1S06023N PREVENTION AND TREATMENT OF CHRONIC DISEASES BY FOOD THAT IS GOOD FOR YOU

A Data Management Plan created using DMPonline.be

Creators: Judith Wellens, First Name Surname, 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: First Name Surname, n.n. n.n.

Data Manager: Judith Wellens

Project Administrator: Judith Wellens

Grant number / **URL:** ZKE3104 - SB/1S06023N Wellens Judith

ID: 192471

Start date: 01-11-2022

End date: 26-12-2024

Project abstract:

In this project, the cooking methods and role of emulsifiers on intestinal and systemic inflammation, the gut barrier and gut microbiota will be investigated.

The first work package on cooking methods will consist of the PreSTEAMM study and the STEAMM study. The PreSTEAMM study assesses whether advanced glycation end products (AGEs) can be measured in the blood by ELISA after a single meal challenge with a meal high vs. low in AGEs in 5 healthy volunteers. The STEAMM study is a double-blind randomised cross-over trial in 20 healthy volunteers (2x2 weeks of intervention) wherein a fixed meal plan that is high in AGEs will be compared to one that is low in AGEs. Measured end points include anthropometric measurements, regular blood work, OLINK proteomics, urine lactulose mannitol testing, faecal calprotectin and the gut microbiota.

The second work package on the impact of emulsifiers on intestinal and general health contains an in vitro part and a clinical in vivo part. In vitro I will use a Caco2/HT29 MTX co-culture cell model and afterwards human-derived organoids of the colon from healthy volunteers and inflammatory bowel disease patients to test the effect of various emulsifiers on intestinal integrity and inflammation.

The clinical trial, the FOAM trial, is a placebo-controlled randomized trial in 60 healthy volunteers during six weeks in which 5 emulsifiers are tested using baked goods on top of an emulsifier-free diet. Measured end points include anthropometric measurements, regular blood work, OLINK proteomics, urine lactulose mannitol testing, faecal calprotectin and the gut microbiota.

Last modified: 10-01-2023

1S06023N PREVENTION AND TREATMENT OF CHRONIC DISEASES BY FOOD THAT IS GOOD FOR YOU

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
FOAM in vitro - microscopy	Confocal microscopy of cell cultures and organoids	☒Generate new data ☐ Reuse existing data	☒Digital ☐Physical	☐Observational ☒Experimental ☐ Compiled/ aggregated data ☐ Simulation data ☐ Software ☐ Other ☐ NA	☐ .por ☐ .xml ☐ .tab ☐ .csv ☐ .pdf ☐ .txt ☐ .rtf ☐ .dwg ☐ .tab ☐ .gml ☒ other: .tif ☐ NA	☐ < 100 MB ☐ < 1 GB ☒ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	None
FOAM in vitro – sequencing	Bulk RNA sequencing and qPCR of RNA transcripts for both cell cultures and organoids	☒Generate new data ☐ Reuse existing data	☒Digital ☒Physical	☐Observational ☒Experimental ☐ Compiled/ aggregated data ☐ Simulation data ☐ Software ☐ Other ☐ NA	☐ .por ☐ .xml ☐ .tab ☐ .csv ☐ .pdf ☐ .txt ☐ .rtf ☐ .dwg ☐ .tab ☐ .gml ☒ other: .count, .tif, .xls ☐ NA	☐ < 100 MB ☐ < 1 GB ☒ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	RNA will be stored in the biobank at TARGID at -20°C, this will take approximately 5 boxes.
FOAM in vitro – transepithelial electrical resistance (TEER)	TEER measurements for barrier integrity in cell cultures and organoids	☒Generate new data ☐ Reuse existing data	☒Digital ☐Physical	☐Observational ☒Experimental ☐ Compiled/ aggregated data ☐ Simulation data ☐ Software ☐ Other ☐ NA	☐ .por ☐ .xml ☐ .tab ☒ .csv ☐ .pdf ☐ .txt ☐ .rtf ☐ .dwg ☐ .tab ☐ .gml ☒ other: .xls ☐ NA	☒ < 100 MB ☐ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	

