TOWARDS DIAGNOSTIC USE OF NEUTROPHIL EXTRACELLULAR TRAP (NET) MEASUREMENTS IN PATIENT SAMPLES BY PAIRING NOVEL BIOMARKER MEASUREMENTS WITH ESTABLISHED MICROSCOPY READOUTS

A Data Management Plan created using DMPonline.be

Creator: Kim Martinod

Affiliation: KU Leuven (KUL)

Template: FWO DMP (Flemish Standard DMP)

Principal Investigator: Kim Martinod

Grant number / **URL:** CELSA/22/024

ID: 194491

Start date: 01-10-2022

End date: 30-09-2024

Last modified: 19-01-2023

TOWARDS DIAGNOSTIC USE OF NEUTROPHIL EXTRACELLULAR TRAP (NET) MEASUREMENTS IN PATIENT SAMPLES BY PAIRING NOVEL BIOMARKER MEASUREMENTS WITH ESTABLISHED MICROSCOPY READOUTS

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)

Physical data: Plasma samples (human) : -80 degree freezer – 10 boxes, Multiwell plates with cells (human) - 4 degree fridge, -20 degree freezer – 30 plates, Cell digest supernatants - -20 degree freezer, 20 boxes

Data sources and experimental data: Researchers involved in the work will use paper lab books, which will include sufficient details to perform experiments and to trace data. Lab books are stored securely in locked offices and after completion in the PI's locked office. All methods used are transposed into a lab template based standard operating procedure (SOP) and maintained on the J drive (total ~ 0.2 GB). Data consists of microscopy images (.tiff, .czi files; total ~500 GB), Excel files (.xlsx, ~10 MB). Figures are generated as GraphPad Prism files (.pzf, .pdf, ~50 MB).

Flow cytometry of immune cells: files exported from analysers in non-proprietary .fcs 3.0 format which includes metadata. ~100 GB of data will be collected, and processed data is stored as FlowJo workspace files (.wsp), ~ 2 GB estimated volume.

Processed data and statistics: Initial data collation in Excel and then analysis in PRISM V9 (.pzf files). total <5 GB. R scripts used in some of this analysis will also be saved, R files, 50 MB.

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: the PI, Kim Martinod
2. Storage capacity/repository
    - during the research:

      KU Leuven shared network drive space is purchased on an as-needed basis, with volumes ranging from 100 GB to several TB. Our current shared drive has a capacity of 1.6 TB, which is sufficient for the remainder of this project.
    - after the research:

      Data will be preserved for the following 5 years after the end of research, and any data related to human subjects for 25 years.

      Long-term storage will be via the server backend storage, large volume storage and archive drives administered by KU Leuven ICTS.

      Physical lab notebooks will be stored in a central location in the offices in the Center for Molecular and Vascular Biology (CMVB).

      Published data will be made available through the use of an open repository, for example the Image Data Resource for microscopy images. This allows for further analysis of public datasets.

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)

We will not deviate from the minimum preservation term of 5 years.

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)

Blood samples from human healthy volunteers will be used, with samples anonymized (no personal data). Use of these samples is covered by the approval S63490 from the Ethics committee research UZ / KU Leuven (EC research). Amendments and extensions will be made to these existing files in the future.

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

N/A

TOWARDS DIAGNOSTIC USE OF NEUTROPHIL EXTRACELLULAR TRAP (NET) MEASUREMENTS IN PATIENT SAMPLES BY PAIRING NOVEL BIOMARKER MEASUREMENTS WITH ESTABLISHED MICROSCOPY READOUTS

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.

Data collection	Data processing	Data analysis
Origin	Type	Format
Protocols	Digital - processed - textual	.docx
Informed consent forms (healthy volunteers)	Physical - raw - textual	Paper
Scripts statistical analyses	Digital - processed - textual	.R
Whole blood samples (humans)	Physical - raw - organic	/
Whole blood samples (humans)	Physical - raw - organic	/
Whole blood samples (humans)	Physical - raw - organic	/
Whole blood samples (humans)	Physical - raw - organic	/
Whole blood samples (humans)	Physical - raw - organic	/
Whole blood samples (humans)	Physical - raw - organic	/
Whole blood samples (humans)	Physical - raw - organic	/
Whole blood (human) + bacteria	Physical - raw - organic	/
Whole blood (human) + bacteria	Physical - raw - organic	/
Whole blood (human) + bacteria	Physical - raw - organic	/
Whole blood (human) + bacteria	Physical - raw - organic	/

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:

N/A

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

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

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.

