Investigating the role of DNA methylation in the failing and in the ageing heart

A Data Management Plan created using DMPonline.be

Creators: Onne Ronda, First Name Surname

Affiliation: KU Leuven (KUL)

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

Template: FWO DMP (Flemish Standard DMP)

Principal Investigator: First Name Surname

Data Manager: First Name Surname

Grant number / URL: 12D6523N

ID: 196954

Start date: 01-10-2022

End date: 30-09-2025

Project abstract:

Ageing is the greatest risk factor for heart failure. Heart muscle cells (cardiomyocytes, CMs) experience ageing-associated alterations leading to a lower functional output, caused by (among others) maladaptive remodelling of the myocardium, including hypertrophy of individual cardiomyocytes. CMs are long-lived (decades) and have limited proliferative capacity. Transcriptional variation (heterogeneity) between CMs increase with age, with potentially detrimental consequences for heart function. Not only does this reflect dysregulation of gene transcriptional control in individual cells, it also likely reflects changes in CM subpopulations/states, such as the ratio of maladapted hypertrophic- to normal CMs. Transcription is regulated in part by DNA methylation (DNAm) (an epigenetic modification). Curiously, global DNAm decreases with age. Our data show that the expression of the TET2 DNA demethylase increases with age, with profound inter-CM heterogeneity. We hypothesize that changes in DNAm patterns induce transcriptional remodelling underlying dysfunction during normal ageing of CMs. We further propose that the increased variability in transcriptomes between CMs in ageing is mediated by stochastic alterations in DNAm brought about by increased TET2 expression.

Last modified: 27-02-2023

Investigating the role of DNA methylation in the failing and in the ageing heart
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)

WP 1: Murine pressure-overload induced heart failure in CM-specific Tet2 KO

New physical, -80°C: Collected tissues in tubes (heart, lung, liver), and blood plasma. Collected heart tissue in cryomolds. Total expected size: 20 cryoboxes.

New digital: echocardiography (.bimg .pimg .mxml .vxml 50 gb), FACS (.fcs 10 gb), immunofluorescence (.nd2 100 gb), mass spectrometry (.csv 100 mb), genotyping (.scn .xlsx 100 mb), SPSS and R analysis (.sav .r 100 mb), manuscript and figures (.docx .xlsx .pzfx .tif .pdf 100 mb). Total expected size: 0.2 TB

WP 2: Human heart array-based methylomics (no personal data)

New physical, -80°C: Isolated DNA and RNA from human myocardial tissue in tubes. Total expected size: 10 cryoboxes.

New digital: FACS (.fcs 10 gb), immunofluorescence (.nd2 100 gb), gene expression qPCR (100 mb), mass spectrometry (.csv 100 mb), Illumina Infinium HumanMethylation850 BeadChip (10 gb), SPSS and R (.sav .r 100 mb), manuscript and figures (.docx .xlsx .pzfx .tif .pdf 100 mb). Total expected size 0.2 TB.

WP 3: Murine transcriptomics and methylomics

Reuse physical: From WP 1

New digital: FACS (.fcs 10 gb), immunofluorescence / RNA-scope (.nd2 100 gb), snNMT-sequencing (.fastq 1.5 tb), R analysis (.r 100 mb). Total expected size: 2 TB.

WP 4: Murine cardiac ageing in cardiac-specific Tet2 absence

New physical -80°C: Collected tissues in tubes (heart, lung, liver), and blood plasma. Collected heart tissue in cryomolds. Total expecte size: 25 cryoboxes.

New digital: echocardiography (.bimg .pimg .mxml .vxml 100 gb), FACS (.fcs 50 gb), immunofluorescence (.nd2 100 gb), snNMT-sequencing (.fastq 2 tb), mass spectrometry (.csv 100 mb), genotyping (.scn .xlsx 100 mb), SPSS and R analysis (.sav .r 100 mb), manuscript and figures (.docx .xlsx .pzfx .tif .pdf 100 mb). Total expected size: 2.5 TB.