FOAM in vitro – cytokine measurement in supernatant of cell cultures and organoids	MSD inflammatory cytokine panel	☒Generate new data ☐ Reuse existing data	☒Digital ☒Physical	☐Observational ☒Experimental ☐ Compiled/ aggregated data ☐ Simulation data ☐ Software ☐ Other ☐ NA	☐ .por ☐ .xml ☐ .tab ☒ .csv ☐ .pdf ☐ .txt ☐ .rtf ☐ .dwg ☐ .tab ☐ .gml ☒ other: .xls ☐ NA	☒ < 100 MB ☐ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	Supernatantwill be stored in the biobank at TARGID at -20°C, this will take approximately 5 boxes.
FOAM in vivo – survey data	Survey data collected in REDCap	☒Generate new data ☐ Reuse existing data	☒Digital ☐Physical	☐Observational ☒Experimental ☐ Compiled/ aggregated data ☐ Simulation data ☐ Software ☐ Other ☐ NA	☐ .por ☐ .xml ☐ .tab ☒ .csv ☐ .pdf ☐ .txt ☐ .rtf ☐ .dwg ☐ .tab ☐ .gml ☐ other: ☐ NA	☒ < 100 MB ☐ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	
FOAM in vivo – blood	60 volunteers x 4 timepoints for DNA extraction and serum	☐Generate new data ☐ Reuse existing data	☐Digital ☒Physical	☐Observational ☒Experimental ☐ Compiled/ aggregated data ☐ Simulation data ☐ Software ☐ Other ☐ NA	☐ .por ☐ .xml ☐ .tab ☐ .csv ☐ .pdf ☐ .txt ☐ .rtf ☐ .dwg ☐ .tab ☐ .gml ☐ other: ☐ NA	☐ < 100 MB ☐ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	Serum and DNA fraction stored at biobank TARGID – space already ascertained
FOAM in vivo – blood DNA analysis	MIP - sequencing	☒Generate new data ☐ Reuse existing data	☒Digital ☐Physical	☐Observational ☒Experimental ☐ Compiled/ aggregated data ☐ Simulation data ☐ Software ☐ Other ☐ NA	☐ .por ☐ .xml ☐ .tab ☒ .csv ☐ .pdf ☐ .txt ☐ .rtf ☐ .dwg ☐ .tab ☐ .gml ☒other:.xls ☐ NA	☐ < 100 MB ☐ < 1 GB ☒ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	
FOAM in vivo – serum OLINK cytokine panel	OLINK inflammatory and metabolic panel	☒Generate new data ☐ Reuse existing data	☒Digital ☐Physical	☐Observational ☒Experimental ☐ Compiled/ aggregated data ☐ Simulation data ☐ Software ☐ Other ☐ NA	☐ .por ☐ .xml ☐ .tab ☒ .csv ☐ .pdf ☐ .txt ☐ .rtf ☐ .dwg ☐ .tab ☐ .gml ☒other:.xls ☐ NA	☐ < 100 MB ☐ < 1 GB ☒ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	

FOAM in vivo – blood analysis by the LAG (UZ Leuven laboratory)	Routine blood work	☒Generate new data ☐ Reuse existing data	☒Digital ☐Physical	☐Observational ☒Experimental ☐ Compiled/ aggregated data ☐ Simulation data ☐ Software ☐ Other ☐ NA	☐ .por ☐ .xml ☐ .tab ☒ .csv ☐ .pdf ☐ .txt ☐ .rtf ☐ .dwg ☐ .tab ☐ .gml ☒other:.xls ☐ NA	☐ < 100 MB ☐ < 1 GB ☒ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	
FOAM in vivo - urine	60 volunteers x 3 timepoints	☒Generate new data ☐ Reuse existing data	☐Digital ☒Physical	☐Observational ☒Experimental ☐ Compiled/ aggregated data ☐ Simulation data ☐ Software ☐ Other ☐ NA	☐ .por ☐ .xml ☐ .tab ☐ .csv ☐ .pdf ☐ .txt ☐ .rtf ☐ .dwg ☐ .tab ☐ .gml ☐ other: .tif ☐ NA	☐ < 100 MB ☐ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	Stored at the biobank at TARGID – space already ascertained
FOAM in vivo - urine	60 volunteers x 3 timepoints for lactulose mannitol testing	☒Generate new data ☐ Reuse existing data	☒Digital ☐Physical	☐Observational ☒Experimental ☐ Compiled/ aggregated data ☐ Simulation data ☐ Software ☐ Other ☐ NA	☐ .por ☐ .xml ☐ .tab ☒ .csv ☐ .pdf ☐ .txt ☐ .rtf ☐ .dwg ☐ .tab ☐ .gml ☒ other: .xls ☐ NA	☒ < 100 MB ☐ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	Stored at the biobank at TARGID – space already ascertained
FOAM in vivo faeces	60 volunteers x 4 timepoints	☒Generate new data ☐ Reuse existing data	☐Digital ☒Physical	☐Observational ☒Experimental ☐ Compiled/ aggregated data ☐ Simulation data ☐ Software ☐ Other ☐ NA	☐ .por ☐ .xml ☐ .tab ☐ .csv ☐ .pdf ☐ .txt ☐ .rtf ☐ .dwg ☐ .tab ☐ .gml ☐ other: .xls ☐ NA	☐ < 100 MB ☐ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	Stored at the biobank at TARGID – space already ascertained

