Unraveling the cellular landscape and (epi)genetic mechanisms at single-cell resolution and the effect of Prdm16 deficiency during the pathogenesis of pulmonary arterial hypertension

A Data Management Plan created using DMPonline.be

Creator: Aernout Luttun

Affiliation: KU Leuven (KUL)

Funder: Fonds voor Wetenschappelijk Onderzoek - Research Foundation Flanders (FWO)

Template: FWO DMP (Flemish Standard DMP)

Principal Investigator: n.n. n.n., Aernout Luttun, Rozenn Quarck

Data Manager: Aernout Luttun, Rozenn Quarck

Project Administrator: Aernout Luttun

Grant number / URL: G060923N

ID: 199185

Start date: 01-01-2023

End date: 30-12-2026

Project abstract:

Our vascular system has two loops, the systemic circulation for distribution of oxygen-rich blood to all organs and a pulmonary circulation for re-oxygenation of blood coming from these organs. Both loops have fundamentally different characteristics and it remains undetermined whether the endothelial cells lining the inside of the pulmonary circulation have a unique reaction to diseases. Pulmonary arterial hypertension (PAH) specifically affects the pulmonary arterial circulation whereby an increase in vascular resistance leads to remodelling and dysfunction of the right heart ventricle and eventually death if left untreated. While genetic defects – mostly mutations in bone morphogenetic (BMP) pathway genes – play a role in the pathogenesis of PAH, additional genes and epigenetic factors have also been implicated. While our previous findings have revealed that deficiency for the arterial transcription factor Prdm16 compromises arterial flow recovery in the systemic circulation, it is unknown whether and how Prdm16 deficiency affects the pathogenesis of PAH. Here, we will use a multi-omics approach at single-cell resolution to (i) unravel the (epi)genetic mechanisms behind PAH in mice; (ii) investigate how these mechanisms are perturbed by (endothelial) Prdm16 deficiency; (iii) where possible, validate our findings in newly developed transgenic rats and human tissue. Altogether, we expect this integrated approach will offer clues to develop more effective PAH treatments.

Last modified: 20-06-2023

Unraveling the cellular landscape and (epi)genetic mechanisms at single-cell resolution and the effect of Prdm16 deficiency during the pathogenesis of pulmonary arterial hypertension
FWO DMP (Flemish Standard DMP)

1. 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.

