LipoMacs: functional Lipidomics to unlock Macrophage-mediated therapy

A Data Management Plan created using DMPonline.be

Creators: Johan Swinnen, Jerome HENDRIKS, Max Mazzone, n.n. n.n., Charlotte Scott https://orcid.org/0000-0003-4914-6580

Affiliation: KU Leuven (KUL)

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

Template: FWO DMP (Flemish Standard DMP)

Principal Investigator: Johan Swinnen, Jerome HENDRIKS, Max Mazzone, n.n. n.n., Charlotte Scott https://orcid.org/0000-0003-4914-6580

Data Manager: Jonas Dehairs, Jaroslav Belotserkovsky, Geert Raes, Marcello Delfini

Project Administrator: Jaroslav Belotserkovsky

Grant number / **URL:** S001623N

ID: 191421

Start date: 01-10-2022

End date: 30-09-2026

Project abstract:

Macrophages are increasingly recognized as key regulators of common diseases such as neurodegeneration, cancer and liver disease. Central to macrophage functionality is their ability to attain different activation states, ranging from inflammation-inducing to inflammation-resolving and healing states. The precise phenotype and functions adopted are highly context dependent and have been shown to be driven by the local microenvironment in which the macrophages reside. Emerging evidence indicates that many pathological conditions are associated with alterations in macrophage phenotypes/functions due to a shift in their activation state, however whether this is a result of true macrophage plasticity or the recruitment of a phenotypically distinct macrophage population is also disease-specific. Crucially, recent studies have demonstrated a role for lipids and altered lipid metabolism in regulating the phenotypes and functions of disease-associated macrophages. In this project, we bring together unique complementary expertise from both academia and industry to design strategies for the functional reprogramming of macrophages by interfering with lipid metabolism. To this end, and instructed by the real-world needs of the industrial partners, we will set up a comprehensive technology pipeline involving state-of-the-art lipidomics approaches to map the heterogeneous landscape of macrophage phenotypes in terms of lipid metabolism in healthy and diseased tissue. Using gold-standard preclinical models, we will identify and validate key enzymes in lipid metabolism as potential targets and will provide proof of concept of pharmacological and/or nutraceutical approaches to regulate macrophage functions. This pipeline can be used for unbiased discovery but also offers multiple entry points to validate predefined targets in a hypothesis-driven manner.

Last modified: 31-03-2023

LipoMacs: functional Lipidomics to unlock Macrophage-mediated therapy 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)

Lab notebooks: hardcopy books and electronic notebooks
Lipidomics and MS Imaging data: .wiff, .raw, .TDF and contain instrument settings metadata (max 5 Tb for the entire project)
Transcriptomics and spatial transcriptomics data: .bcl, .fastq, .bam, .mtx, .csv (max 5 Tb) (including secondary data)
Microscopy pictures and gel scans: .tif/.tiff, .jpeg. (max 1 Tb)
Metabolomics data: .xls, .spss
Functional data: .xlsx or .pzfx
Flow cytometry data: .fcs, .wsp
Biological samples from human and mouse origin stored in biobanks
Manuscripts and illustrations: .doc/.docx, .pdf, .jpeg

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)

At each participating institution and research team a person responsible for data storage has been designated. For KU Leuven: J Swinnen, M Mazzone, J Dehairs. For VUB: J Van Ginderachter. For U Hasselt: J Hendriks; and for U Ghent: C Scott.
The different partner's institutes have sufficient (>5 Tb) and scalable storage capacity available, both during the research and during at least 5 years after the end of the research. Data storage and back-up is based on a combination of internal servers and external storage with commercial providers such as Microsoft and with public databases and repositories. Non-public data are login-protected and accessible to designated staff members only. After 5y, data will be transferred to semi-cold storage.

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)

As a general rule, data will be preserved for a minimum term of 5 years. Datasets collected in the context of clinical research, which fall under the scope of the Belgian Law of 7 May 2004, will be archived for 25 years, in agreement with the University Hospitals policies and the European Regulation 536/2014 on clinical trials of medicinal products for human use.

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)

All experiments on mice and on human clinical samples are conducted according to institutional, national and European regulations. Work with laboratory animals that is not yet included in ongoing approvals will be subjected to prior formal approval by the Ethical Committee on Animal Experimentation of the respective institutes before the start of the respective work packages.
Work on human biological material will be subjected to approval by the respective Ethics Committee.

