SORDD polyneuropathy: unraveling a polyol pathway disorder

A Data Management Plan created using DMPonline.be

Creators: Bram Decru, n.n. n.n., n.n. n.n.

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., n.n. n.n.

Data Manager: Bram Decru

Project Administrator: Bram Decru

Grant number / **URL:** 1126523N

ID: 191081

Start date: 01-11-2022

End date: 30-10-2026

Project abstract:

Sorbitol dehydrogenase (SORD) deficiency was only identified in 2020 as the most frequent recessive form of hereditary neuropathy. SORD enzyme loss-of-function generates an intracellular accumulation of sorbitol in the polyol pathway. This pathophysiologic mechanism has also been suggested as a major determinant in diabetic polyneuropathy. The aim of this project is to unravel the polyol pathway and metabolic rewiring in SORDD polyneuropathy using novel techniques such as tracer metabolomics and high-resolution respirometry in fibroblast cells. The mechanism of underlying alterations will be elaborated at the transcriptional and post-translational level to elucidate the pathophysiologic mechanism. Next, to support our *in vitro* findings, we will study neuron performance by electroretinogram in a Drosophila model I will use enzyme selective pharmacological and genetic modifications to support the *in vitro* findings. To date, no in-depth studies using state-of-the-art high throughput mass spectrometry technologies have been performed in this area. If successful, this project will novel insights and could provide ground-breaking information identifying novel treatment strategies for SORDD-mediated polyneuropathy. Moreover, these novel insights about polyol metabolism could also impact research into other disorders associated with polyol accumulation such as diabetic polyneuropathy

Last modified: 27-02-2023

SORDD polyneuropathy: unraveling a polyol pathway disorder
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)

The SORDD study, as approved by the UZ/KU Leuven ethical committee (S66609) will only collect new data. This study has not been performed before and shall therefor only use new data, i.e. not reuse existing data. Data will be pseudonymized with a letter referring to the status of patient or control and a 4 digit code. This coding will be locked in a password protected excel file on a long term storage drive of the UZ Leuven with limited acces for PhD students of prof P. Vermeersch. The password is only known by persons connected to the research (i.e. the PhD investigator and promotor).

Description	Extension	Estimated volume	Notes
Mass spectrometry data	.raw	High estimation* per 0.2 GB average sample. Low estimate of 520 samples: **104GB** in total.	This is the raw files extracted from the mass spectrometer. This data is converted into ".mzML" files to be analyzed. Both ".raw" and ".mzl" data are stored. *The high estimation is based on the average memory necessary for a run of 12 samples including the additional memory for the run of 3 blanks and 2 pre-runs necessary to interpret 12 samples (total of 2.4 GB). Since runs of more than 12 samples will be run, this is an overestimation.
Mass spectrometry data	.mzML	High estimation* of 0.12GB per sample Low estimate of 520 samples: **62.4 GB**	After processing these files, these data are extracted into an ".xlsx" data form.
Mass spectrometry data	.xlsx	0.6MB per file: Low estimate of 52 runs: **31.2MB**	4 files (saving intermediate processing steps) per run.
Mass spectrometry output	.png	Megabytes	Processing by Microsoft Excel, Graphpad Prizm and Rstudio (".Rmd" format) to have data graphs.
Respirometry	.datlab	+/- 20MB per run. Total: up to 1GB.	Raw respirometry files. Amount cannot be estimated. Somewhere between 20-50
Respirometry	.xlsx	Megabytes	Data extracted from raw .datlab files
Western blotting	.tiff .xlsx	Megabytes	Pictures in ".tiff" format Intensities will be processed using excel
RT-qPCR	.txt .xlsx	Megabytes	Processing in excel
Manuscript and presentation	.doc .pptx	Megabytes	
Literature	.pdf	Megabytes	

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: my promotor prof. Dr. Pieter Vermeersch.
    1. In accordance with the KU Leuven policy, data will be stored for 10 years after research.
2. Storage capacity/repository
    - during the research:
        - use of KU Leuven onedrive for business for pseudonymized data (up to 2TB of storage capacity)
        - Back-up data of data on KU Leuven long term storage drive (storage capacity is hired in blocks of 5 TB)
        - Password protected coding file: restricted acces folder on longterm storage drive of UZ Leuven
    - after the research:
        - storage on longterm storage drive of KU/UZ Leuven, accessible by prof. dr. Pieter Vermeersch.