Overall (whole-project) expected digital size: 5 TB.

Overall (whole-project) expected physical -80°C size: 45 cryoboxes

Overall (whole-project) expected physical size: 15 bound lab books.

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)

1. **Designation of responsible person:** Roxane Menten (lab manager), Hywel Llewelyn Roderick (PI)
2. Storage capacity/repository
    - **during the research**: Daily experimental data are written down into a bound physical lab book, each page paginated and dated, in chronological order. Physical samples are stored in the lab's -80°C freezer. Digital data are stored on the lab's L drive (GBW-0050_Llew_Cardio) and raw data are transferred onto the lab's K drive (GBW-0088_Exp_Cardio_Archive) when appropriate.
    - **after the research**: Lab books are stored in locked storage cabinets owned by the department. Biological samples are kept inside -80°C freezers for at least 5 years after the end of the research. Digital data are kept on the K and L drives for at least 5 years after the end of the research. Bulk raw data (>1 TB sequencing data) will be transferred to 2 separate external harddrives. Samples and data may be kept for longer, in case they can be reused by other researchers within the department.

Derivatives of human samples (e.g. isolated RNA, DNA) will be deposited into the biobank. The
necessary storage room for the expected amount of biological material is already ascertained at the biobank.

All digital data are stored on the lab's L and K drive. Only members of the lab who were granted permission to access these data can access these data; e.g. students are not granted access to data which is not relevant to their training.

Upon manuscript acceptance, raw sequencing data are deposited onto NCBI Sequence Read Archive (SRA) and publicly accessible. Scripts (R) are deposited onto github and publicly accessible.

During the research, proper storage of data is the responsibility of dr. Ronda. After the research, this responsibility shifts to prof. Roderick. Human samples and derivatives are the responsibility of Ms Menten.

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)

Derivatives of human samples (e.g. isolated RNA, DNA) are deposited in the biobank stored at -80°C, with an expected storage time beyond 5 years duration.

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)

Yes, human subject data are being used and the personal data of healthy volunteers will be stored. This includes 'regular' personal data such as age, weight, sex, but also 'special' or 'sensitive' personal data such as health data. To dr. Ronda, these data are anonimized, with Ms. Menten serving as the pseudo-anonimizer key holder. For these human samples, approval from the ethical committee is mandatory. This has already been granted: S58824.

A follow-up ethical approval for future human sample collection is currently pending and is expected to be granted H2 2023. This follow-up approval is not required for the continued use of the samples already collected under S58824.

Animal experiments are being performed, for which approval from the ethical committee is mandatory. This has already been granted: P182/2021 - Tet2 in hartspierverdikking en hartpathologie tijdens ziekte en veroudering.

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

None.

Investigating the role of DNA methylation in the failing and in the ageing heart
DPIA

DPIA

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

- Not applicable

Investigating the role of DNA methylation in the failing and in the ageing heart
GDPR

GDPR

Have you registered personal data processing activities for this project?

