The impact of the administration mode on the physiological effects of SCFA.

A Data Management Plan created using DMPonline.be

Creator: Riet Rosseel

Affiliation: KU Leuven (KUL)

Funder: Fonds voor Wetenschappelijk Onderzoek - Research Foundation Flanders (FWO)

Template: FWO DMP (Flemish Standard DMP)

Grant number / URL: 1SHGN24N

ID: 206257

Start date: 01-11-2023

End date: 31-10-2027

Project abstract:

The research project focuses on short-chain fatty acids (SCFA) which are the main metabolites produced during the bacterial fermentation of indigestible fibers in the human colon. SCFA are proposed as potential mediators of the health benefits associated with fiber intake as they affect the physiology either locally in the gut or via the systemic circulation. Furthermore, SCFA are present in plant-based fermented foods which are also linked with health promoting effects. However, while SCFA are absorbed in the colon upon fiber consumption, SCFA are absorbed in the small intestine upon consuming fermented foods. The different site of absorption might affect the health effects of SCFA.

This research project consists of 3 human intervention studies (referred to as work packages, WP) investigating how the site of SCFA absorption might influence their physiological effects/ health benefits.

In WP1, SCFA will be administered either into the small intestine or in the colon of healthy individuals and the systemic availability of the SCFA will be quantified using stable isotope technology. WP 2 will evaluate to what extent targeted delivery of SCFA affects the release of gut hormones GLP-1 and PYY.- Both WP1 and 2 involve healthy volunteers attending 2 study visits during which SCFA are orally administered (acutely) in the colon or the small intestine using dedicated capsules and blood samples and questionnaires are collected.

In WP3, the results from previous work packages will be further investigated in a parallel, nutritional intervention study. Participants at risk for metabolic syndrome will supplement their diets with either a fiber-rich, fermented or a control food product for 12 weeks. The impact of the supplementation on cardiometabolic risk factors will be assessed, with the postprandial glucose response as primary outcome. Blood and fecal samples will be collected before and after the intervention period during study visits.

Last modified: 10-04-2024

The impact of the administration mode on the physiological effects of SCFA.
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, .csv,.pdf, .txt, .rtf, .dwg, .gml, … • NA	*Please choose from the following options:* • <100MB • <1GB • <100GB • <1TB • <5TB • <10TB • <50TB • >50TB • NA	
Demographic data	Demographic data of the participants of the three WPs (acronym, age, height, weight, BMI, medication use, medical history, alcohol and tobacco use)	Generate new data	Digital	Experimental	.xlsx; .R	< 100 MB	
Code file	The file contains the name, email address and bank account of the participants linked with the assigned pseudonym.	Generate new data	Digital	Experimental	.xlsx	< 100 MB	
Informed consents	Informed consent forms of the participants of the three WPs.	Generate new data	Physical	Experimental			~10 pages for each informed consent form

Parameters screening	During the screening visit, some parameters are assessed to evaluate eligibility. For WP 1 and 2, the hemoglobin level is assessed in blood sample taken at screening. For WP 3, cardiometabolic risk factors are assessed in a blood sample including LDL cholesterol, HDL cholesterol, postprandial glucose and triglyceride levels. Also blood pressure and waist circumference are determined to evaluate the presence of hypertension or central obesity.	Generate new data	Digital	Experimental	.xlsx; .R	< 100 MB	
Food diaries	Report of the food intake of participants 3 days prior to test day (WP 1 and 2) or of 3 days every week during the intervention period (study 3).	Generate new data	Digital	Experimental	.xlsx	< 1 GB	
Biological samples	Biological samples (blood, feces) collected at study visits. Includes all blood samples collected for all WPs and fecal samples of WP3.	Generate new data	Physical	Experimental			5700 x blood samples; 200 fecal samples
rawdata_metabolomics	Data from targeted metabolomics analysis, ~ 200 blood samples (WP 3).	Generate new data	Digital	Experimental	.mzML .xlsx	< 100 GB (~ 20 GB)	

processed data_metabolomics	The file contains the processed metabolomics data. Data is preprocessed with normalization and alignment and processed with filtering steps.	Generate new data	Digital	Experimental	.xlsx	< 100 GB (~ 10 GB)	
rawdata_shotgun sequencing_gutmicrobiota	Shotgun sequencing data of 200 fecal samples (WP3): sequences generated via the Raes lab and the Nucleomics core (VIB/KU Leuven)	Generate new data	Digital	Experimental	.fastq	< 5 TB (~ 2 TB)	
rawdata_16S rRNA_gutmicrobiota	16S rRNA sequencing of the microbiome of ~200 fecal samples (WP 3) : generated via the Raes lab and the Nucleomics core (VIB/KU Leuven)	Generate new data	Digital	Experimental	.fastq	< 1 TB (~ 200 GB)	
processed_data_gutmicrobiota	Processed data from 16S rRNA sequencing and shotgun sequencing of the fecal samples (alignment, normalization, filtering).	Generate new data	Digital	Experimental	.xlsx .counts	< 2 TB (~ 1.65 TB)	
short-chain fatty acids	Data of the analysis of SCFA (all WPs)	Generate new data	Digital	Experimental	.xlsx	< 100 MB	
GLP-1 and PYY	Results of GLP-1 and PYY analysis (WP 2 and 3)	Generate new data	Digital	Experimental	.xlsx	< 1 GB	
VAS_appetite	VAS questionnaire about appetite (WP2 and 3)	Generate new data	Digital	Experimental	.xlsx	< 100 MB	

