Leica brightfield. Zen (Zeiss), LAS X (Leica), Photoshop & Illustrator (Adobe), ImageJ and QuPath softwares	New	Digital	Images	.zen/.lif/.tif/.jpeg/ .psd	<100 GB	
Proteomics	Reads generated by the SomaLogic platform (data available Govaere et al Nature Metabolism 2023). Analysis using IBM SPSS.	Reuse existing data	Digital	Numeric	.xlsx/.sav/.spv	<1GB	
Transcriptomics	Data generated by Illumina Nextgen 500 platform (available at	Reuse existing data	Digital	Numeric	.fastq/.fast/.xlsx	<1TB	

	GSE135251). Analysed using R.						
Publications and presentation data	Publications, and presentation of data Publications generated using Microsoft Word, Endnote, Powerpoint, and Illustrator.	New	Digital	Numerical	docx/.pptx/.enl/. pzfx/.psd/.ai	<1GB	

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.	• Proteomics DOI: 10.1038/s42255-023-00775-1 • Transcriptomics DOI: 10.1126/scitranslmed.aba4448

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: S67418 ☐ Yes, animal data; provide ECD reference number: ☐ Yes, dual use; provide approval number: ☐ No Additional information:
Will you process personal data[4]? If so, please refer to specific datasets or data types when appropriate and provide the KU Leuven or UZ Leuven privacy register number (G or S number).	☒ Yes (provide PRET G-number or EC S-number below) ☐ No Additional information: S67418
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:
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*.	1. Microscopy and MILAN data from human slides will be saved on our KULeuven NAS server at IT O&N4 with back-up provided by KU Leuven ICTS. 2. Microscopy data from single slides and ex vivo experiments will be saved in a shared drive accessible by all lab members involved in the project. The number of experiment, protocols and stainings will be described in detail in lab notebooks and will also be available in the shared drive. 3. Raw data from the nanoString GeoMX and CosMX will be saved in a shared drive identified by number of experiment and also indicated in the lab book. Protocols and methodology used will be attached with a clear description to facilitate reproducibility at any time.
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: All datasets will be described and summarized in an excel file. In addition, all lab members will have access to this file to be able to find, interpret, use and reproduce the data generated if necessary. Metadata will be saved onto KULeuven LabCollector for our lab. 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:
How will the data be backed up? *WHAT STORAGE AND BACKUP PROCEDURES WILL BE IN PLACE TO PREVENT DATA LOSS?*	☒ Standard back-up provided by KU Leuven ICTS for my storage solution ☐ Personal back-ups I make (specify) ☐ Other (specify)
Is there currently sufficient storage & backup capacity during the project? If yes, specify concisely. If no or insufficient storage or backup capacities are available, then explain how this will be taken care of.	☒ Yes ☐ No Yes: An unlimited storage space is available and maintained by the ICTS-IT department. 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*	Research data are stored and managed by the KU Leuven IT department and are accessible only by the researchers working on the project.

What are the expected costs for data storage and backup during the research project? How will these costs be covered?	Back-up costs of 1 TB (KU Leuven ICTS) 113.84 euros/year. The lab budget will cover storage and back up costs.

5. Data Preservation after the end of the Research Project	
Which data will be retained for at least five years (or longer, in agreement with other retention policies that are applicable) after the end of the project? In case some data cannot be preserved, clearly state the reasons for this (e.g. legal or contractual restrictions, storage/budget issues, institutional policies...). *Guidance on data preservation*	☒ All data will be preserved for 10 years according to KU Leuven RDM policy ☐ All data will be preserved for 25 years according to CTC recommendations for clinical trials with medicinal products for human use and for clinical experiments on humans ☐ Certain data cannot be kept for 10 years (explain)
Where will these data be archived (stored and curated for the long-term)? *Dedicated data repositories are often the best place to preserve your data. Data not suitable for preservation in a repository can be stored using a KU Leuven storage solution, consult the interactive KU Leuven storage guide.*	☐ KU Leuven RDR ☒ Large Volume Storage (longterm for large volumes) ☐ Shared network drive (J-drive) ☐ Other (specifiy):
What are the expected costs for data preservation during the expected retention period? How will these costs be covered?	Yearly storage costs of 1TB data on the K-drive: 56.92 euros. Costs will be covered by internal lab funding.

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:
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:

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 will be published using open access publications and will be available at dedicated data in supplementary files, in particular for GeoMX and CosMX nanoString data, which will be provided as normalised reads. Unpublished research data will be accessible to the PI's group and all scientific collaborators involved in the project
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?	We don't expect any costs regarding data sharing.

7. Responsibilities	
Who will manage data documentation and metadata during the research project?	Prof Olivier Govaere (PI) and Dr Asier Antorantz (head MILAN bioinformatics team TCTR lab)
Who will manage data storage and backup during the research project?	ICTS-IT department (KU Leuven)
Who will manage data preservation and sharing?	Prof Olivier Govaere (PI)
Who will update and implement this DMP?	Prof Olivier Govaere (PI) bears the end responsibility of updating & implementing this DMP.