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	Berkehür Abaylı (0000-0002-8530-4227)
Contributor name(s) (+ ORCID) & roles	Hugo Vankelecom (0000-0002-2251-7284) Laura Van Gerven (0000-0002-5325-7956) Emma Laporte (0000-0003-0799-3116)
Project number [1] & title	11P8X24N Unraveling the pituitary stem cells' biology in the injured, regenerating and aging gland
Funder(s) GrantID [2]	Fonds voor Wetenschappelijk Onderzoek – Research Foundation Flanders (FWO)
Affiliation(s)	■ KU Leuven ☐ Universiteit Antwerpen ☐ Universiteit Gent ☐ Universiteit Hasselt ☐ Vrije Universiteit Brussel ☐ Other: ROR identifier KU Leuven: 05f950310
Please provide a short project description	We aim to decipher the processes that unfold in the pituitary upon local damage, in particular the molecular and cellular mechanisms that underlie pituitary stem cell activation and subsequent regeneration, and their deterioration at aging. Here, we will focus on specific appealing aspects recently uncovered through our single-cell (sc) transcriptomic interrogations and will further expand this to in vivo (mouse) and in vitro (organoid) models. Moreover, we will define the translational power of our investigation by transposing the mouse findings to the human pituitary.

[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.

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
		☐ 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	
Cryopreserved human pituitary tissue.	Pituitary biopsy of healthy post-mortem patients.	New	Physical				24 relevant biological samples with each approximately 2 vials of primary tissue/cells
Cryopreserved human pituitary organoids.	Organoids derived from healthy post-mortem patients	New	Physical				24 relevant biological samples with each approximately 15 vials of organoids

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

C57Bl6/J male mice	Dissection of pituitary for downstream analyses, including: - Validation of transcriptomic findings - Establishment of pituitary organoid cultures.	New	Physical				40 male mice
GhCre/+;R26iDTR/+ male mice	Dissection of pituitary for downstream analyses, including: - Validation of transcriptomic findings - Establishment of pituitary organoid cultures.	New	Physical				40 male mice
Sox2CreERT2/+ mice	Breed animals with pituitary stem cell-specific knock-out (KO) for	New	Physical				80 mice

	Stat3 gene (by crossing them with *Stat3fl/fl* mice).						
Stat3fl/fl mice	Breed animals with pituitary stem cell-specific KO for *Stat3* gene (by crossing them with Sox2CreERT2/+ mice). Assess the role of *Stat3* in pituitary stem cell behaviour.	New	Physical				80 mice
Sox2eGFP/+ mice	FACS sort SOX2+ pituitary stem cells.	New	Physical				20 mice
Fixed samples	Paraformaldehyde (PFA)-fixed mouse/human pituitaries and derived organoids.	New	Physical				<400
Microtome sections	Rotary microtome sections obtained from	New	Physical				<1000

	PFA-fixed samples.						
RNA	RNA from pituitary samples and derived organoids.	New	Physical				<1000
cDNA	cDNA from pituitary samples and derived organoids.	New	Physical				<1000
RNA/DNA concentration/ quality	Information obtained after RNA extraction via measurement with Nanodrop.	New	Digital	Experimental	.xlsx	<100 MB	
PCR results	Gel electrophoresis (gel image) obtained via Image Lab software.	New	Digital	Experimental	.tif	<1 GB	
Light, epifluorescence, THUNDER, multi-photon, light-sheet and	Images from sections of organoids and pituitary samples.	New	Digital	Images	.lif, .lsm and .tiff files	< 200 GB	

confocal images							
RT-qPCR data/graphs	RT-qPCR data/graphs from gene expression analysis of pituitary tissue and organoids.	New	Digital	Numerical	xlsx, .eds, .pzfx	< 10 GB	
scRNA-seq dataset	Single-cell (sc) RNA-seq data of pituitary and organoids from both human and mice.	New & reuse of existing data	Digital	Textual	.fastq files	< 1 TB	
scRNA-seq output	scRNA-seq output of pituitary and organoids from both human and mice.	New & reuse of existing data	Digital	Numerical Textual	.html, .txt, .pdf files	< 100 MB	
Experimental analysis data and manuscripts	Analysis of obtained data summarized in presentations/ excel/word files.	New	Digital	Numerical	.xls, .txt files	< 100 MB	
Biopsy and organoid	Biopsy and organoids	New	Digital	Textual	.xls files	< 100 MB	

biobank database	biobank database.						
Lab books	Notes on experiments, observations in the lab.	New	Physical				< 20 books

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.	Use of published scRNA-seq datasets of mouse and human pituitary biopsies. Mouse : - DOI: 10.1073/pnas.2100052118 - DOI: 10.7554/eLife.75742 Human : - DOI: https://doi.org/10.1016/j.celrep.2022.110467 - DOI : https://doi.org/10.1038/s41467-020-19012-4
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: S68961 ☒ Yes, animal data; provide ECD reference number: P165/2023 ☐ Yes, dual use; provide approval number: ☐ No Additional information: EC application for human pituitary samples is under revision by the Ethical Committee.