1. Newly generated data (overall estimated volume: **5 Tb**)
1.1. Digital:
- single-nuclei RNA and assay for transposase accessible chromatin sequencing (snRNA/ATACseq) raw data sets generated in rodents (.fastq files, .gz files, .bam files), snRNA/ATACseq processed datasets (.xlsx files, .jpeg files, .pdf files, .txt files, .cvs files) (work package(**WP**)**1** and **WP3** studies)
- echocardiography data (.bimg files generated by Vevo Lab software) and images in rodents (WP2 studies)
- hemodynamics data (.adicht files) in rodents (**WP2** studies)
- vaso-reactivity data on rodents (.xlsx files) (**WP3** studies)
- microscopic images of immunofluorescence stainings or in situ hybridisation on rodent tissues (.jpeg files, Zen .zvi files) (**WP1, 2** and **WP3** studies)
- microscopic images of human cells (.jpeg files, .zvi files, .tiff files) (**WP2** studies)
- microscopic images of immunofluorescence stainings or in situ hybridisation on human tissues (.jpeg files, .zvi files, .tiff files) (**WP1** and **WP3** studies)
- electron microscopic images in rodents (.jpeg files, .tiff files) (**WP3** studies)
- blood chemistry reports (.xlsx files) (WP2 studies)
- morphometric analysis data (.xlsx files) (**WP1, 2** and **WP3** studies)
- Western-blot gel pictures (.jpeg files) (**WP3** studies)
- mouse genotyping gel pictures (.jpeg files) (**WP1, 2** and **WP3** studies)
- pseudonymised personal data humans in different formats (anthropometric data, clinical data) (**WP1** and **WP3** studies)
- statistical analysis (Prism .pzfx files, R package .r files) (**WP1, 2** and **WP3** studies)
- composition figures, digital images (.eps files, Acrobat .pdf files, Adobe Indesign .indd files, Adobe Illustrator .ai files, .TIFF files, .jpeg files, .PNG files) (**WP1, 2** and **WP3** studies)
- electronic lab note books (OneNote) (**WP1, 2** and **WP3** studies)
- hypoxia cage monitoring data in rodents (generated by Anawin software)
- standard operation procedures (SOPs; Microsoft word .doc files) (**WP1, 2** and **WP3** studies)
- informed consent forms (.pdf files) (**WP1** and **WP3** studies)
- sample inventories (.xlsx files) (**WP1, 2** and **WP3** studies)
- quantitative (q)RT-PCR data on mouse tissues or cells (.cvs files) (**WP1** and **WP2** studies)
- qRT-PCR data on human tissues or cells (.cvs files) (**WP1** and **WP3** studies)
1.2. Physical:
- single-nuclei suspensions (rodent origin) (**WP1** and **WP3** studies)
- tissues (paraffin-embedded, rodent origin) (**WP1, 2** and **WP3** studies)
- protein/RNA/cDNA (rodent origin) (**WP1,2** and **WP3** studies)
- protein/RNA/cDNA (human origin) (**WP1** and **WP3** studies)
- genomic DNA samples (rodent origin) (**WP1, 2** and **WP3** studies)
- lentiviral particles (**WP3** studies)
- paraffin sections (rodent origin) (**WP1, 2** and **WP3** studies)
- frozen pulmonary endothelial cells (human origin) (**WP1** and **WP3** studies)
- cryo sections (rodent origin) (**WP1, 2** and **WP3** studies)
- tissues (snap-frozen or paraffin-embedded, human origin) (**WP1** studies)
- paraffin sections (human origin) (**WP1** and **WP3** studies)
- cryo sections (human origin) (**WP1** and **WP3** studies)
- plasma (human origin; two samples per patient procedure) (**WP1** and **WP3** studies)
- blood (human origin; 3 samples per patient procedure) (**WP1** and **WP3** studies)
- Swan-Ganz pulmonary arterial catheters (1 per patient procedure) (**WP3** studies)

2. Existing data (generated by the host lab):
Single-cell RNA sequencing raw data sets (.fastq files, .gz files, .bam files) published and generated in the context of a previous FWO project.

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:

The reused scRNAseq raw data sets (.fastq files, .gz files, .bam files) published and generated in the context of a previous FWO project are uploaded on the ArrayExpress database at EMBL-EBI (www.ebi.ac.uk/arrayexpress) under public accession number E-MTAB-9703).

Are there any ethical issues concerning the creation and/or use of the data (e.g. experiments on humans or animals, dual use)? Describe these issues in the comment section. Please refer to specific datasets or data types when appropriate.

- Yes, human subject data
- Yes, animal data

Ethical committee approval: is required for rodent-related studies: two ethical commission dossiers (ECDs) for rodent-related studies have been approved by the Ethics Committee of the KU Leuven Animalium (ECDN° P094/2023 and ECDN° Creation Luttun/2023); ethical approval for use of human tissues will be submitted in the second half of the first year (2023) to the Ethical Committee of UZ Leuven after obtaining an S number from the Clinical Trial Center. Approval will also depend on compliance with GDPR/PRET regulations. Online questionnaires related to GDPR and PRET will be submitted. In parallel, approval from the UZ Leuven Biobank for sample storage and general use has already been obtained (S57114). French PAH patients, whose biological samples and clinical data will be used in this project, are part of the French Network on Pulmonary Hypertension and have provided written informed consent (Protocol N8CO-08-003, ID RCB: 2008-A00485-50, approved on June 18, 2008).

Will you process personal data? If so, briefly describe the kind of personal data you will use in the comment section. Please refer to specific datasets or data types when appropriate.