- Not applicable

Investigating the role of DNA methylation in the failing and in the ageing heart
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
WP1 Murine TET2 AB - physical samples	Collected tissues in tubes: heart, lung, liver, blood plasma. Collected heart tissue in cryomolds.	**Generate new data**	**Physical**	**Experimental**			20 cryoboxes stored at -80°C
WP1 Murine TET2 AB - digital data, echocardiography	In vivo echocardiography measurements	**Generate new data**	**Digital**	**Experimental**	.bimg .pimg .mxml .vxml	10 gb	
WP1 Murine TET2 AB - digital data, FACS	Flow cytometry sorting of cardiomyocyte nuclei	**Generate new data**	**Digital**	**Experimental**	.fcs	10 gb	
WP1 Murine TET2 AB - digital data, immunofluorescence	Confocal microscopy images	**Generate new data**	**Digital**	**Experimental**	.nd2	100 gb	
WP1 Murine TET2 AB - mass spectrometry	Mass-spec quantification of 5-mC	**Generate new data**	**Digital**	**Experimental**	.csv	100 mb	
WP1 Murine TET2 AB - genotyping	Determine mouse genotypes	**Generate new data**	**Digital**	**Experimental**	.scn .xlsx	100 mb	
WP1 Murine TET2 AB - SPSS and R analysis	Bio-informatics and statistical analysis of data	**Generate new data**	**Digital**	**Software**	.r .sav	100 mb	
WP2 human heart ageing - physical samples	Generate derivatives (isolated RNA, DNA) from existing human samples present in the biobank	**Generate new data**	**Physical**	**Experimental**			10 cryoboxes stored at -80°C
WP2 human heart ageing - digital data, FACS	Flow cytometry sorting of cardiomyocyte nuclei	**Generate new data**	**Digital**	**Experimental**	.fcs	10 gb	
WP2 human heart ageing - digital data, immunofluorescence	Confocal microscopy images	**Generate new data**	**Digital**	**Experimental**	.nd2	100 gb	
WP2 human heart ageing - digital data, methylation array	Illumina Infinium HumanMethylation850 BeadChip for array-based analysis of DNA methylation status	**Generate new data**	**Digital**	**Experimental**	.csv	10 gb	
WP2 human heart ageing - digital data, SPSS and R	Bio-informatics and statistical analysis of data	**Generate new data**	**Digital**	**Software**	.r .sav	100 mb	
WP3 Murine transcriptomics and methylomics - digital data, FACS	Flow cytometry sorting of cardiomyocyte nuclei for the purpose of plate-based sequencing	**Generate new data**	**Digital**	**Experimental**	.fcs	10 gb	
WP3 Murine transcriptomics and methylomics - digital data, immunofluorescence	Confocal microscopy images, RNA-scope	**Generate new data**	**Digital**	**Experimental**	.nd2	100 gb	
WP3 Murine transcriptomics and methylomics - digital data, sequencing	Multi-omics analysis of biological samples obtained in WP1	**Generate new data**	**Digital**	**Experimental**	.fastq	1.5 tb	
WP3 Murine transcriptomics and methylomics - digital data, R analysis	Bio-informatics and statistical analysis of data	**Generate new data**	**Digital**	**Software**	.r .sav	100 mb	
WP4 Murine ageing - physical samples	Collected tissues in tubes: heart, lung, liver, blood plasma. Collected heart tissue in cryomolds.	**Generate new data**	**Physical**	**Experimental**			25 cryoboxes stored at -80°C
WP4 Murine ageing - digital data, echocardiography	In vivo echocardiography measurements	**Generate new data**	**Digital**	**Experimental**	.bimg .pimg .mxml .vxml	100 gb	
WP4 Murine ageing - digital data, FACS	Flow cytometry sorting of cardiomyocyte nuclei, for the purpose of bulk analysis and sequencing	**Generate new data**	**Digital**	**Experimental**	.fcs	50 gb	
WP4 Murine ageing - digital data, immunofluorescence	Confocal microscopy images	**Generate new data**	**Digital**	**Experimental**	.nd2	100 gb	
WP4 Murine ageing - digital data, sequencing	Multi-omics analysis of biological samples	**Generate new data**	**Digital**	**Experimental**	.fastq	2 tb	

WP4 Murine ageing - digital data, mass spectrometry	Mass-spec quantification of 5-mC	**Generate new data**	**Digital**	**Experimental**	.csv	100 mb	
WP4 Murine ageing - digital data, genotyping	Determine mouse genotypes	**Generate new data**	**Digital**	**Experimental**	.scn .xlsx	100 mb	
WP4 Murine ageing - digital data, SPSS and R analysis	Bio-informatics and statistical analysis of data	**Generate new data**	**Digital**	**Software**	.r .sav	100 mb	
Lab books	Day-to-day reporting of procedures, data, observations	**Generate new data**	**Physical**	**Experimental**		15 lab books	
PCR primers	Designed and ordered primers for PCR and qPCR		**Physical & Digital**	**Experimental**	.xlsx	100 mb	4 cryoboxes stored at -20°C
Standard operating procedures and (study) protocols	Optimized procedures described and documented in protocols		**Digital**		.docx	100 mb	

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:

WP2 reuses human samples generated under ethical agreement S58824. No human samples are collected specifically under this DMP's study.