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)

Not applicable.
Data will be stored for 10 years in accordance with KU Leuven policy.

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)

Sensitive data will be stored as described above in a password protected file in a folder with highly restricted acces. This is deemed as in accordance with current GDPR directives and approved by the ethical committee of KU/UZ Leuven.
There is no data with misuse potential other than the aforementioned personal data.

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

Not applicable

SORDD polyneuropathy: unraveling a polyol pathway disorder
DPIA

DPIA

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

- No

SORDD polyneuropathy: unraveling a polyol pathway disorder
GDPR

GDPR

Have you registered personal data processing activities for this project?

- Yes

SORDD polyneuropathy: unraveling a polyol pathway disorder
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
Fibroblast culture cells of patients	Derivatives obtained from the CME after skin puncture	Generate new data	Physical				Cellines will be obtained from +/- 10 patients and kept in a steady stock of 5-10 vials
Fibroblast culture cells of controls	Derivatives obtained form the CME after skin puncture. Controls from S58977 will be used, after new IC is approved by the individuals	Reuse existing data	Physical				10 controls, Kept in a steady stock of about 5-10 vials of 1mL
urine	urine of patiënt collected at time of punction biopt	Generate new data	Physical				20mL of urine volume
blood	red blood cells	Generate new data	Physical				Red blood cells collected from 4-8mL of blood cells
Mass spectrometry samples	Samples after preparation. This includes medium, cellysate, protein pellet for BCA and methanol supernatans	Generate new data	Physical				Variabel volume, stored in -80°C
Mass spectrometry & chromatograms data	This is the raw files extracted from the mass spectrometer. This data is converted into ".mzML" files to be analyzed. Both ".raw" and ".mzl" data are stored.	Generate new data	Digital	experimental	.raw	<1TB	
Mass spectrometry & chromatograms data	After processing these files, these data are extracted into an ".xlsx" data form.	Generate new data	Digital	compiled/aggregated data	.mzML	<100GB	
Mass spectrometry & chromatograms data	4 files (saving intermediate processing steps) per run.	Generate new data	Digital	compiled/aggregated data	.xlsx	<1GB	
Mass spectrometry & chromatograms data	Processing by Microsoft Excel, Graphpad Prizm and Rstudio (".Rmd" format) to have data graphs.	Generate new data	Digital	compiled/aggregated data	.png	<1GB	
Respirometry	Raw respirometry files. Amount cannot be estimated. Somewhere between 20-50	Generate new data	Digital	experimental	.datlab	<1GB	
Respirometry	Data extracted from raw .datlab files	Generate new data	Digital	compiled/aggregated data	.xlsx	<1GB	
Western blotting	Pictures in ".tiff" format	Generate new data	Digital	experimental	.tiff	<100MB	
Western blotting	Intensities will be processed using excel	Generate new data	Digital	compiled/aggregated data	.xlsx	<100MB	
RT-qPCR	Processing in excel	Generate new data	Digital	experimental	.txt	<1GB	
RT-qPCR		Generate new data	Digital	compiled/aggregated data	.xlsx	<100MB	
Manuscript and presentation		Generate new data	Digital	compiled/aggregated data	.doc	<100MB	
Literature		Generate new data	Digital	compiled/aggregated data	.pdf	<1GB	
Drosphila flies	flie model	Generate new data	Physical				
electroretinogram		Generate new data	Digital	experimental	.abf	<100MB	
electroretinogram	processed data, with igorPro	Generate new data	digital	compiled/aggregated data	.xlsx	<100MB	

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:

Not applicable

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

Ethical approval was obtained from the KU/UZ Leuven ethical committee (S66609). The study includes the collection of human blood, urine and fibroblast punctures.

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

In the coding file the EADnumber (administrative identification number used in NEXUZ hospitals) is coupled to the pseudonym. At this point, no other personal data is obtained or will be 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.

- No

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

A material transfer agreement is coupled to S66609 between the KU Leuven acting on request of prof Vermeersch en prof Claeys and the VIB at facilities under the direction of prof. Bart Ghesquière for the mass spectrometry analysis. De facto the aggreement balances the intellectual property rights, confidentiality obligations towards oneother and concerning personal data. Since the provider and recipient are promotor and co-promotor respectively, I do not foresee major restrictions to be in place. Data and research data can be uploaded to a data repository on the condition that the upload, storage and further use of any personal data complies with data protection laws. Appropriate contractual arragnements with the Data repository must be made to ensure appropriate mechanisms are put in place that allow access control.