GSRS	Gastrointestinal symptom rate scale of all three studies. Questionnaire was completed after each study visit (WP1 and 2). For WP3 the questionnaires were completed at before, during and after the intervention.	Generate new data	Digital	Experimental	.xlsx; .R	< 100 MB	
Rscripts	Rscripts for the statistics (all WPs).	Generate new data	Digital	Experimental	.R	< 1 GB	
Metadata	Data and scripts are explained in excel files.	Generate new data	Digital	Experimental	.csv	< 1 GB	
Standard operating procedures and lab book.	Protocols for analysis.	Generate new data	Digital	Experimental	.docx	< 100 MB	

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:

/

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, there are ethical issues in this research project as we recruit human subjects. Ethical approval has been obtained for the first two studies (S-number: S67256) and a protocol for the ethical approval of WP3 is being prepared. In general, only information relevant for the project will be collected. Data will be pseudonymized and stored and processed in coded form.

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 KU Leuven privacy number for this study is G-2023-6266. We will collect personal data such as participant characteristics, questionnaires, blood and fecal samples. The data will be pseudonymized. The code file will contain the personal data (name/surname, email address, bank account) linked to the pseudonym which is used to organize research. This file will be stored separately from the research data and secured with a password.

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).

To keep the data understandable in the lab, each excel file will contain a sheet explaining the data in the file. Rscripts, sequencing and metabolomics data files will be explained in a separate excel file linked to the corresponding file. After publication of our research, metadata will also be provided in excel files next to the data.

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

Metadata will be provided in excel files next to the data. The mandatory and recommended properties of the DataCite's metadata schema 4.4 will be added.

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

Where will the data be stored?

Demographic data and questionnaires (VAS_appetite and GSRS) will be kept on RedCap software to which only assigned employees of our lab have access to. During the research project, data folders containing pseudonymized clinical data, processed data, sample lists and scripts will be stored on the KU Leuven shared J-drive assigned to our group. The code file with the name, email and bank account of participants and the assigned pseudonym will be stored in a separate file on the shared drive, secured with a password known by the principle and executing investigator. Sequencing (microbiome) and metabolomics data will be kept on the KUL L- large volume storage assigned to our group. The informed consents will be kept in a locked cabinet. Once the project has ended, all data will be stored on the KUL large volume storage (assigned to our group) for long term preservation.

How will the data be backed up?

Standard back-up provided by KU Leuven ICTS.

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 still 1 TB available on the shared J drive which is enough for the clinical data of ~ 5 GB. The large volume storage contains 8 PB which is enough to store the metabolomics and microbiome data (4TB) during the project and all data for long term preservation.

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

To secure our data, we store it within KU Leuven's shared drive and Large Volume Storage both compliant with high and medium confidentiality standards, respectively. KU Leuven's ICT solutions adhere to stringent university-wide information security protocols. Access permissions for the raw network storage managed by the faculty's ICT service are strictly regulated, delegated, and audited by designated data managers who are trained for this role, regardless of their IT expertise. For the code files, which include pseudonyms, access is tightly controlled by a dedicated data manager, with the Principal Investigator serving as an alternate overseer. Researchers involved in the project are granted access exclusively to pseudonymized data, ensuring that personal identifiers are not disclosed.

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

We expect that we need around 5 GB storage space on the shared drive for the clinical data of this project. This corresponds with a cost of 503.66 euro/TB/year or in total 2014.64 euro (4x 503.66 euro x 1 TB). For metabolomics and microbiota data we will use the large volume storage which will cost 104.42 euro/ year/TB. The data will only be collected during the final part of the project and therefore we expect that the large volume storage will cost in total 835.36 euro (2 years x 4 TB x 104.42 euro). These expenses will be covered using our current grant money and through future grant applications.

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...).

The digital data of this research project, biological samples and the informed consents will be preserved for 25 years as according to CTC recommendations for clinical experiments on humans.

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

For long term preservation the digital data of the project will be stored on the large volume storage provided by KU Leuven. Biological samples will be managed and stored by the UZ/KU Leuven Biobank.

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

The expected costs for 25 years data preservation on the large volume storage are 13502.5 euros (104.42 euro/year/TB). The costs will be covered with our current grant money and through future grant applications.

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 an Open Access repository

The data will be shared on Zenodo in a pseudonymized format.

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

/

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, Ethical aspects

We use data from human subjects and personal data so both privacy and ethical aspects have to be taken into account. Data will only be shared in pseudonymized format. The code file will not be shared.

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

All digital data (except the code file) will be made available on the repository Zenodo.

When will the data be made available?

The data will be made available upon publication publication of research results.

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

Data will be deposited under the Creative Commons Attribution International Public License (CC BY) 4.0.

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

A DOI number will be provided.

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

Data sharing via Zenodo is free.

6. Responsibilities

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

Riet Rosseel

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

Riet Rosseel

Who will manage data preservation and sharing?

Riet Rosseel (during project) and Kristin Verbeke (long term)

Who will update and implement this DMP?

Riet Rosseel