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

We foresee that the proposed work wlll result in research data with potential for tech transfer and valorization. The partner institutes have a policy to actively monitor research data for such potential. If there is substantial potential, the invention will be thoroughly assessed, and in a number of cases the invention will be IP protected (mostly patent protection or copyright protection). As such the IP protection does not withhold the research data from being made public. In the case a decision is taken to file a patent application it will be planned so that any delay to publications is minimal.

LipoMacs: functional Lipidomics to unlock Macrophage-mediated therapy
DPIA

DPIA

Have you performed a DPIA for the personal data processing activities for this project?

- Not applicable

LipoMacs: functional Lipidomics to unlock Macrophage-mediated therapy
GDPR

GDPR

Have you registered personal data processing activities for this project?

- Not applicable

LipoMacs: functional Lipidomics to unlock Macrophage-mediated therapy
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.

				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
		Please choose from the following options: • Generate new data • Reuse existing data	*Please choose from the following options:* • Digital • Physical	*Please choose from the following options:* • Observational • Experimental • Compiled/aggregated data • Simulation data • Software • Other • NA	*Please choose from the following options:* • .por, .xml, .tab, .cvs,.pdf, .txt, .rtf, .dwg, .gml, … • NA	*Please choose from the following options:* • <100MB • <1GB • <100GB • <1TB • <5TB • <10TB • <50TB • >50TB • NA	
KUL-JS-Bulk lipidomics	Bulk lipidomics data	Generate new data	Digital	Experimental	.xlsx .wiff	<100GB	
KUL-JS-MSImaging	MS Imaging data (TimsTOF, DESI-MRT)	Generate new data	Digital	Experimental	.imzML	<10TB	
KUL-JS-microscopy	Raman/fluorescence/visible light microscopy data	Generate new data	Digital	Experimental	.tiff	<10TB	
KUL-JS-Functional cell culture	Incucyte data	Generate new data	Digital	Experimental	.tiff	<100GB	
KUL-JS-Tumor growth	Caliper and toxicity data of mouse experiments	Generate new data	Digital	Experimental	.xlsx	<100GB	
KUL-MM-Flow cytometry	Flow cytometry data	Generate new data	Digital	Experimental	.fcs	<1TB	
KUL-MM-Functional assay	Phagocytosis, cytotoxicity, Tcell suppression	Generate new data	Digital	Experimental	.xlsx .pfz .fcs	<1GB	
KUL-MM-Metabolomic	Metabolomic data	Generate new data	Digital	Experimental	.raw .pmd	<10GB	
KUL-MM-Tumour growth	tumour growth curves	Generate new data	Digital	Experimental	.xlsx .pfz	<1GB	
VUB-JVG-tumor growth	Tumor growth data	Generate new data	Digital	Experimental			
VUB-JVG-transcriptomics	Transcriptomics data	Generate new data	Digital	Experimental			
VUB-JVG-flow cytometry	Flow cytometry data	Generate new data	Digital	Experimental			
VUB-JVG-functional assays	Phagocytosis, cytotoxicity, Tcell suppression	Generate new data	Digital	Experimental			
UH-JH-Flow cytometry	Flow cytometry	Generate new data	Digital	Experimental	.fcs	<1TB	
UH-JH-microscopy	Zeiss LSM880, Leica	Generate new data	Digital	Experimental	.czi ;Jpeg	<1GB	
UH-JH-Functional cell culture	Functional cell culture	Generate new data	Digital	Experimental	.xlsx .fcs	<1GB	
UH-JH-in vivo data	EAE scores, mice weights, other data of mouse experiments	Generate new data	Digital	Experimental	.xlsx	<1GB	
UG-CS-Flow cytometry	Flow cytometry	Generate new data	Digital	Experimental	.fcs	<1TB	
UG-CS-microscopy	Zeiss LSM780,880 Leica	Generate new data	Digital	Experimental	.czi ;Jpeg	<1GB	
UG-CS-Spatial transcriptomics/proteomics	Milltenyi MACSima, Resolve Molecular Cartography, Vizgen MerFISH, !0X Visium HD	Generate new data	Digital	Experimental	.tar.gz; .fastq.gz; Jpeg; mtx.gz; bam.bai	<1TB	
UG-CS-transcriptomics	Transcriptomics data	Generate new data	Digital	Experimental	.fastq.gz; .tar.gz	<1TB	
UG-CS-Functional assays	Phagocytosis, cytotoxicity, Tcell suppression	Generate new data	Digital	Experimental	.xlsx .pfz .fcs	<1GB	