- 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).

Research carried out during this project will be published in academic journals that will be openly accessible to our peers as well as 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 MSs and/or their supplementary files. 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/).
Methods used for our analysis are written up as SOPs and maintained in a shared database in the lab as .docx. All SOP have a version number with the author of the version indicated. Experiments performed will be documented in lab books – personal or for animal experiments a shared lab book on the shared KUL J drive. Data and its location will be referenced to in lab books. Care is now also taken to include lot/batch numbers of reagents used in experiments.
On a 6 monthly basis, data is transferred to the L drive of KUL. This data is stored under a laboratory agreed format, including with metadata. This is regularly backed up by KU Leuven IT.

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. The following metadata will be created for each dataset/image acquisition:

1. Project information: project title, project funder (name of funder, type of grant, grant number)
2. Numbers and names of files(s) or dataset(s) that this README file describes: names of files, description, date of creation of file, file format, software used to generate the data, software necessary to open the file, relationship between the files
3. Description of the dataset: date and place of data collection, data collecting method, information about data processing methods, information about the instrument/calibration
4. Information about limitations of the dataset, information that ensures correct interpretation of the dataset
5. Date of creation/last update of the README file
6. Name and contact information of Principal Investigator
7. Institution of Principal Investigator
8. Other involved researchers
9. Data information: thesaurus or controlled vocabulary keywords, full names and definitions for columns and rows, explanation of abbreviations, units of measurement, symbols for missing data
10. Storage information: where are the data stored
11. Publications based on this dataset
12. References of publications used to create the datasets

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

Where will the data be stored?

Data will be stored on the J drive of the Martinod Lab (GBW-0449_CMVB_Martinod_Lab) or L drive of the Martinod Lab (GBW-0123_Martinod_Lab).
We are also in the process of investigating how to migrate data to a Sharepoint site.

How will the data be backed up?

Automatic backups of the central server storage space is organized by KU Leuven ICTS. A copy of all the data on this central network drive is stored on a second location, which guarantees their availability even in the event of problems with the hardware (“disaster recovery”).

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 shared network drive space is purchased on an as-needed basis, with volumes ranging from 100 GB to several TB. Our current shared drive has a capacity of 1.6 TB, which is sufficient for the remainder of this project.
OneDrive for Business provides a 2 TB online storage capacity with the possibility to upload files up to 100 GB. If necessary, the 2 TB limit of the OneDrive for Business account can be extended up to 5 TB if motivated and requested through ICTS Service Point. Desktop file storage can also be expanded up to 5 TB.

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

KU Leuven requires login and multifactor authentication to protect unauthorized access to all data.

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

Costs for OneDrive are supported centrally by KU Leuven.
Desktop file storage costs €864 per TB per year and is supported by this CELSA funding.

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 be preserved for the following 5 years after the end of research, and any data related to human subjects for 25 years.
The following will not be retained past the end of the study period (after publication of all results): stained multiwell cell culture plates. However, all related electronic documentation on these data will be saved.
Reused human samples will be stored for a 10-year period according to Biobank approval.

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

Long-term storage will be via the server backend storage, large volume storage and archive drives administered by KU Leuven ICTS.
Physical lab notebooks will be stored in a central location in the offices in the Center for Molecular and Vascular Biology (CMVB).
Published data will be made available through the use of an open repository, for example the Image Data Resource for microscopy images. This allows for further analysis of public datasets.

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

Costs are 2246 euros per year and will be funded from future research grants from Prof. Martinod

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, …)
- No (closed access)

Data will be available on request to the PI after signing a data sharing agreement. The procedure for requesting access to data is reported in each published manuscript.
Flow cytometry or microscopy data raw files will be uploaded via a data respository upon study publication.

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

Data will be available on request to the PI after signing a data sharing agreement.

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

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

Not yet known

When will the data be made available?

- Upon publication of the research results

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

N/A

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

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

N/A

6. Responsibilities

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

PhD researchers, postdoctoral fellows, and technical staff will be responsible for data documentation and metadata. Prof. Martinod is responsible for the oversight of this documentation.

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

PhD researchers, postdoctoral fellows, and technical staff will be responsible for data storage and backup. Prof. Martinod is responsible for the oversight of this storage and backup.

Who will manage data preservation and sharing?

Professor Martinod

Who will update and implement this DMP?

Professor Martinod