FOAM in vivo Faeces - microbiota	60 volunteers x 4 timepoints	☒Generate new data ☐ Reuse existing data	☒Digital ☐Physical	☐Observational ☒Experimental ☐ Compiled/ aggregated data ☐ Simulation data ☐ Software ☐ Other ☐ NA	☐ .por ☐ .xml ☐ .tab ☐ .csv ☐ .pdf ☐ .txt ☐ .rtf ☐ .dwg ☐ .tab ☐ .gml ☒ other: .fastq ☐ NA	☐ < 100 MB ☐ < 1 GB ☒ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	
PreSTEAMM - faeces	2x 5 samples	☒Generate new data ☐ Reuse existing data	☐Digital ☒Physical	☐Observational ☒Experimental ☐ Compiled/ aggregated data ☐ Simulation data ☐ Software ☐ Other ☐ NA	☐ .por ☐ .xml ☐ .tab ☐ .csv ☐ .pdf ☐ .txt ☐ .rtf ☐ .dwg ☐ .tab ☐ .gml ☐ other: .tif ☐ NA	☐ < 100 MB ☐ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	Stored at the biobank at TARGID – space already ascertained
PreSTEAMM – blood	Serum for AGE ELISAs	☒Generate new data ☐ Reuse existing data	☒Digital ☐Physical	☐Observational ☒Experimental ☐ Compiled/ aggregated data ☐ Simulation data ☐ Software ☐ Other ☐ NA	☐ .por ☐ .xml ☐ .tab ☒ .csv ☐ .pdf ☐ .txt ☐ .rtf ☐ .dwg ☐ .tab ☐ .gml ☒ other: .xls ☐ NA	☒ < 100 MB ☐ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	
PreSTEAMM – blood	Serum for AGE ELISAs	☒Generate new data ☐ Reuse existing data	☐Digital ☒Physical	☐Observational ☒Experimental ☐ Compiled/ aggregated data ☐ Simulation data ☐ Software ☐ Other ☐ NA	☐ .por ☐ .xml ☐ .tab ☐ .csv ☐ .pdf ☐ .txt ☐ .rtf ☐ .dwg ☐ .tab ☐ .gml ☐ other: .tif ☐ NA	☐ < 100 MB ☐ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	Stored at the biobank at TARGID – space already ascertained

