Interorgan Crosstalk in CKD-MBD: On the road to Precision Medicine

A Data Management Plan created using DMPonline.be

Creators: Dieter Smout, 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: Dieter Smout

Data Manager: n.n. n.n.

Project Administrator: n.n. n.n.

Grant number / URL: 11J1623N

ID: 198629

Start date: 31-10-2022

End date: 29-09-2025

Project abstract:

Chronic kidney disease (CKD) is a common health problem affecting 10-15% of the population. Patients with CKD suffer multiple metabolic complications, among them mineral and bone disorders (MBD). CKD-MBD results in excessive morbidity and mortality due to fractures and cardiovascular disease. Progress within the field of CKD-MBD is hampered by a lack of validated non-invasive diagnostic tools, as well as an incomplete understanding of the pathophysiology involved. These obstacles need to be overcome in order to establish evidence-based treatment and improve patient outcomes. Growing experimental evidence points at an important role of the gut microbial environment and immune system as a mediator of CKD-MBD. Newly discovered molecules such as microRNAs may clarify obscure pathways in this complex disorder or act as diagnostic markers. This project seeks to advance the field by an in-depth investigation into the role of the gut microenvironment and immune system as well as exploration of microRNAs for diagnosis, prognosis, and novel treatment targets for CKD-MBD. In WP1, I will investigate whether CKD and hyperparathyroidism associate with a disturbed Th17/Treg balance, and how this disturbance affects bone health. I will use retrospective clinical data and an experimental rat model. In WP 2, I will explore the involvement of a disturbed gut microbial environment in the pathogenesis of bone loss in clinical CKD and explore potential therapeutic strategies in a uremic rat model. In WP3, I will use RT-qPCR Taqman microarrays to discover miRNAs in bone (marrow) and circulation that associate with bone turnover. In a validation cohort I will test the diagnostic performance of a selected miRNA panel.

Last modified: 31-07-2023

Interorgan Crosstalk in CKD-MBD: On the road to Precision Medicine
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		
Rat biosamples	blood, urine, faeces and tissues (liver, kidney, bone, aorta) from experimental rat studies	Generate new data	Physical				Approximately 10 boxes of 9x9 2 mL tubes.	
Analytical rat data	Data derived from biochemistry tests including ELISA, PCR, LCMS-MS	Generate new data	Digital	Experimental	.csv .xls	<100MB		
microCT	Images and analyses using microCT on rat bones	Generate new data	Digital	Experimental	.bmp .xlsx			
miRNA	Raw data of miRNAs	Generate new data	Digital	observational	.EDT			
Analytical rat data	Data derived from biochemistry tests including ELISA, PCR, LCMS-MS	Generate new data	Digital	Experimental	.csv .xls	<100MB		
microCT	Images and analyses using microCT on rat bones	Generate new data	Digital	Experimental	.bmp .xlsx TBD	<100GB		
Histology	Images from histology	Generate new data	digital	Experimental	TBD	<100GB		
miRNA	Raw data of miRNAs	Generate new data	Digital	observational	.EDT	<1GB		
databases	Clinical data from databases	Reuse existing data	digital	observational	.xlsx	<1GB		
R	R scripts and markdowns	Generate new data	digital	aggregated	.R .rmd	<100GB		
Tables and figures	Tables and figures created by R	Generate new data	digital	aggregated	.ppt .xlsx .docx	<100GB		

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:

We will reuse datasets and biosamples collected in the frame of completed or ongoing clinical cohort studies locally stored in our databases (S63818, S63813, S51935, S55578, S59393, S52091, S56905).

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

We will perform secondary analyses on data that has been collected under ethical approval of the KU Leuven. The data sets that will be used in this project are mentioned above and are pseudonymized.
The animal experiments are awaiting ethical approval before starting.

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

Personal data collected in ongoing studies will be pseudonymized before stored in the databases and further analytical processing. I will work with the pseudonymized data.

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).

Each database has their own legends, similar datasets are bundled in the same database to get better overview. Databases are constructed as similar as possible (identical names for columns describing the same variables). Additional rows and columns are added in the databases to explain variables or add comments. The datasets include a logbook to track changes made.
README.txt files are created for interpreting secondary data files.
R scripts are created with the necessary explanations using # inside the scripts or with R.markdown files.

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

- Yes

Within our lab we have created our own standards for metadata which will be used for the databases.

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

Where will the data be stored?

All electronic data will always be stored in the password-protected UZ Leuven network environment and personal UZ Leuven OneDrive. If needed, anonymized data will temporarilly be stored on a local drive.

Longtemr storage (archived data) will be done on the lab's UZ Leuven K drive.

How will the data be backed up?

The UZ Leuven servers and UZL OneDrive benefit from automatic back-up services.

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

There is sufficient storage and back-up capacity on the KUL OneDrive server and all UZ Leuven and KU Leuven servers:

- KUL OneDrive provides 2TB storage for non-confidential data, which is estimated to be adequate for the entire research project.
- The KUL/UZL “L-drive” is an easily scalable system, has unlimited maximal size and is expandable in blocks of 5TB.
- The KUL/UZL “J-drive” is an easily scalable system, has unlimited maximal size and is expandable in blocks of 100GB.
- The Staging and Archive on VSC are also sufficiently scalable (petabyte scale).

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

The UZ Leuven server of the Department of Nephrology is password-protected and benefits from the firewall services provided by the UZ Leuven IT department. Access to the data as well as the access level will be limited on a project need and individual basis.
Personal data is always pseudonymized. The coding key to patient information of linked pseudonymized data does not carry any personal identifiers and all records containing the identity of each participant will be kept private and confidential.