- Yes

Personal data include demographic information (age, gender, length, body weight, body mass index), medical information (hemodynamics, exercise capacity (6-min walk distance or 6MWD), echocardiography, aetiology, mortality/survival, NYHA functional class) and general health status (smoking behaviour) and therapy (medical treatment). Data are exported from UZ Leuven medical records ('KWS'). M. Delcroix (co-promoter) and Rozenn Quarck (clinical data manager of BREATHE) are responsible for data collection and management. These personal data are categorised as 'special personal data' of vulnerable individuals and will therefore be treated with caution.

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

Potentially; dataset /data type currently not yet defined.

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 in the comment section to what data they relate and what restrictions are in place.

- Yes

- Since this project encompasses a collaboration with a non-KU Leuven partner (Inserm, France), agreements related to exploitation and dissemination have been written down in a collaboration agreement signed on March 22nd, 2023 by all parties involved in the project.
- Dissemination and exploitation of patient data to 3rd parties obtained in collaboration with Prof. M. Delcroix is restricted, and we will ensure privacy of the patients by pseudonymisation. A link between the patient and the dataset will be encoded. This code will not contain information that could lead to the identification of the patient and will be stored in an encrypted way on a different location than the pseudonymised data. The unique 'eenmalig administratief dossiernummer' (EAD number) in the UZ Leuven patient database will not be used. The code will only be accessible to Marion Delcroix, co-promoter and Rozenn Quarck, the clinical data manager of BREATHE.
- Human data from French PAH patients will not be sent to other participants of the project. Generated raw data from pseudonymised samples will be sent to Prof. David Montani who will depseudonymise and analyse the data. Prof. Montani, partner of this project with Dr. Frédéric Perros, is Professor of Respiratory Medicine at the French National Referral Centre for Pulmonary Hypertension in the Department of Respiratory Medicine, at the Hôpital Bicêtre, Le Kremlin-Bicêtre, Paris-Sud University, France and team leader in the INSERM U999 laboratory « Pulmonary Hypertension: Pathophysiology and Novel Therapies», Hôpital Marie Lannelongue, Le Plessis Robinson.

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 in the comment section to what data they relate and which restrictions will be asserted.

- Yes

Since this project encompasses a collaboration with a non-KU Leuven partner (Inserm, France), agreements related to IP rights and ownership have been written down in a collaboration agreement signed on March 22nd, 2023 by all parties involved in the project.

2. 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).

Main results and methods will be published in peer-reviewed journals (open access as required by FWO regulations) 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 (at room temperature or frozen) and location and preservation method of the biological samples (tissues, tissue sections, blood plasma, genetic material) will be documented digitally (.xlsx files).

Will a metadata standard be used to make it easier to find and reuse the data? If so, please specify (where appropriate per dataset or data type) which metadata standard will be used. If not, please specify (where appropriate per dataset or data type) which metadata will be created to make the data easier to find and reuse.

- Yes

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/). RNA/ATACseq data will be uploaded to the GEO repository which uses the MIAME standard.

3. Data storage & back-up during the research project

Where will the data be stored?

Data storage:

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 for this project (the J drive for data that needs to be accessible daily and is exchangeable between KU Leuven-affiliated project participants 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.

In regards to our Inserm partner and the storage and accessibility of the digital data for the project, the data will be securely stored on the server of Université Claude Bernard Lyon 1 (UCBL). Our Inserm partner will have the ability to access and analyse this data using a desktop computer that is specifically provided by the CarMeN laboratory (INSERM U.1060/Université Lyon1/INRAE). This computer is connected to the UCBL server. Our Inserm partner possesses an account with login credentials and a password, which have been provided by the server host. This account grants the beneficiary access to their individual work session on their computer, as well as their personal directory on the server. It is important to note that this directory is solely accessible by the beneficiary, and remains invisible to any other user of the server. This setup ensures the confidentiality and privacy of the data, as only the beneficiary possesses the necessary rights and permissions to access their personal directory.

All **physical data** collected during the course of the project will be stored at designated storage places. An inventory of each storage place is available.

How will the data be backed up?