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:

KU Leuven-JS: Not applicable
KU Leuven-MM: Not applicable
UHasselt-JH: Not applicable
VUB-JVG: not applicable
UGent-CS: most data will be generated in this project, but also possibility to reuse data from www.livercellatlas.org

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 approvals are required for animal experimentation and the use of human material.
Approvals are in place in part at the start of the project and/or will a obtained from the local ethical committees by each partner before the start of the experiments.
The following approvals are in place:

- KU Leuven: S-62275; S-61508; S-65339; P-086-2021; P012/2022
- VUB: 22-220-20
- U Gent: EC2021-016; EC2021-107 and EC2022-065
- U Hasselt: 201904BA1; 201904B; 201904A1; 202001BA1

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.

- No

We will process biological samples from patients. All samples and associated clinical information are anonymized and are provided by a code that cannot be traced back to an individual by the researchers.

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

We foresee potential for valorization in at least 4 areas:

- Lipidomics-related analytical and data processing technologies
- Assays and models for target identification and validation
- Targets and inhibitors for the development of therapeutics
- Nutraceutical approaches targeting lipid metabolism

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.

- No

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

Certain models are obtained under MTA restricting their use and distribution.

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).

KU Leuven-JS: The data will be documented in lab notebooks, recorded data will include its source, format, structure, and any other relevant metadata. Initially, physical lab books will be used and this will be transitioned to electronic notebooks (which one is to be decided and depends on KU Leuven initiatives)
KU Leuven-MM: Protocols and details related to data collection and processing will be recorded in physical lab books and transcribed to Word or Excel files by the applicants. Long-term storage of the lab books is supervised by the lab manager. The lab book describes in detail the experimental setting and any deviation from the original experimental design. Experiments are organized by date and therefore fully searchable. Data folders containing raw and processed data will be hierarchically organized and labeled based on the source of the data, the type of experiment, the date of data generation, and the different experimental conditions analyzed. Data analysis methods and particularities (including metadata) will be described in text documents and Excel files included in these folders. Standard operating procedures (SOP) are constantly updated and safely stored as PDF to ensure proper replication of the biological experiments. SOPs are backed up to the J-drive once per month, where they are available for all members of the lab. Similarly, all the experimental data (raw and processed data), are safely stored in dedicated drives.
VUB-JVG: Data will be generated following standardized protocols, that are stored in a central CMIM Teams Sharepoint folder. Documentation and metadata linked to each experiment will be documented by the technical and research staff, via electronic files in dedicated Teams Sharepoint folders that refer to specific data sets and in hard copy lab notebooks.
UH-JH-: An electronic lab notebook (ELN FTW) will be used to share data. Metadata is used to provide information about the generated data. Information is documented both on individual datasets and on how the datasets are (inter)linked. Each dataset will contain a clear overview of all information needed to understand and reuse the data collected/generated in this project (e.g. the protocol that was used, information on used abbreviations, calculations, …).
UG-CS: Documentation and metadata linked to each experiment w ill be documented by the technical and research staff in hard copy lab notebooks and in electronic lab notebooks (ELN) in this project. This includes the research design, protocol, context of data collection, data collection methods, quality control procedures, processing and analysis procedures. Cryotubes of biological samples stored at -80°C or -150°C w ill be labelled with a reference number that links to an entry in its corresponding database.

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.

- No

KU Leuven MM: Text documents and Excel files stored within each experiment folder will respectively contain guidelines describing data collection/analysis methods and all relevant metadata (including experimental conditions, sample keys, computational analysis pipelines and their parameters) to ensure the reusability of the data and the reproducibility of any further data generation.

Experiments are organized by date so they can be compared with lab books for additional information. References of reagents used are listed on an excel table together with the reference and the dilution used.