STEAMM – survey data	Survey data collected in REDCap	☒Generate new data ☐ Reuse existing data	☒Digital ☐Physical	☐Observational ☒Experimental ☐ Compiled/ aggregated data ☐ Simulation data ☐ Software ☐ Other ☐ NA	☐ .por ☐ .xml ☐ .tab ☒ .csv ☐ .pdf ☐ .txt ☐ .rtf ☐ .dwg ☐ .tab ☐ .gml ☐ other: ☐ NA	☒ < 100 MB ☐ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	
STEAMM in vivo – blood	20 volunteers x 3 timepoints for DNA extraction and serum	☐Generate new data ☐ Reuse existing data	☐Digital ☒Physical	☐Observational ☒Experimental ☐ Compiled/ aggregated data ☐ Simulation data ☐ Software ☐ Other ☐ NA	☐ .por ☐ .xml ☐ .tab ☐ .csv ☐ .pdf ☐ .txt ☐ .rtf ☐ .dwg ☐ .tab ☐ .gml ☐ other: ☐ NA	☐ < 100 MB ☐ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	Serum and DNA fraction stored at biobank TARGID – space already ascertained
STEAMM – blood DNA analysis	MIP - sequencing	☒Generate new data ☐ Reuse existing data	☒Digital ☐Physical	☐Observational ☒Experimental ☐ Compiled/ aggregated data ☐ Simulation data ☐ Software ☐ Other ☐ NA	☐ .por ☐ .xml ☐ .tab ☒ .csv ☐ .pdf ☐ .txt ☐ .rtf ☐ .dwg ☐ .tab ☐ .gml ☒other:.xls ☐ NA	☐ < 100 MB ☐ < 1 GB ☒ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	
STEAMM – serum OLINK cytokine panel	OLINK inflammatory and metabolic panel	☒Generate new data ☐ Reuse existing data	☒Digital ☐Physical	☐Observational ☒Experimental ☐ Compiled/ aggregated data ☐ Simulation data ☐ Software ☐ Other ☐ NA	☐ .por ☐ .xml ☐ .tab ☒ .csv ☐ .pdf ☐ .txt ☐ .rtf ☐ .dwg ☐ .tab ☐ .gml ☒other:.xls ☐ NA	☐ < 100 MB ☐ < 1 GB ☒ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	
STEAMM– blood analysis by the LAG (UZ Leuven laboratory)	Routine blood work	☒Generate new data ☐ Reuse existing data	☒Digital ☐Physical	☐Observational ☒Experimental ☐ Compiled/ aggregated data ☐ Simulation data ☐ Software ☐ Other ☐ NA	☐ .por ☐ .xml ☐ .tab ☒ .csv ☐ .pdf ☐ .txt ☐ .rtf ☐ .dwg ☐ .tab ☐ .gml ☒other:.xls ☐ NA	☐ < 100 MB ☐ < 1 GB ☒ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	

STEAMM - urine	20 volunteers x 3 timepoints	☒Generate new data ☐ Reuse existing data	☐Digital ☒Physical	☐Observational ☒Experimental ☐ Compiled/ aggregated data ☐ Simulation data ☐ Software ☐ Other ☐ NA	☐ .por ☐ .xml ☐ .tab ☐ .csv ☐ .pdf ☐ .txt ☐ .rtf ☐ .dwg ☐ .tab ☐ .gml ☐ other: .tif ☐ NA	☐ < 100 MB ☐ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	Stored at the biobank at TARGID – space already ascertained
STEAMM urine	20 volunteers x 3 timepoints for lactulose mannitol testing	☒Generate new data ☐ Reuse existing data	☒Digital ☐Physical	☐Observational ☒Experimental ☐ Compiled/ aggregated data ☐ Simulation data ☐ Software ☐ Other ☐ NA	☐ .por ☐ .xml ☐ .tab ☒ .csv ☐ .pdf ☐ .txt ☐ .rtf ☐ .dwg ☐ .tab ☐ .gml ☒ other: .xls ☐ NA	☒ < 100 MB ☐ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	Stored at the biobank at TARGID – space already ascertained
STEAMM faeces	20 volunteers x 3 timepoints	☒Generate new data ☐ Reuse existing data	☐Digital ☒Physical	☐Observational ☒Experimental ☐ Compiled/ aggregated data ☐ Simulation data ☐ Software ☐ Other ☐ NA	☐ .por ☐ .xml ☐ .tab ☐ .csv ☐ .pdf ☐ .txt ☐ .rtf ☐ .dwg ☐ .tab ☐ .gml ☐ other: .xls ☐ NA	☐ < 100 MB ☐ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	Stored at the biobank at TARGID – space already ascertained
STEAMM Faeces - microbiota	20 volunteers x 3 timepoints	☒Generate new data ☐ Reuse existing data	☒Digital ☐Physical	☐Observational ☒Experimental ☐ Compiled/ aggregated data ☐ Simulation data ☐ Software ☐ Other ☐ NA	☐ .por ☐ .xml ☐ .tab ☐ .csv ☐ .pdf ☐ .txt ☐ .rtf ☐ .dwg ☐ .tab ☐ .gml ☒ other: .fastq ☐ NA	☐ < 100 MB ☐ < 1 GB ☒ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	

