DMP title

Project Name My plan (FWO DMP) - DMP title
Project Identifier 12Z0622N
Grant Title 12Z0622N
Principal Investigator / Researcher Katleen Martens
Institution KU Leuven

1. General Information

Name applicant

Katleen Martens

FWO Project Number & Title

12Z0622N, The bittersweet taste of bacteria in the nose of patients with chronic rhinosinusitis with nasal polyps

Affiliation

- KU Leuven
- Universiteit Antwerpen

Co-promotor of this project, Prof Dr. S. Lebeer, is located at the University of Antwerp and part of the project will be performed at the University of Antwerp.

2. Data description

Will you generate/collect new data and/or make use of existing data?

- Generate new data

Describe in detail the origin, type and format of the data (per dataset) and its (estimated) volume. This may be easiest in a table (see example) or as a data flow and per WP or objective of the project. If you reuse existing data, specify the source of these data. Distinguish data types (the kind of content) from data formats (the technical format).

Type of data	Format	Volume	How created
Shotgun sequencing data	.txv; .rds; .csv in R: rmd file	300-900GB	Next generation sequencing on bacterial DNA isolated from nasal swabs
Microscopy images	.tif	100GB	Immunofluorescent microscopic images of nasal epithelial cells
Qualitative data (RT-qPCR)	Gene expression profiles of nasal tissue stored transcribed into MS excel format	2GB	Gene expression profiles obtained from human and mouse tissue with CFX connect (BioRad)
Qualitative data (Nitrite/Nitrate Assay kit)	Protein expression levels of Nitrite/Nitrate in supernatant of stimulated nasal epithelial cells transcribed into MS excel format	2GB	Data created with classical ELISA reader
Qualitative data (calcium imaging)	.tiff	100GB	The intercellular calcium concentrations will be monitored using the ratio of fluorescence of Fura-2 measured on the alternating illumination at 340 and 380 nm using an MT-10 illumination system and CellM software (Olympus).
Qualitative data (ciliary beat frequency)	.tiff	100GB	a high-speed digital imaging system that detects the frequency of variation in light intensity of the image because of repetitive motion of the cilia.
Qualitative data ((LC)-MS/MS analysis)	.bcf; .rtm; .raw; .rtx	100GB	Fractionation/purification of bacterial ligands

3. Legal and ethical issues

Will you use personal data? If so, shortly describe the kind of personal data you will use. Add the reference to your file in KU Leuven's Register of Data Processing for Research and Public Service Purposes (PRET application). Be aware that registering the fact that you process personal data is a legal obligation.

- Yes

Privacy Registry Reference: Registration will be done by filling in the GDPR questionnaire integrated in the application for submission to the Clinical Trial center of the University Hospitals

Leuven.

Short description of the kind of personal data that will be used:

1/ Personal information for research purposes, consisting of socio-demographical data (e.g. gender, date of birth, smoker/non-smoker) and data concerning medical status (e.g. disease severity, treatments and/or medication intake). This is needed to define the disease phenotypes to investigate the interaction between the microbiome and taste receptor function.

Are there any ethical issues concerning the creation and/or use of the data (e.g. experiments on humans or animals, dual use)? If so, add the reference to the formal approval by the relevant ethical review committee(s)

- Yes

Nasal swabs will be taken from patients with chronic rhinosinusitis with nasal polyps and healthy controls. Ethical approval has been obtained (S65918).

Does your work possibly result in research data with potential for tech transfer and valorisation? Will IP restrictions be claimed for the data you created? If so, for what data and which restrictions will be asserted?

- No

Do existing 3rd party agreements restrict dissemination or exploitation of the data you (re)use? If so, to what data do they relate and what restrictions are in place?

- No

4. Documentation and metadata

What documentation will be provided to enable reuse of the data collected/generated in this project?

1/ Microscopy images the following information will be noted: dimensions, image type, bit-depth, pixel sizes and microscope settings. The methodology and protocol will be described in detail in the lab book. The original images togheter with a ReadMe file are stored on OneDrive.

2/ Raw experimental data will be collected per experiment, including a txt file with a clear description of what the data represent and how they were generated. The input files used for the stimulation will be kept in the same folder.

3/ For the shot gun sequencing, a ReadMe file will be created for the different analysis. All steps are written down in the lab book.

Will a metadata standard be used? If so, describe in detail which standard will be used. If no, state in detail which metadata will be created to make the data easy/easier to find and reuse.

- No

5. Data storage and backup during the FWO project

Where will the data be stored?

1/ We will collect patient samples and thus patient information (i.e. two nasal swaps from the anterior nares, residual human tissue, questionnaire, etc).

All this information is coded. Each patient receives a unique code in order to anonymize the collected information. The questionnaires, that collect information on demographics, previous history and symptoms are also blinded for each patient with the same unique patient code.

2/ Shotgun sequencing data of the nasal swaps of the patients will be collected and will be reused by our research group. A ReadMe file will be created to help with the reuse of this data.