VUB-JVG: Metadata will include the following elements: Title (free text); Creator (Last name, first name, organization); Date and time reference; Subject (Choice of keywords and classifications); Description (Text explaining the content of the data set and other contextual information needed for the correct interpretation of the data, the software(s) used to produce and to read the data, the purpose of the experiment, etc); Resource Type (data set, image, audio, etc); Identifier (DOI; when applicable)

UHasselt-JH: Experiment metadata will be organized using the following elements: Experiment number, Title, Date (YYYYMMDD), Creator abbreviation (First 3 letters last name, first 2 letters first name), description. References of reagents used are listed in the electronic lab books. All data files including text documents, Excel files and raw data files will be stored within one experiment folder and attached to the electronic lab book.

UGent-CS: Metadata will include the following elements: Title (free text); Creator (Last name, first name, organization); Date and time reference; Subject (Choice of keywords and classifications); Description (Text explaining the content of the data set and other contextual information needed for the correct interpretation of the data, the software(s) used to produce and to read the data, the purpose of the experiment, etc); Resource Type (data set, image, audio, etc); Identifier (DOI; when applicable)

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

Where will the data be stored?

KU Leuven-JS: Data will be stored on personal laptops with automated backup on KU Leuven servers. Omics data are stored on a local 90Tb Synology NAS.

KU Leuven-MM: Data will be backup on KU Leuven servers. (1) J-Drive (Minimum 100Gb, Expandable, Unlimited Size, Stored data can be modified); (2) L-Drive (Minimum 5Tb, Unlimited Size, Stored data can be modified); (3) K- Drive (Archive storage, Minimum 100Gb, Expandable, Unlimited Size, Stored data can NOT be modified or deleted).

VUB-JVG: Experiment details are recorded in lab notebooks. Electronic records (raw and analyzed data) are stored on Sharepoint and on encrypted external hard drives. All personal digital research data are stored in personal folders on the VUB OneDrive, with access by the team leader(s). Centralized data that need to be accessible for multiple team members are stored on a common Teams/Sharepoint site. Omics data generated during the project will either be stored on VUB servers or on The Flemish Supercomputer Centre (VSC), initially in the staging area and later in the archive area.

UH-JH:All digital data (e.g. Excel files, Graphpad files, …) will be stored in a password protected Google Drive provided by the institution. All samples (e.g. tissue, RNA, cDNA, …) will be stored in access-restricted fridges and/or freezers.

UG-CS: All research data are stored in personal folders on the S drive on the servers of the VIB-UGent Center for inflammation Research (IRC) according to our data policy. Raw data obtained from flow cytometry and microscopy are stored on separate drives on the IRC servers. Data are stored three times on two different locations.

How will the data be backed up?

KU Leuven-JS: Data will be backed-up on external hard drives, on KU Leuven servers and at the central data storage at LISCO. Mirroring of the Synology NAS at Biomed and an iRODS pilot are planned. The NAS will be configured in RAID5, meaning should one of the hard drives fail it can be replaced without the loss of data.

KU Leuven-MM: Data storage and backup is managed by KU Leuven IT services. Specifically, mirror copies of the stored data are made immediately upon upload, for safety backup purposes. Data storage and backup is based on a combination of internal servers and external storage with commercial providers such as Microsoft, and with public databases and repositories. Long- term storage, is ensured by the L-drive (stored data can NOT be modified or deleted. This, will guarantee the preservation of data over the minimum term of 10 years. Regular computer backups through the inSync platform (unlimited storage) secure the recovery of unsaved data.

VUB-JVG: The back-up of the data present on VUB Teams/SharePoint/OneDrive is performed automatically via Microsoft.

UH-JH: At Hasselt University, digital data is backed up using the automatic back-up service and version control system of Google Drive Enterprise with a guaranteed back-up time period of 6 month, hosted by Hasselt University (data will be available 30 days after removal). Additionally, data will be backed up regularly using a hard drive of 1 terabyte, which will be password protected. Physical data cannot be backed up but they can be repeated at any time using the protocols and processes described in the associated (meta)data.

UG-CS: Data are backed up on two different locations, one in the main building of the host institute (FSVM I building) and one in the iGent building on the Tech Lane Ghent site. The frequency of backups depends on the type of data: research data every 24 hours, emails: every 8 hours and databases every 24 hours.

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

KU Leuven-JS: sufficient data storage is available at the local internal and external hard drives and the Synology NAS (90Tb).

KU Leuven-MM: The institutes (KU Leuven and VIB) have sufficient and scalable storage capacity available, both during the research and during at least 10 years after the end of the research.

VUB-JVG: There is sufficient storage and back-up capacity at VUB. Storage and backup capacity is scalable depending on the need of the project.