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

Human: This applies to WP2. Data are being used and the personal data of healthy volunteers will be stored. This includes 'regular' personal data such as age, weight, sex, but also 'special' or 'sensitive' personal data such as health data. To dr. Ronda, these data are anonimized, with Ms. Menten serving as the pseudo-anonimizer key holder. For these human samples, approval from the ethical committee is mandatory. This has already been granted: S58824.

A follow-up ethical approval for future human sample collection is currently pending and expected to be granted H2 2023. This approval is not required for the continued use of the samples already collected under S58824.

Animal: This applies to WP1, WP3, and WP4. Animal experiments are being performed, for which approval from the ethical committee is mandatory. This has already been granted under project number P182/2021.

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

This applies to WP2. Data are being used and the personal data of donors and patients will be stored. This includes 'regular' personal data such as age (years), weight, sex, but also 'special' or 'sensitive' personal data such as health data. To dr. Ronda, these data are anonimized, with Ms. Menten serving as the pseudo-anonimizer key holder. Data concerning name, family name, address, e-mail address, rijksregisternummer, date of birth, are not stored and not processed.

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

The work proposed in the application is highly discovery orientated. However, biomarkers of cardiac ageing or of susceptibility of aged heart to disease may be of interest for development as a prognostic of cardiac health, for example in determining suitability of hearts for heart transplantation.

When appropriate, this work may be developed under a university research fund for translational research.

Mechanistic/sequencing studies may also yield new targets for therapeutic intervention. These findings will be a long way from translation requiring validation. This will be carried out under a subsequent project. Should any IP be generated that has potential for exploitation, we will seek to protect this IP. It is unlikely this will be the case for findings distal to translation – such as the mechanistic studies.

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.

- No

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).

Day-to-day documentation is performed in a paper bound lab book. Methods are documented as digital SOPs and maintained in a shared database in the lab as .docx. All SOPs have a version number with the author of the version indicated.
Experiments will be documented in lab books. Each book contains the name and contact information of the author. As multiple books exist, each is numbered. Each page inside the lab book is numbered and dated. Each pahe contains a header containing the experimental code (e.g. OR53) and a running title describing the subsequent experimental notes (e.g. cDNA synthesis from RNA using SS3). An SOP is referenced. For experiments which use biological reagents such as antibodies, care is now also taken to include lot/batch numbers of reagents.
For animal experiments, procedures are documented in the LAIS (Leuven Animal Information System), whereby the origin/heritage and exact experimental procedures and timings are documented.
The lab increasingly makes use of videography and photography to document e.g. murine surgeries for training and documentation purposes. Other laboratory methodologies e.g. FACS enjoy increasing use of photography and including these images into SOPs for improved clarity by follow-up or novice users.
On a 6 monthly basis, data is transferred to the Archive K drive of KUL. This data is stored under a laboratory-agreed format, including with metadata, and is read-only after copy transfer.
Research carried out during this project will be published in academic journals that will be openly accessible to our peers, as well as openly accessible to the public. Detailed methods and links to deposited data that will allow repeat of experiments and data re-interrogation respectively will be included in the manuscripts. Large data sets, metadata and non published findings will be deposited in online repositories such as https://www.ebi.ac.uk/ena/ and the KUL repository RDR (https://rdr.kuleuven.be/).