3/ Our results will be published in different manuscripts. This data will be blinded in order to protect patient information.

How is backup of the data provided?

All the collected data and results will be stored on the university's central servers (K:\GBW-0041_KlinImm and J:\GBW-0057_Labo Klinische Immunologie) with automatic daily back-up procedures and is only accessible for registered KU Leuven personnel. In addition, all patient information is blinded and coded. The key for unblinding is stored on a different secured account

(i.e. the account of the PI).

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

OneDrive: Standard every user has 2 TB. This capacity can be extended to 5 TB without costs.
K:-drive: is an easily scalable system with blocks of 5 TB. Over 100 TB, requests can be done per block of 10 TB.

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

Large volume storage (K:-drive) can be purchased in blocks of 5 TB and from 100 TB on in blocks of 10 TB. 113,84 euro is the price per TB per year. For large volume storage it is standard procedure to make a disaster recovery (mirror) copy to have a second copy available. For this an additional cost of 51,30 euro is charged. These costs are covered by the laboratory.

Data security: how will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?

Access to the J drive is only possible for registered KU Leuven personnel. Each coworker at the KU Leuven has his/her personal ID & password allowing access to certain folders only. Especially folders containing keys to patient ID codes, will be retrievable with secure restricted access only. The Ethical Committee of UZ Leuven is particularly inquiring on this topic before approving a clinical study and carefully checks this procedure, according to the national and EU legislation. Pseudonymised data will be accessible to coworkers of the Allergy & Clinical Immunology Research Group only. Access to the folders can furthermore be checked.

6. Data preservation after the FWO project

Which data will be retained for the expected 5 year period after the end of the project? In case only a selection of the data can/will be preserved, clearly state the reasons for this (legal or contractual restrictions, physical preservation issues, ...).

All data will be retained during 5 years following the end of the project. Pseudonomized data and physical data will also be stored for longer than 5 years. Permission is requested in the informed consent (IC) documents. Legal obligations will be discussed at different moments with the legal department of the clinical Trial Center (CTC) and the Ethical Committee Research UZ/KU Leuven. If any change is needed, we will inform participants by an adapted IC document and ask for eventual prolongation of the preservation of the data. Preservation of biomaterial is done in close collaboration with the UZ Leuven Biobank and will also be done at least for 5 years, but longer upon IC.

Where will the data be archived (= stored for the longer term)?

The data of the shotgun sequencing will be temporary stored on the local server of the Univeristy of Antwerp and locally be destroyed once admitted to the KU Leuven where it will be stored on the KU Leuven central servers (with automatic back-up procedures) for at least 5 years, conform the KU Leuven RDM policy. Data exchange will be done be by the Belnet filesharing tool Filesender only. Large files will be stored on the K-disc devoted for long term storage (see above). The investigators will work with staff of the KU Leuven Libraries to determine what to archive.

What are the expected costs for data preservation during the retention period of 5 years? How will the costs be covered?

The expected costs for the network K-drive are 11,38 euro/100GB/year and will be covered for 50% by the Group Biomedical Sciences and for 50% by the lab . If storage of extremely large files is required, then the use of the L-drive will be discussed with the PI and IT department, as it is currently not in use at the laboratory of Allergy and Clinical Immunology. The cost price for the L-drive (5TB) is 569,2 euro/year.

7. Data sharing and reuse

Are there any factors restricting or preventing the sharing of (some of) the data (e.g. as defined in an agreement with a 3rd party, legal restrictions)?

- No

Which data will be made available after the end of the project?
I am keen to present the results of this project to peers and to a wide audience. All research outputs supporting publications will be made openly accessible. Depending on their nature, some data will be anonemyzed prior to publication, either on an individual basis to interested researchers and/or potential new collaborators, or publicly via respositories.

Where/how will the data be made available for reuse?

- In an Open Access repository
- In a restricted access repository
- Upon request by mail

As soon as the manuscript is publicaly available, the repository will be changed to a public repository. Researcher can ask permission to reuse the data upon reasonable request via email.

When will the data be made available?

- Upon publication of the research results

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.

Who will be able to access the data and under what conditions?
Data will be shared between the PI and co-PI and within the research unit but always respecting the pseudonymisation. Sequence data will be uploaded in an open access repository and shared upon request. Only resuse for research purposes will be allowed and commercial reuse will be excluded. Data without sharing restrictions will be shared through peer reviewed publications.

What are the expected costs for data sharing? How will the costs be covered?
No costs for data sharing are expected

8. Responsibilities

Who will be responsible for data documentation & metadata?
The PI and co-PI (P. Hellings and S. Lebeer)

Who will be responsible for data storage & back up during the project?
In every lab of the PI and co-PI, a person will be dedicated to be responsible for the data storage.

Who will be responsible for ensuring data preservation and reuse ?
The PI and co-PI.

Who bears the end responsibility for updating & implementing this DMP?
The PI Peter Hellings bears the end responsibility of updating & implementing this DMP.