UH-JH: Sufficient data storage is available. UHasselt provides and manages password protected Google suite packages which allow the storage of more than 1 terabyte per group drive. Furthermore, all facilities necessary for proper treatment and safe storage of data during and after completion of the project are available through the infrastructure at BIOMED.

UG-CS: Yes. Storage is scalable depending on the need of project (similar for backup capacity). To prevent bad storage policy, personal and team folders are subjected to quota.

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

KU Leuven-JS: Access to laptops and data storage facilities is secured by personal passwords.

KU Leuven-MM: Non-public data are login-protected (2FA) and accessible to designated staff members only. Network security is ensured by the KULeuven IT services.

VUB-JVG: Physical access to the lab is controlled by electronic access cards. Access to the VUB OneDrive folders containing the data for this project is personal, but access is granted to the other lab-members working on the project. Access to the CMIM Sharepoint/Teams folder is set by the Administrators of the Teams site (the PIs and responsible post-docs).

UH-JH: Concerning digital data, Hasselt University's Google Drive Enterprise data is encrypted when it is on a disk, stored on backup media, moving over the Internet, or travelling between data centres. Encryption is an important piece of the Google drive security strategy, helping to protect emails, chats, Google Drive files, and other data. Specific access needs to be granted before data is visible for others. Concerning physical data, all samples (e.g. tissue, RNA, cDNA, …) will be stored in access-restricted fridges and/or freezers.

UG-CS: The following measurements are taking to ensure secure data storage and to prevent modification by unauthorized persons: controlled physical access to the building, firewalling (on both departmental and individual server levels), switched network & encrypted communications (prevents eavesdropping), VLAN-ing (network compartimentalisation) & MAC authentication, least-know n ports for well-known services (security through obscurity), brute-force intrusion detection & isolation, individual account expiry in accordance with contract of employment, ACL's (Access Control Lists; application of Principle of Least Privilege).

What are the expected costs for data storage and backup during the research project? How will these costs be covered?

KU Leuven-JS: The costs are estimated at 10k euro. This is covered by internal grants (LISCO data core) and fee-for-service revenue.

Ku Leuven-MM: The annual cost of storage is approximately 569.2 € per 5TB of storage space per year. This cost and capacity include the performance of mirror copies of the stored data, for safety backup purposes. These costs will be covered by the budget of the project leaders.

VUB-JVG: The Sharepoint/OneDrive license is covered by VUB, without any specific costs for the research groups. The total storage cost is estimated to be at 600€/TB and 50€/TB backup.

UH-JH: Concerning digital data, costs for Google Drive enterprise data and back up are covered by UHasselt. Costs for storage of physical data will be covered by the Biomedical research institute (BIOMED).

UG-CS: Digital data storage costs are included in general lab costs. The total storage cost is estimated to be at 600€/TB and 50€/TB backup

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...).

KU Leuven-JS: All lipidomics and spatial lipidomics data will stored at the LISCO data storage facility for at least 5 years.

KU Leuven-MM: Biologicals will be retained when possible. The storage will be at (1) room temperature, (2) 4°C, (3) -20°C, (4) -80°C, (5) -150°C, according to the type of biological material. As a general rule, data will be preserved for a minimum term of 10 years. Datasets collected in the context of clinical research, which fall under the scope of the Belgian Law of 7 May 2004, will be archived for 25 years, in agreement with the University Hospital policies and the European Regulation 536/2014 on clinical trials of medicinal products for human use.

VUB-JVG: The minimum preservation term of 5 years after the end of the project will be applied to all datasets.

UH-JH: All the obtained data on our drive is accessible to Prof. dr. Hendriks, responsible for its preservation and any necessary transfer after the end of the research. All digital data (e.g. Excel files, Graphad files, …) will be retained for the expected 5-year period after the end of the project. All biological samples that can be stored in access-restricted freezers and fridges (e.g. RNA, tissue, …) will also be retained for the expected 5-year period after the end of the project.

UG-CS: The minimum preservation term of 5 years after the end of the project will be applied to all datasets.

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

KU Leuven-JS: Long-term storage is foreseen at the LISCO facilities.

KU Leuven-MM: Physical space in the lab will be provided to store the biological material (refrigerators, cold rooms as well as cupboards for non-degradable materials). Digital data will be stored on KU Leuven storage space: the KU Leuven L-Drive (virtually unlimited size) has sufficient storage capacity for the outlined project.