Data back-up:

Researchers' computers are backed-up through constant synchronisation to the Enterprise Box cloud of the KU Leuven (which provides 100 Gb storage per KU Leuven researcher). The KU Leuven network drives and the UZ Leuven central server are automatically backed-up on a daily basis. Continuous KU Leuven network drive back-up is guaranteed and supervised by the KU Leuven ICT service and makes use of 'snapshot' technology. Once data sets do no longer need to be accesses and/or modified, (e.g., after publication of manuscripts), archiving to a read-only KU Leuven network drive (the K drive) will be done to maintain a copy.

At Inserm, the data will be backed-up locally on the server, creating a retention point for each day, both in the evening and in the morning. This ensures that each modified file will store its previous version securely on the local server. To provide additional protection against potential hardware failures of the storage array, a second server has been implemented. This server is connected to a separate storage array housed in a different server room. As a precautionary measure, the data will be replicated to this secondary server on a weekly basis, further safeguarding against potential data loss.

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

Digital data: Each researcher will be given access to 100 Gb storage space on the KU Leuven Enterprise Box. This should be sufficient, except for large volume datasets (e.g., microscopy images, sequencing datasets). For large volume datasets, space will be reserved on the editable KU Leuven network drives (J or L). Also, space will be reserved on the read-only K drive for storage after the end of the project or after publication of manuscripts. Storage on the J and K KU Leuven network drives is extendable by blocks of 100 Gb), storage on the L drive is extendable per 5 Tb, hence, by acquiring additional storage space based on the project's requirements, sufficient storage can be made available.

Our Inserm partner has been granted a substantial storage allocation of 20 Terabytes (TB) on the UCBL server. This ample storage capacity allows him to efficiently store and manage a significant volume of digital data related to the project. With such extensive storage capabilities, our partner can ensure the preservation and accessibility of the project's data without encountering any limitations.

For storage of **physical data**, sufficient storage space is available.

How will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?

Digital data: The KU Leuven Enterprise Box cloud is suitable for secure storage of non-personal data. Both UZ Leuven and KU Leuven network servers are compatible with GDPR regulations and allow for secure storage of personal data. The access to the KU Leuven server is u-number and password controlled. KU Leuven ICTS services provide the option to control data access for authorised persons only (in this case, KU Leuven affiliated 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 clinical data manager Rozenn Quarck and by co-promoter Prof. M. Delcroix.

At Inserm, it is important to note that the computer connected to the secure data storage system will not be used for hosting any websites. Likewise, no data from the project will be made available on any website. The responsibility for maintenance lies with the IT Department of the UCBL, and the server itself is covered by a manufacturer's warranty from DELL. To ensure smooth operations and support, our faculty benefits from the dedicated assistance of a DSI agent who promptly addresses any technical issues or concerns that may arise. The IT Department of the UCBL takes charge of maintaining the security of the IT channels. Their comprehensive network infrastructure, known as the Lyres network, provides the necessary safeguards for protecting data during transmission and storage. Furthermore, the CarMeN laboratory, where our partner is based, has an internal VLAN within the LYRES network. This VLAN effectively isolates the laboratory's network from the rest of the network, enhancing security measures and ensuring restricted access solely to authorised personnel. Furthermore, it is important to note that no data is stored on the individual work stations used within the laboratory. Instead, all data is securely protected and stored on the server. This practice ensures the centralisation and proper management of data, minimising the risk of data loss or unauthorised access. To access the university's LYRES network, the work stations have dedicated sessions that are opened when needed. These sessions provide authorised personnel with secure access to the network, allowing them to retrieve and interact with the stored data as required for their work. In line with the university's security procedures, passwords are regularly changed to maintain a high level of security. This proactive approach to password management helps mitigate potential risks associated with unauthorised access. Additionally, when the computer is not actively used, an automatic screen saver is activated, requiring a password to resume normal operation. This added layer of protection ensures that sensitive data remains secure even when the workstation is temporarily unattended. These robust measures collectively contribute to safeguarding the integrity, confidentiality, and accessibility of the project's data.