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: PRET: G-2023-7086 EC: S68961
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:

3. Documentation and Metadata

[4] See Glossary Flemish Standard Data Management Plan

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*.	- Daily lab activities are recorded in detail in the lab book and/or Microsoft OneNote. - Documentation of the mice: in an excel file the following information will be noted for every mouse: cage number, date of birth, sex, derived from which breeding couple, genotype, euthanising date, used in which experiment. - Overview of experiments (date, topic, sample origin, experimental set-up, read-outs) are saved on the KU Leuven One drive. - For documentation of microscopy images (of organoid cultures) the following information will be noted: date, experimental condition, passage of organoid culture, number of days in culture, magnification used, antibodies/dyes. Images will be saved on the shared drive of the lab and KU Leuven One drive in a designated folder of the particular experiment. Within the experiment folder, additional folders are labelled in a clearly structured way (according to different experimental conditions or different timepoints within the experiment). The setup of an experiment is written down in the lab book. A meta data file, generated by the microscope programme, is saved automatically together with the image. - For RNA and cDNA concentration and quality measurements using the nanodrop: 260/230 and 260/280 ratios (quality measure) and concentrations are written down in lab book and later transferred manually to an excel file where all previous RNA/cDNA measurements are stored. Date of measurement together with name of the sample is included. The location of each sample (which freezer and which box) is also included in this excel file. - For qPCR data: excel file containing sample setup, raw data, results, melt curve data are given the name: "Date_experiment number_sample name(s)_general gene list_qPCR". The qPCR data is saved in a "qPCR folder" within the 'Raw data' folder, together with the template of the particular qPCR reaction. The analysed data can be found in the designated folder of the particular experiment within the 'Analysed data' folder. Graphs from the data are made using Graphpad Prism (.pzfx file). File is named: "Experiment number_sample name(s)_general gene list_graphpad" and saved in the same folder. - For scRNA-sequencing scripts/figures: .html, .pdf, .txt files containing scripts or figures of typical scRNA-sequencing workflow. Each script is saved as 'Experiment number_vignette type used'. Each figure is saved as 'Experiment number_sample name(s)_figure type'. - Methodology and protocols for RNA extraction, cDNA preparation, immuno-histochemistry stainings, organoid culture, medium preparation, ... are all included in the lab book. In the table of

	contents of the lab book, the page number of each protocol included in the lab book can be found. In addition, the start of each experiment is indicated in the table of contents. - Biobank documentation: Cryopreserved tissue samples and organoid lines from patients will be stored in UZ/KU Leuven Biobanks. For each patient, the following data will be collected: sex, date of birth/age, BMI, co-morbidities, patient ethnicity, cause of death, brain damage, period spent in ICU, surgery-related information (period of post-mortem interval, intracranial pressure, internal bleeding, etc.), anatomical abnormalities and medical/medication history. All pituitary biopsies will be clearly labeled with an original code: PB_#_age_sex_date in which: PB = pituitary biopsy, # = number of patients, age = age of the patient, sex = sex of the patient, date = respective date of the sample collection.
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: At the moment, metadata standards are not implemented in the research group. Metadata generated by microscopy, RT-PCR analyses and from sequencing data, is saved to KU Leuven OneDrive. Saved data is further subdivided in a clearly structured way (e.g. specific folders for different experiments). Additionally, in the lab books, a description of every experiment can be found including all the experimental conditions.

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: The large scRNA-seq data will be stored on the lab's storage space of the Flemish Super Computer VSC.
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 (Personal external hard drive) ☐ 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*	Access to the shared KU Leuven OneDrive of our lab is secured by a login with the personal u-number and password. An extra layer of precaution will be taken by activating the multifactor authenticator app provided by the KU Leuven. There is also a password on the personal computer of the FWO fellow.
What are the expected costs for data storage and backup during the research project? How will these costs be covered?	As long as the data does not exceed the 2 TB of storage of the KU Leuven OneDrive, no additional costs for data preservation are expected. If the storage capacity unexpectedly exceeds 2 TB, KU Leuven provides a large volume storage for research data in a cost-efficient manner: 104.42 euro/TB/year (to be purchased in blocks of 5 TB).

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?	As long as the digital data does not exceed the 2TB of storage of the KU Leuven OneDrive, no additional costs for data preservation are expected.

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 The following datasets will be made available: - Fluorescence/brightfield images - RT-qPCR data - scRNA-seq datasets ☐ Yes, as embargoed data (temporary restriction) ☐ Yes, as restricted data (upon approval, or institutional access only) ☒ No (closed access) Following data will remain closed: - Personal information of patients (such as name, surname, address) ☐ Other, please specify:
If access is restricted, please specify who will be able to access the data and under what conditions.	Only staff of Intensive Care Unit in UZ Leuven will be able to reach the personal information about patients (such as the name and surname of the patient).
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 information about patients.

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) The obtained data (fluorescence/brightfield images, qPCR data) in the project will be made available through publications and the PhD Thesis. The scRNA-seq data will be made available on ArrayExpress after publication.
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) Data can be requested after signing a data sharing agreement (Attribution 4.0 International (CC by 4.0)). Public availability after publishing the data will also depend on the journals policy (postpublication data repository).
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?	There are currently no expected costs for data sharing.

7. Responsibilities	
Who will manage data documentation and metadata during the research project?	The grant holder, Berkehür Abaylı.
Who will manage data storage and backup during the research project?	The grant holder, Berkehür Abaylı.
Who will manage data preservation and sharing?	The PI (Prof. Dr. Hugo Vankelecom) will be responsible for ensuring data preservation and sharing.
Who will update and implement this DMP?	The grant holder (Berkehür Abaylı) will be responsible for updating this DMP. The PI (Prof. Dr. Hugo Vankelecom) bears the end responsibility for updating and implementing this DMP.