VUB-JVG: As a general rule, datasets will be made openly accessible, whenever possible via existing platforms that support FAIR data sharing (www.fairsharing.org), at the latest at the time of publication.
For all other datasets, long term storage will be ensured as follows:
-Digital datasets: files will be stored on Sharepoint.
-Tissue samples: Tissues will be stored locally in the laboratory.
-Omics data generated during the project will either be stored on VUB servers or on The Flemish Supercomputer Centre (VSC), initially in the staging area and later in the archive area
-Animal cell lines are stored in liquid Nitrogen
-Other biological and chemical samples: storage at 4°C and/or as frozen samples in cryovials as appropriate.
- Hard copy notebooks will be archived in the host institute's building.

UH-JH: All digital data (e.g. Excel files, Graphpad files, …) will be stored in a password protected Google Drive Enterprise, hosted by Hasselt University. All samples (e.g. tissue, RNA, cDNA, …) will be stored in access-restricted fridges and/or freezers. Experimental notes will be digitalized in electronic lab books.

UG-CS: Hard copy notebooks w ill be archived in the host institute's building. Digital data will be archived in team folders on the T-drive (IRC server). In addition, data w ill be stored and backed up offline on hard disks on two different locations.

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

KU Leuven-JS: These costs are included in the estimated 10k euro for storage and backup. This is covered by internal grants (LISCO data core) and fee-for-service revenue.

KU Leuven-MM: The annual cost of storage is approximately 569.2 € per 5TB of storage space per year. This cost and capacity include the performance of mirror copies of the stored data, for safety backup purposes. These costs will be covered by the budget of the project leaders.

VUB-JVG: Data storage and backup costs via Sharepoint/OneDrive are covered by VUB. The total storage cost, including backups, is estimated to be at 0,66€/GB. Offline storage is estimated at 0,10€/GB.

UH-JH: Concerning digital data, costs for Google Drive enterprise data and back up are covered by UHasselt. Costs for storage of physical data will be covered by the Biomedical research institute (BIOMED).

UG-CS: Digital data storage costs are included in general lab costs. The total storage cost, including backups, is estimated to be at 0,66€/GB. Offline storage is estimated at 0,10€/GB.

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
- Yes, in a restricted access repository (after approval, institutional access only, …)

KU Leuven-JS: selected data will be deposited in Open Access repositories and in the KU Leuven restricted access institutional repository RDR.

KU Leuven-MM & UH-JH: The key findings of the project and their interpretation will be made available through publication of journal articles in established, peer-reviewed (non-predatory) academic journals. Relevant raw data will be made publicly available through upload to well-established open-access data repositories.

VUB-JVG: Key data and findings will be made available upon publication of the results. Bioinformatics data and software scripts are put on public repositories, such as GEO and Github. Biological material will be distributed under MTA to other parties if requested.

UG-CS: All data, with the exception of personal/otherwise confidential data, are available upon publication of the results. Manuscripts will be published in Open Access journals or made Green Open Access using the repository UGent
biblio. Bioinformatics data and software scripts are put on public repositories, such as GEO and Github as well as on www.livercellatlas.org. Biological material will be distributed under MTA to other parties if requested.

If access is restricted, please specify who will be able to access the data and under what conditions.

KU Leuven-JS: Only the PI and the data manager (Jonas Dehairs) will have access to the data.

KU Leuven-MM: Data may be shared externally upon reasonable requests from collaborating scientists, which will be reviewed and approved on a case-by-case basis by the project leaders. We do not exclude that the proposed work could result in research data with potential for tech transfer and valorization. The partner institutes have a policy to actively monitor research data for such potential. If there is substantial potential, the invention will be thoroughly assessed, and in a number of cases the invention will be IP protected (mostly patent protection or copyright protection). As such the IP protection does not withhold the research data from being made public. In the case a decision is taken to file a patent application it will be planned so that any delay to publications is minimal.

VUB-JVG: Published data are accessible to all. Unpublished data that are considered confidential will only be accessible to the PI and the people within his research team. For data shared directly by the PI, a material transfer agreement (and a non-disclosure agreement if applicable) will be concluded with the beneficiaries in order to clearly describe the types of reuse that are permitted.