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

We currently have no metadata standard. However, where data is collected on specific equipment, e.g. confocal (imaging), FACS (cytometry), LICOR (westernblot), metadata relating to acquisition parameters are collected. In general, we collect and store as much metadata as possible, and the digital files are stored inside a directory structure indicating a static experimental code (e.g. OR53), inside sub-directories indicating analysis type (e.g. qpcr, imaging, sequencing). The static experimental codes are used throughout the data structure and lab books, so that its code + date are sufficient to cross-reference digital data to paper lab books and vice-versa.
In brief, raw data files are always dated, annotated with the experiment's code, the sample's ID, and any relevant parameters.

Example 1: a typical confocal microscopy file will be named as below. The filename indicates the acquisition date in a numerically machine-sortable way, the species (mouse), tissue type (heart), the animal's ID no (126), lens configuration (40x), section and field number (technical replicates), dyes or antibodies used on each channel. Each microscopy experiment is accompanied by a notes.txt which describes the exact dyes and antibodies and at which dilution they were used, and any specific details related to the experiment.
2022-11-28_Mouse_heart_126_40x_section3_field3_DAPI_(blue)_Tet2_Abe364_(green)_ACTN2_(red)_WGA_(cyan).nd2
Example 2: a typical FACS experiment file will be named as below. The filename indicates the acquisition date in a numerically machine-sortable way, indicates the sample's ID, and indicates whether it relates to an analysis or a sort. Each sample is stored as as .fcs as well as a .pdf which is an export from the facs machine's software, which describes the optical and laser configuration, the sensor's amplification and thresholding settings, the gating, and the (sorted) counts. Each experiment is accompanied by a document containing screenshots on the machine's droplet stream, system parameters and other remarked info.
2022-12-01_06_HD26_SORT

In addition:
For **patient samples**, metadata are noted in a shared database with criteria related to the isolation and use of derivatives included as well as quality of sample.
For **animal samples**, metadata are recorded in LAIS (https://lais.cidercone.com/), and in excel sheets on a per-experiment basis.

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

Where will the data be stored?

Data is temporarily stored during collection on related equipment (e.g. confocal microscopy PC, echocardiography PC). Data is immediately transferred to network drives – KUL L drive (large volume storage). On a 6 monthly basis, and upon publishing, data are transferred to the KUL archive K drive.
Data is shared between labs using Belnet FileSender (https://filesender.belnet.be/). There are no costs (to the researcher) associated with the use of this service.

How will the data be backed up?

Digital data are backed up and versioned automatically by their storage on the L drive (large volume storage). ICTS is responsible for the automatic versioning and backup (mirroring) of all data stored on the shared network drives including the large volume drive.
Physical data (e.g. biological samples) cannot be backed up, but are continously monitored using the aeroscout temperature monitoring system. At all times, there is a person on-call in case of -80°C and -20°C freezer failure and over-temp. At all times, we have access to a backup -80°C and -20°C freezer, to move samples from a failing freezer to an operational one.

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.

- No

The lab currently has 10 TB storage space on the L large volume network drive, of which approx 2 TB is available. This would not be sufficient for the projected data of this project.
However, I foresee that (over time) additional space is made available by moving old datasets to the K archive drive, and/or by requesting additional storage space. Through these means, sufficient space will be available for this DMP's project. Costs of digital data storage will be paid for using lab funds.

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

Data are securely stored by their storage on the KUL network drives. Data are not stored on personal computers or external harddrives / devices.
Access to the lab's shared L drive is restricted to members of the lab. Students and non-lab members are not given access to these resources. Inside the shared drive are personal

directories for each lab members. It is agreed upon in the lab that these directories can not be modified by other lab members. Upon archival of raw data onto the K archive drive, data can be read but cannot be modified by researchers.

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

The current price of the L drive is €569,2 / 5.0 TB / year. The current price of the K drive is €284,6 / 5.0 TB / year.