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

The total estimated cost of data storage during the 4 years of this FWO project is less than 1,000 EUR. This estimation is based on the following costs:

- The costs of digital data storage are as follows: €868/5 TB/Year for the “L-drive” and €519/TB/Year for the “J-drive”.
- The cost of VSC archive is €70/TB/Year and staging €130/TB/Year.

The physical samples will be stored within available freezers in our laboratory.
The costs for data storage will be covered by our own 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...).

All data generated and the biological material taken will be retained for minimally 10 years after ending the FWO project. This is because of the possibility of reuse of samples or data
for new research projects conducted by our research group.

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

long term storage will be ensured as follows:

- Large data will be stored on VSC archive.
- Small digital files and database extracts will be stored on the KUL/UZL “L-drive”.
- Developed algorithms and software will be stored on VSC archive and/or L-drive, as well on public repositories such as Github.com.

The physical samples will remain stored within available freezers in our laboratory. Most samples will have been used and destroyed by this time.

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

The total estimated cost of electronic data storage for 5 years after the end of the project is < €700. This estimation is based on 1 TB in total, at 70EUR/Tb/year.
All costs for data preservation will be covered by our own funding.

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, ...)

Regarding digital data, only published data will be available in the form of publications or other dissemination of scientific work. All data will be anonymised when disseminated. More data can be made available or shared after permission of the responsible person (prof. Pieter Evenepoel). Non-published data will remain confidential until a final decision on publication of the data has been taken.
Physical data will remain avialable within the lab of nephrology and renal transplantation for the duration of 10y. R scripts may be made publicly available on GitHub.

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

Data will be stored on servers only accessible to members of the nephrology and renal transplantation group. Outside this group, data can only be shared with permission of prof. Evenepoel and after signing a data sharing agreement, when required for colaborative research.

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

Some of the analyses will be performed in colaboration with other laboratories. In that case, an agreement has been or will be made prior to disclosure of data.

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

- Double-coded raw microarray data (linked to double-coded patient data) will be deposited in open access repositories with restricted access control such as the EBI European Genome-phenome Archive (EGA). The EGA is a repository for personally identifiable genetic and phenotypic data. microarray data at EGA will only be available upon reasonable request via our institutional data access committee and, if necessary, a material transfer agreement will be concluded with the beneficiaries in order to describe the types of reuse that are permitted.
- Double-coded patient data: Upon publication, all double-coded patient details supporting a manuscript will be made publicly available as supplemental information.
- Research documentation: All protocols used to generate published data will be described in the corresponding manuscript(s), and the related documentation will be included as supplementary information. These data and all other documents (raw data) deposited in the E-Notebook are accessible to the research staff and will be made available upon request.
- Manuscripts: All scientific publications will be shared openly. Manuscripts submitted for publication will be deposited in a pre-print server such as bioRxiv. Before the end of the embargo or in cases where sharing the post-print is not allowed due to copyright agreements, a pre-print version of the manuscript will be made available. (Pre-print) publications will also be automatically added to our institutional repository, Lirias 2.0, based on the authors name and ORCID ID.
- Algorithms, scripts and software: All the relevant algorithms, scripts and software tools driving the project will be described in manuscripts and/or on GitHub (https://github.com).
- Data that do not support publication will be either deposited in an open access repository or made available upon request by email. Data will be reused by transfer via Belnet Filesender or secure copy.

When will the data be made available?

All research outputs will be made openly accessible, at the latest, at the time of publication (or preprint deposition). No embargo will be foreseen unless imposed e.g. by pending publications.

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

Data usage licences will be discussed with involved parties before any licences are granted.

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

Depending on the data repository and the type of data that would be made available, a unique identifier will be added to the data set.

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

If shipment of data or material is required by an other study group abroad, after approval the costs of drafting of MTA/DTA and shipment itself will be covered by the requesting party.

6. Responsibilities

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

Dieter Smout and Pieter Evenepoel

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

Dieter Smout and Pieter Evenepoel

Who will manage data preservation and sharing?

Pieter Evenepoel

Who will update and implement this DMP?

Dieter Smout and Pieter Evenepoel

Interorgan Crosstalk in CKD-MBD: On the road to Precision Medicine 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 current project will reuse existing pseudonymized clinical data from available databases including clinical information, laboratory results, pathology results and radiographic results.
The current project will create new data from biosamples stored in our biobank, including microRNA, mRNA, protein and metabolite levels in blood, urine, bone and bone marrow.
During the rat experiments, new data will be generated with laboratory tests.
During the project, we will generate new data based on existing data by using analyses with R.

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 (If already designated, please fill in his/her name.) Dieter Smout (during the project) and Prof Pieter Evenepoel (during and after the project)
2. Storage capacity/repository
    - during the research: All the electronic data, including raw data, excel files, figures, tables and R scripts, will be stored on the internal UZ Leuven OneDrive network. The used biosamples are stored in the Leuven Renal Research Biobank.
    - after the research: All the electronic data will remain stored in an archive folder the internal UZ Leuven network for at least 5 years. The used biosamples remain stored in the Leuven Renal Research Biobank.

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)

I do not wish to deviate.

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)

Current project only reuses personal data and does not collect new personal data. The newly created data is low-sensitive and has a low risk of misuse. Nonetheless, this newly ceated data will be handled and stored by the same principles as the primary data, i.e. stored on the UZ Leuven Network with restricted permission to members of the laboratory of nephrology.

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

NA

Interorgan Crosstalk in CKD-MBD: On the road to Precision Medicine
DPIA

DPIA

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

- Not applicable

Interorgan Crosstalk in CKD-MBD: On the road to Precision Medicine GDPR

GDPR

Have you registered personal data processing activities for this project?

- Not applicable