All **physical data**, printed forms and notebooks are stored in the labs in locked cabinet. Access to the lab is secured and badge controlled. In case samples are stored at the Biobank, secure storage 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: 25€/y/100 Gb block (KU Leuven Enterprise Box); 503.66€/y/1 Tb block (J-drive) and 6.4€/y/100Gb block (K-drive), 869€/y/5Tb block (L-drive).

At Inserm, the costs associated with digital storage have already been fully covered for the next five years. As a result, there will be no additional charges or expenses anticipated during this period. This prepayment ensures that our project has secure and uninterrupted access to the required digital storage resources, without any financial burden or unexpected costs. We can confidently rely on the allocated storage capacity without concerns about future financial implications.

Costs for **physical data** are only applying for biological samples of human origin in case they are stored in the Biobank. The UZ Leuven Biobank is currently still in the process of calculating the yearly storage cost per sample. At Inserm, the costs associated with the physical storage are already incorporated within the laboratory's day-to-day operational expenses. Therefore, no specific charges have been allocated or imposed for our project.

4. 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...).

All digital data and metadata will be retained for 10 years after the project (per the requirements of Research Data Management policy of the KU Leuven). The same term will be applied to physical data. Long-term storage of personal data additionally requires GDPR clearance, which will be obtained upon approval from the Ethics Committee of UZ Leuven. The same procedure applies for our Inserm partner.

Where will these data be archived (stored and curated for the long-term)?

For the KU Leuven partners, digital data will be archived on the KU Leuven K drive for storage of read-only data. At Inserm, the digital data are stored on the server until 2 years after the last publication in the active database, then the data generated by the project will be archived on magnetic tapes for 20 years before destruction.

What are the expected costs for data preservation during the expected retention period? How will these costs be covered?

For the KU Leuven partners, cost rate for storage on the K drive is 6.4€/year/100Gb. To store a total of 5 Tb for 10 years, the estimated cost hence is 3,200 €. Costs will be allocated to the project budget.
The long-term storage costs at Inserm are not yet determined, and they will be factored into the laboratory's day-to-day operational expenses. As a result, no specific charges have been assigned or applied to our project at this stage. The CarMeN laboratory is committed to managing and accommodating the costs associated with long-term storage as part of its regular operational budget. This ensures that our project can benefit from the necessary storage resources without incurring any separate or additional fees.

5. Data sharing and reuse

Will the data (or part of the data) be made available for reuse after/during the project? In the comment section please explain per dataset or data type which data will be made available.

- Yes, in an Open Access repository
- 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 regulations). All manuscripts will also be deposited in the KU Leuven Lirias digital repository. Raw RNA/ATACseq 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.

All KU Leuven-affiliated researchers involved in the project will have access to non-personal data on the KU Leuven servers through their u-number and accompanying personal password. Since this project involves a collaboration with a third partner outside of KU Leuven (Inserm, France), data that can be shared will be deposited on a Teams-based storage platform that can be made accessible to external users.
Personal data will only 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.

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 in the comment section per dataset or data type where appropriate.

- Yes, Intellectual Property Rights
- Yes, Privacy aspects
- Yes, Other

All KU Leuven-affiliated researchers involved in the project will have access to non-personal data on the KU Leuven servers through their u-number and accompanying personal password. Since this project involves a collaboration with a third partner outside of KU Leuven (Inserm, France), data that can be shared will be deposited on a Teams-based storage platform that can be made accessible to external users.
Since this project encompasses a collaboration with a non-KU Leuven partner (Inserm, France), agreements related to **exploitation and dissemination** have been written down in a collaboration agreement signed on March 22nd, 2023 by all parties involved in the project.
Since this project encompasses a collaboration with a non-KU Leuven partner (Inserm, France), agreements related to **IP rights and ownership** have been written down in a collaboration agreement signed on March 22nd, 2023 by all parties involved in the project.
Personal data will only 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.

☒ In an Open Access repository
☐ In a restricted access repository
☒ Upon request by mail
☒ Other (specify): Raw snRNA/ATACseq data will be made available publicly through the GEO repository upon acceptance of the manuscript.