This research project will gradually generate data. As soon as possible, the raw data are transferred to the K drive. For the 3 year duration of this project, the projected digital data storage costs are €284.6 × 0.5 × 3 = €426.90. These digital data storage costs are covered by lab funds.

The expected costs to store 25 cryoboxes at -80°C is more difficult to assess, but will be borne by the host laboratory.

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...).

Digital data:
In the case of relatively small digital data, data will be preserved for longer than 5 years. The decision when to delete these data is made by prof. Roderick.
For large digital datasets, such as raw microscopy images, unused datasets will be deleted. Raw sequencing and (methylation) microarray results are submitted to the NCBI SRA and deleted from our local storage after the 5 year period. To the researchgroup and the KUL, there are no storage costs associated with the SRA.

Physical data / biological samples:
Physical data stored at -80°C is costly, however beyond the 5 year period, such samples are stored until higher priority samples require the occupied space. These decisions are made by prof. Roderick.
Physical data storage space at -20°C is much less costly but beyond the 5 year period, preservation or disposal of samples is similarly decided upon at the discretion of prof. Roderick.
Physical data stored at RT (lab books) are stored until the retirement of prof. Roderick.

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

Digital datasets are stored on the KUL L drive (large volume shared network drive), and are moved to the KUL K drive (archive drive) as soon as possible. The K drive is write-once + read-only for researchers, i.e. data cannot be modified after being written. As these network drives are managed by the KUL ICTS, and backed up in 2 physical locations, storage of digital data on the K archive drive protects data from (accidental) loss, degredation, and change.

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

The current price of the L drive is €284,6/ 2.5 TB / year. The current price of the K drive is €142,3 / 2.5 TB / year.

The expected costs to store the ~5 TB generated in this project for 5 years, given that the raw data will reside mainly on the K drive, the costs to store these data for 5 years is €142.30 × 2 × 5 = €1423. These digital data storage costs are covered by lab funds.

The expected costs to store 25 cryoboxes at -80°C is more difficult to assess, but will be borne by the host laboratory.

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

All data included in or referred to in published manuscripts will be made publicly available/accessible. Any data to be used to applications for future TBM applications may be restricted, although may not be published.
We foresee that all data will be published in peer reviewed academic journals in compliance with KUL Open access policy, and where possible fully open access. All manuscripts will be uploaded to LIRIAS.

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

Any human personal data will remain strictly confidential. Only anonimized data will be used in manuscripts. Ms. Roxane Menten holds the key to de-anonimize these data. Students and non-lab members are not granted access to these personal data. We do not anticipate a situation where external researchers need to access these personal data.

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

Should any commercially exploitable IP be generated, these data will not be shared. Due to the discovery nature of the project however, we do not anticipate, significant restrictions.

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

We foresee that all data will be published in peer reviewed academic journals in compliance with KUL Open access policy, and where possible fully open access. All manuscripts will be uploaded to LIRIAS.
Raw sequencing reads and DNA methylation microarray data will be deposited into NCBI SRA.

When will the data be made available?

After publishing.

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

Creative Commons Attribution-ShareAlike (CC-BY-SA)

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

Where possible and appropriate, an accession number will be added to the datasets.

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

The use of Belnet FileSender for transfer of data between researchers is free to researchers. A maximum of 6 TB can be transferred simultaneously using the Belnet Filesender.
Depositing data onto SRA is free to researchers.

6. Responsibilities

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

The main researcher (dr. Onne Ronda) is responsible for data documentation, supervised by Ms. Roxane Menten and prof. Llewelyn Roderick.

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

The main researcher (dr. Onne Ronda) is responsible for data storage, supervised by Ms. Roxane Menten and prof. Llewelyn Roderick.

Who will manage data preservation and sharing?

Dr. Onne Ronda, and prof. Llewelyn Roderick.

Who will update and implement this DMP?

The principal investigator, prof. Llewelyn Roderick, is utimately responsible for implementing this DMP.