UH-JH: Data may be shared exeternally upon reasonable requests from collaborating scientists. Published data are accessible to all. Unpublished data that are considered confidential will only be accessible to the PI and the reseach team. Unpublished research data will be actively monitored for tech transfer and/or valorization, and may be IP protected (patent or copyright protection). In the case a decision is taken to file a patent application it will be planned so that any delay to publications is minimal.

UG-CS: As a general rule all research outputs will be made openly accessible at the latest at the time of publication. No embargo will be foreseen unless imposed e.g. by pending publications, potential IP requirements – note that patent application filing will be planned so that publications need not be delayed - or ongoing projects requiring confidential

data. In those cases, datasets will be made publicly available as soon as the embargo date is reached.

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

Certain models are obtained under MTA restricting their use and distribution.
All patient samples and associated clinical information are anonymized and are provided by a code that cannot be traced back to an individual by the researchers.
We do not exclude that the proposed work could result in research data with potential for tech transfer and valorization. If there is substantial potential, the invention will be thoroughly assessed, and in a number of cases the invention will be IP protected (mostly patent protection or copyright protection). As such the IP protection does not withhold the research data from being made public. In the case a decision is taken to file a patent application it will be planned so that any delay to publications is minimal.

Where will the data be made available? If already known, please provide a repository per dataset or data type.

Depending on the nature of the dataset both open-access and restricted-access repositories will be used to store data. All requests and approvals for reuse of data other than those deposited in open-access repositories will be assessed on a case-by-case basis by the project leaders.
Lipidomics data: Metabolomics Workbench
Transcriptomics data (UH): NCBI Gene Expression Omnibus (GEO) database

When will the data be made available?

Study leaders will monitor data sharing requests. Data will be automatically made available after the publication of results to any requestor using the data for non-commercial purposes. Commercial use of the data will be negotiated through the VIB tech transfer office.

Which data usage licenses are you going to provide? If none, please explain why.

Lipidomics data: Metabolomics Workbench states the data will be in the Public Domain Mark (PD), and will be free of known restrictions under copyright law, including all related and neighboring rights.

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

When relevant PID/DOI/accession numbers will be added to datasets

What are the expected costs for data sharing? How will these costs be covered?

Lipidomics data: Metabolomics Workbench does not charge any costs.
Mainly publication fees. Covered with the FWO budget. Other costs for data sharing will be discussed with collaborators on a case-by-case basis.

6. Responsibilities

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

KU Leuven:JS: Jonas Dehairs. KU Leuven-MM: Data documentation, and metadata acquisition and storage will be performed by the students and postdocs associated with this project. Prof. Massimiliano Mazzone and Dana Liu (Lab Manager) will monitor.; VUB-JVG & UG-CS: The individual PhD students and postdocs will be responsible for the data documentation and metadata. Jo Van Ginderachter/Charlotte Scott will perform regular checks. UH-JH: Data documentation, and metadata acquisition and storage will be performed by the PhD students and postdocs associated with this project, and will be checked regularly by Jerome Hendriks.

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

KU Leuven:JS: Jonas Dehairs. KU Leuven-MM: Data management, storage and back up will be performed by the students and postdocs associated with this project. Prof. Massimiliano Mazzone and Dana Liu will monitor; VUB-JVG &UG-CS: The research and technical staff will ensure data storage and back up, with support from the VUB Research Information and Data Management Office or VIB-IRC (UG-CS). UH-JH Data management, storage and back up will be performed by the students and postdocs associated with this project with regular checks by Jerome Hendriks

Who will manage data preservation and sharing?

KU Leuven:JS: Jonas Dehairs. KU Leuven-MM: Prof. Massimiliano Mazzone will be responsible for data preservation. He will also monitor data sharing requests; VUB-JVG & UG-CS: The PIs are responsible for digital preservation and sharing, with support from the VUB Research Information and Data Management Office or VIB-IRC (UG-CS). UH-JH: Jerome Hendriks will be responsible for the data preservation and sharing with support from the biomedical research institute (BIOMED)

Who will update and implement this DMP?

KU Leuven:JS: Jonas Dehairs and Johan Swinnen KU Leuven-MM: Dana Liu Marcello Delfini and Massimiliano Mazzone; VUB-JVG: Jo Van Ginderachter; UH-JH: Jerome Hendriks and Sanne Verberk; UG-CS: Charlotte Scott