When will the data be made available?

Data will be made available as publications at logical points during the course of the project when the research questions have been sufficiently addressed. 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.

All papers will be published in open access journals (according to FWO regulations) under a Attribution-NonCommercial-NoDerivatives 4.0 International (CC BY-NC-ND 4.0) licence.

Do you intend to add a PID/DOI/accession number to your dataset(s)? If already available, you have the option to provide it in the comment section.

- Yes

For raw snRNAs/ATACseq data, a GEO accession number will be provided.

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.

6. Responsibilities

Who will manage data documentation and metadata during the research project?

PhD students and technicians will have the daily responsibility of record keeping of all data (digital, paper and biological samples). They will also be responsible for a correct and accurate data entry and recording of metadata. The PI's (Aernout Luttun, Marion Delcroix and Frédéric Perros) and the clinical data manager (Rozenn Quarck) will be responsible for management of data documentation and metadata during the project.

Who will manage data storage and backup during the research project?

PhD students and technicians will have the daily responsibility of record keeping of all data (digital, paper and biological samples). They will also be responsible for a correct and accurate data entry and recording of metadata. The PI's (Aernout Luttun, Marion Delcroix and Frédéric Perros) and the clinical data manager (Rozenn Quarck) will be responsible for data storage and back-up during the project.

Who will manage data preservation and sharing?

The principal investigators (A. Luttun, F. Perros, R. Quarck through delegation by M. Delcroix).

Who will update and implement this DMP?

The principal investigators (A. Luttun, F. Perros, R. Quarck through delegation by M. Delcroix).

Unraveling the cellular landscape and (epi)genetic mechanisms at single-cell resolution and the effect of Prdm16 deficiency during the pathogenesis of pulmonary arterial hypertension
Application DMP

Questionnaire

Describe the datatypes (surveys, sequences, manuscripts, objects …) the research will collect and/or generate and /or (re)use. (use up to 700 characters)

During the course of this project different data types will be generated and stored on a dedicated server (see below), including microscopic/macroscopic images, morphometric analyses, mouse genotyping data, single-cell sequencing results. The results will be reported in open access manuscripts which will be deposited in Lirias, which is a document repository to archive the KU Leuven Association research output in a digital way. Lirias captures, stores, indexes, preserves, and distributes digital research material of all members of the KU Leuven Association. RNA/ATAC sequencing data will be made publically available through deposition to publicly accessible websites, i.e., the GEO database.

Specify in which way the following provisions are in place in order to preserve the data during and at least 5 years after the end of the research? Motivate your answer. (use up to 700 characters)

The person responsible for data storage/preservation will be the project' s coordinator (A. Luttun). Data will be stored from the starting date until 5 years after the ending date in a designated folder on 2 internal KU Leuven servers, the 'J-drive' (for data that needs to be accessible daily and is exchangeable between group members) or the ' L-drive (for longer-term storage of large data files), in which space is allocated to the Luttun group and which is automatically backed-up continuously. Only personnel involved in the project will have access to these folders. Costs related to data storage during and 5 years beyond the tenure of this grant are accounted for in the requested budget.

What's the reason why you wish to deviate from the principle of preservation of data and of the minimum preservation term of 5 years? (max. 700 characters)

NA.

Are there issues concerning research data indicated in the ethics questionnaire of this application form? Which specific security measures do those data require? (use up to 700 characters)

For this project, we will collect human samples and personal data. Handling of the latter requires specific security measures in order to comply with GDPR regulations. The data will be pseudonimized and stored on the internal KU Leuven servers mentioned above, which are compatible with GDPR regulations. Obtaining ethical approval from the UZ Leuven Ethics Committee also is dependent on compliance with GDPR rules (e.g. by adding a GDPR clause in the informed consent form).

Which other issues related to the data management are relevant to mention? (use up to 700 characters)

The co-promoter at the non-Flemish university (F. Perros, INSERM UMR_S 999, University of Paris-Saclay) will also put in place a Data Management Plan for data collected at this partner university.