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

A material transfer agreement is coupled to S66609 between the KU Leuven acting on request of prof Vermeersch en prof Claeys and the VIB at facilities under the direction of prof. Bart Ghesquière for the mass spectrometry analysis. The agreement compasses joint ownership of the parties.

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).

Experiments accompanied by their data are kept in a specific map. This map is truncated as follows:
- Sxxxxx (ethical administrative number) brief name of the study
--Workpackage as described in the FWO application
--- Experimentdescription

workpackage

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

For the metadata a standard vocabularium for folders, and specific metadata will be used to keep the data accessible. However, given the currently low use of these standard metadatasets in our lab and collaborating offices, a clear consistent use of vocabularium to describe the data will be used.

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

Where will the data be stored?

1. During the research:
    - use of KU Leuven onedrive for business for pseudonymized data (up to 2TB of storage capacity)
    - Back-up data of data on KU Leuven long term storage drive (storage capacity is hired in blocks of 5 TB)
    - Password protected coding file: restricted acces folder on longterm storage drive of UZ Leuven
2. After the research:
    - storage on longterm storage drive of KU/UZ Leuven, accessible by prof. dr. Pieter Vermeersch.

How will the data be backed up?

Data will be backupped using Druva InSync software available via the VIB.

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

Yes, there is enough data storage capacity.
The KU Leuven onedrive server allows the storage of 2 TB of data per person.
The longterm storage is hired per 5TB.

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

To acces the KU Leuven onedrive, one needs to login with the multistep authentificator of the KU Leuven (fingerprint or code specific on the personnel smartphone).
Back-ups are password protected.
The L-drive has restricted acces for the research group and is protected by the KU Leuven.

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

The expected costs are beared by the KU Leuven and the research group.
Onedrive for business is offered free to KU Leuven personnel.
Long term storage is available for € 100,86 / TB / year

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...).

Data will stored for at least 10 years according the KU Leuven RDM policy.

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

These data will be stored on a longterm storage drive of the KU Leuven (L-drive) and/or UZ Leuven (UZDATA) with restricted acces to the team involved with the project (i.e. (co)promotors, research collegues). The expected volume will be <2TB.

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

The currect costs for preservation on the longterm storage drive of the KU Leuven is € 100,86 / TB / year. These costs are covered by the Ghesquière lab. Potential costs for long term storage on the long term storage of the university hospital is covered by the hospital.

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

Data will be available upon request by e-mail to the corresponding author.

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

Data will be available upon request by e-mail to the corresponding author.

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, Ethical aspects

All (meta)data is subject to the GDPR and Belgian law concerning data protection and will be shared according these laws.
All data originated from blood, fibroblast, liver or muscle obtained in the context of S58977 or fibroblast mass spectrometry data obtained in the context of S66609 is subject to agreements as described in the material transfer agreements of the KU Leuven and VIB represented by Prof. Pieter Vermeersch- prof. David Cassiman and prof. Bart Ghesquière respectively.

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

All (meta)data will be made available in the long term storage for (co)promotors and research collegues from the Bart Ghesquière laboratory as well as research fellows of prof. P. Vermeersch.
Relevant reproducible data will be made available via LIRIAS, the repository made available by the KU Leuven. If requiered or opportune, the data will be attached as addendum to a publication.

When will the data be made available?

The relevant data will be made available upon publication of the research results.
All (meta)data will be made available to the research group after completion of the PhD project.

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

Data can be used on request after signing a data sharing agreement. The request
can be made by sending an e-mail to the corresponding author of the resulting paper. Only uses for research purposes will be allowed and commercial reuse
will be excluded.

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.

- No

A permanent identifier is added to the data upon deposit in a repository. Check which identifier is used by your (candidate) repository, most repositories including KU Leuven RDR use a DIO, in life sciences accession numbers are commonly used.

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

We do not expect costs for data sharing.

6. Responsibilities

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

Bram Decru

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

Bram Decru

Who will manage data preservation and sharing?

Bram Decru

Who will update and implement this DMP?

Bram Decru

.