STEAMM – blood	Serum for AGE ELISAs	☒Generate new data ☐ Reuse existing data	☒Digital ☐Physical	☐Observational ☒Experimental ☐ Compiled/ aggregated data ☐ Simulation data ☐ Software ☐ Other ☐ NA	☐ .por ☐ .xml ☐ .tab ☒ .csv ☐ .pdf ☐ .txt ☐ .rtf ☐ .dwg ☐ .tab ☐ .gml ☒ other: .xls ☐ NA	☒ < 100 MB ☐ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	
STEAMM – blood	Serum for AGE ELISAs	☒Generate new data ☐ Reuse existing data	☐Digital ☒Physical	☐Observational ☒Experimental ☐ Compiled/ aggregated data ☐ Simulation data ☐ Software ☐ Other ☐ NA	☐ .por ☐ .xml ☐ .tab ☐ .csv ☐ .pdf ☐ .txt ☐ .rtf ☐ .dwg ☐ .tab ☐ .gml ☐ other: .tif ☐ NA	☐ < 100 MB ☐ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	Stored at the biobank at TARGID – space already ascertained

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:

No existing data will be reused for this project.

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

For the human trials (and use of the human-derived organoids from colonic biopsy material), approval from the ethical committee is mandatory. For all projects, this has already been granted:
STEAMM: S65600 - approved 17/12/2021
PRET: G-2021-3896
PreSTEAMM: S66021 - approved 20/12/2021
PRET: G-2021-4168
FOAM: S66308
Pret: G-2022-4740 – approved 13/07/2022
VLECC/biopsy material for organoid research: S53684

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

The personal data of 85 healthy volunteers will be collected. This includes 'regular' personal data such as age, weight, sex, and email addresses, but also 'special' or 'sensitive' personal data such as genetic information, ethnicity, health and biometric data.
STEAMM: S65600 - approved 17/12/2021
PRET: G-2021-3896
PreSTEAMM: S66021 - approved 20/12/2021
PRET: G-2021-4168
FOAM: S66308
Pret: G-2022-4740 – approved 13/07/2022
VLECC/biopsy material for organoid research: S53684

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

After finishing the project and before dissemination of the results, we will look carefully if any valorisable material is present. Additional IP discussions might start from that point on. Regarding planned collaborations, IP has been discussed already:
- For the RAES lab (who will perform the the microbiota analysis), IP will be shared.
- For the Olink cytokine panels: they will perform a service.
All of this has been discussed with LRD already and the necessary contracts have been drafted and signed by all parties as well. Of course, after finishing the project, LRD will be contacted again to rediscuss any potential for research valorisation at that stage.

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).

The methodology and protocol will be described in detail in the physical lab book that will be stored at the lab at all times. An accompanying key file (.xls) to decipher which result files match which protocol (which will reference the pages in the lab book) will also be available on a shared drive. Only members from the team will have access to these folders.
Completely analogue for clinical studies, the protocol, ethical documents, and results will also be stored in a shared drive. An accompanying file (.xls) will also state which folders contain which information (eg. Survey data from Redcap in a file named xxx, with the extension .csv.)
Informed consents (physical data) will be safely stored behind lock and key by the primary investigator (João Sabino), only he will have access to these documents.

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

Metadata standards are typically not used within our lab group. We do have a minimal set of requirements that will be followed in order to ensure standardization and possibility to reinterpret and reuse the data when necessary and permitted.
All collected data will be labelled with (1) title, (2) author, (3) data type, (4) data created and date modified, (5) file size, (6) equipment reference (such as manufacturer and model identification). Depending on the nature of data additional metadata are collected.

- Sequencing/genotyping:

All experimental data will be noted every day in the lab book including all detailed experimental data. Once the sequencing data will be available, all files and folders will be labeled in a clearly structured way. The explanation of the labeling and the performed analysis will

also be written down in the lab book.
For sequencing, sequencing depth and analysis cut-offs will also be noted.

- Microscopy:

The labeling of every individual picture taken will contain the patient culture number, the passage number, the experimental stimulation and the magnification.

- Electrical resistance (TEER)

Every measurement will be noted every day and contain all experimental information (patient culture number, passage number, experimental stimulation). Analysis will be performed in excel files. The excel file will also contain all experimental data and refer to the according pages in the labbook.

- Biobank

All patient information (age, gender, disease) will be registered in an anonymized way in the file containing all collected samples. Every patient will receive an identification nummer which can only be decoded by the responsible data manager.
ELISA: all information on the samples (plate layouts that indicate which samples will be used) and the used kits will be registered in the lab book and electronically in .xls (which will go into the shared drive). The protocol will be stored in the lab book as well. The data analysis will be done in R for which the script (that also specifies every step) will be made available in the shared folder.

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

Where will the data be stored?

All experimental data will be written down in a physical lab book with the chronological reporting of all related experiments and results including a cross reference to electronic storage of data. (2) All data will be stored electronically on the personal KU Leuven One Drive for daily back-up. Larger and RAW data files (RNA sequencing) will be stored on the KU Leuven K drive. Only members of the lab group that have permission to access the files will have the ICT permission to do so. (e.g. students won't be granted access to sensitive data).
The biological samples (blood, faeces, urine) will be stored in our biobank for potential later use or repeating of certain test. The necessary storage room for the expected amount of biological material is already ascertained at the biobank.

How will the data be backed up?

Backup is secured daily on central servers of the university. All data will be stored electronically on the personal KU Leuven One Drive, larger and RAW data files (RNA sequencing) will be stored on the KU Leuven K drive.

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

The OneDrive has a limitation in storage capacity (2TB) which will be sufficient for this project. Larger data files can in addition be stored on the K-drive.

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

Since we will be working with sensitive personal data that will only be fully anonymized at the end of the project, the data will be stored in the university's secure environment for private data. Checks and balances are currently in place to ensure that only authorized personnel have access to sensitive data.
For example: sensitive physical documents will be kept by the PI behind lock and key (to which only he has access to), and students won't have access to the electronically sensitive files on the OneDrive. Regarding REDCap, the PI and myself will monitor access to the platform.

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

The OneDrive (2TB) comes without charge, and should be enough for completion of this project. If needed TARGID will provide for an additional budget for data management. Lastly, the benchfee could be used to pay for extra storage capacity.

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 TARGID. The informed consent includes a clause that permits later reuse of the obtained data. The cell cultures/organoids obtained will not be stored since this is biologically not possible. Data obtained through the experiments and coupes will be stored.

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

The data will be stored on the university's central servers – (OneDrive with automatic back-up procedures) for at least 10 years, conform the KU Leuven RDM policy.

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

The OneDrive (2TB) comes without charge, and should be enough for completion of this project. If needed TARGID will provide for an additional budget for data management. Lastly, the benchfee could be used to pay for extra storage capacity.

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, …)

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. João Sabino). Non-published data will remain confidential until a final decision on publication of the data has been taken.

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

Data could be reused by other members of the TARGID team, after consultation and approval of the head of our lab group and/or the head of TARGID (currently Séverine Vermeire and Kristin Verbeke respectively).

Data can possibly be accessed by a third party after signing a data sharing agreement and approval of the head of TARGID. An appropriate DTA and/or MTA will be in place. Costs for shipment are to be covered by the requesting party. Access will be considered after a request is submitted explaining the planned reuse. Only uses for research purposes will be allowed and commercial reuse will be excluded. Exceptions are to be submitted to the head of our lab group and/or TARGID (Séverine Vermeire and Kristin Verbeke).

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

Regarding the microbiota results, the IP is shared with the RAES lab and dissemination of the results should be discussed with them first for this particular part of the project. However, this has already been discussed with LRD and we will keep in touch for this aspect of the project.

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

A specific repository will be chosen after the publication strategy is known as some journal request specific repositories.

When will the data be made available?

Only after publication of the research results in a peer-reviewed journal.

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

Data usage licences will be discussed with LRD before any licences are granted. Similarly, when DTAs or MTAs are discussed, this will always be after consulting and collaborating with LRD.

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?

Judith Wellens and João Sabino

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

Judith Wellens and João Sabino

Who will manage data preservation and sharing?

João Sabino

Who will update and implement this DMP?

The PI (João Sabino) bears the end responsibility of updating & implementing this DMP.