FWO DMP Template - Flemish Standard Data Management Plan

Version KU Leuven

Project supervisors (from application round 2018 onwards) and fellows (from application round 2020 onwards) will, upon being awarded their project or fellowship, be invited to develop their answers to the data management related questions into a DMP. The FWO expects a **completed DMP no later than 6 months after the official start date** of the project or fellowship. The DMP should not be submitted to FWO but to the research co-ordination office of the host institute; FWO may request the DMP in a random check.

At the end of the project, the **final version of the DMP** has to be added to the final report of the project; this should be submitted to FWO by the supervisor-spokesperson through FWO's e-portal. This DMP may of course have been updated since its first version. The DMP is an element in the final evaluation of the project by the relevant expert panel. Both the DMP submitted within the first 6 months after the start date and the final DMP may use this template.

The DMP template used by the Research Foundation Flanders (FWO) corresponds with the Flemish Standard Data Management Plan. This Flemish Standard DMP was developed by the Flemish Research Data Network (FRDN) Task Force DMP which comprises representatives of all Flemish funders and research institutions. This is a standardized DMP template based on the previous FWO template that contains the core requirements for data management planning. To increase understanding and facilitate completion of the DMP, a standardized **glossary** of definitions and abbreviations is available via the following link.

1. General Project Information	
Name Grant Holder & ORCID	**Eva Vissers**
Contributor name(s) (+ ORCID) & roles	**João Sabino: promotor** **Christophe Matthys: co-promotor**
Project number [1] & title	Making healthier and safer foods: development of a preclinical platform for testing of food components
Funder(s) GrantID [2]	1SH1L24N
Affiliation(s)	x KU Leuven ☐ Universiteit Antwerpen ☐ Universiteit Gent ☐ Universiteit Hasselt ☐ Vrije Universiteit Brussel ☐ Other: ROR identifier KU Leuven: 05f950310

[1] "Project number" refers to the institutional project number. This question is optional. Applicants can only provide one project number.

[2] Funder(s) GrantID refers to the number of the DMP at the funder(s), here one can specify multiple GrantIDs if multiple funding sources were used.

Please provide a short project description	The consumption of ultra-processed foods (UPFs) has increased significantly over the last decades. Epidemiological studies have shown associations between the consumption of UPFs and chronic diseases with a rising prevalence, such as obesity, cancer and inflammatory bowel diseases. Recently, non-nutritional components of UPFs, including food additives (e.g. emulsifiers) and compounds formed during industrial processing (e.g. dAGEs), are suggested to negatively affect intestinal homeostasis. Despite those possible detrimental effects, a considerable amount of these food components are authorized. Therefore, we hypothesize that the current methods and evaluated outcomes, used in the pre-marketing safety testing of food components, do not capture their possible involvement in chronic health conditions, related to intestinal barrier disruption. In the first work package of this project, we will develop a preclinical platform based on human-derived intestinal organoids to detect intestinal hazards of specific food components. Specifically, we will look into the use of readouts evaluating intestinal permeability, inflammatory pathways, and mucus production. In work package 2 and 3 we will validate our platform by applying it to test the effects of emulsifiers (WP2) and dAGEs (WP3) on the intestinal barrier. To identify the optimal experimental setup (e.g. concentrations, stimulation time…) on a faster and less expensive model, we will use two immortalized intestinal epithelial cell lines, Caco-2 cells and HT29-MTX cells. This project will generate a useful tool for the food industry to test the intestinal effects of new ingredients.

2. 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 [3].

				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
Cells	Cells (from Caco-2/HT29 coculture and from organoid-derived monolayers) are collected after an experiment for further analysis	☒ Generate new data ☐ Reuse existing data	☐ Digital ☒ Physical	☐ Audiovisual ☐ Images ☐ Sound ☐ Numerical ☐ Textual ☐ Model ☐ Software ☐ Other:		☐ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	Cells will be stored at -80°C in the biobank at TARGID. The volume will be 10-15 boxes.
Supernatant	Supernatant (from Caco-2/HT29 coculture and from organoid-derived monolayers) is collected after an experiment for further analysis	☒ Generate new data ☐ Reuse existing data	☐ Digital ☒ Physical	☐ Audiovisual ☐ Images ☐ Sound ☐ Numerical ☐ Textual ☐ Model ☐ Software ☐ Other:		☐ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	Supernatant will be stored at -80°C in the biobank at TARGID. The volume will be 15-20 boxes.
Permeability – TEER	Measurements of the transepithelial electrical resistance (TEER)	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images ☐ Sound ☒ Numerical ☐ Textual	.csv .xlsx	☒ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB	

	of cell monolayers.			☐ Model ☐ Software ☐ Other:		☐ NA	
Permeability – FITC	The passage of FITC-dextran molecules over cell monolayers are measured with a spectrophotometer	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images ☐ Sound ☒ Numerical ☐ Textual ☐ Model ☐ Software ☐ Other:	.DAT .xlsx	☒ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	
qPCR	RNA analysis of cell monolayers via qPCR	☒ Generate new data ☐ Reuse existing data	☒ Digital ☒ Physical	☐ Audiovisual ☐ Images ☐ Sound ☒ Numerical ☐ Textual ☐ Model ☐ Software ☐ Other:	.eds .xlsx	☐ < 1 GB ☒ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	Extracted RNA and cDNA will be stored at -80°C in the biobank at TARGID. The volume will be 20-30 boxes.
Microscopy	Brightfield and confocal microscopy images of cell monolayers	☒ Generate new data ☐ Reuse existing data	☒ Digital ☒ Physical	☐ Audiovisual ☒ Images ☐ Sound ☒ Numerical ☐ Textual ☐ Model ☐ Software ☐ Other:	.tif .jpg .czi .ims .png	☐ < 1 GB ☐ < 100 GB ☒ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	Microscopy slides will be stored at 4°C in the biobank at TARGID. The volume will be 2-3 boxes with slides.
Cytokine measurements – MSD	Analysis of cytokine levels in supernatants via	☒ Generate new data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images	.txt .spe .xlsx	☐ < 1 GB ☒ < 100 GB	

	Mesoscale Discovery platform	☐ Reuse existing data		☐ Sound ☒ Numerical ☐ Textual ☐ Model ☐ Software ☐ Other:		☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	
Lab note books	Dated written notes describing all experimental procedures	☒ Generate new data ☐ Reuse existing data	☐ Digital ☒ Physical	☐ Audiovisual ☐ Images ☐ Sound ☐ Numerical ☐ Textual ☐ Model ☐ Software ☐ Other:		☐ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	Lab books (± 4) will be stored in a closed office at TARGID.

GUIDANCE:
The data description forms the basis of your entire DMP, so make sure it is detailed and complete. It includes digital and physical data and encompasses the whole spectrum ranging from raw data to processed and analysed data including analysis scripts and code. Physical data are all materials that need proper management because they are valuable, difficult to replace and/or ethical issues are associated. Materials that are not considered data in an RDM context include your own manuscripts, theses and presentations; documentation is an integral part of your datasets and should described under documentation/metadata.
RDM Guidance on data

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)? If so, refer to specific datasets or data types when appropriate and provide the relevant ethical approval number.	☒ Yes, human subject data; provide SMEC or EC approval number: ☐ Yes, animal data; provide ECD reference number: ☐ Yes, dual use; provide approval number: ☐ No Additional information: This project involves the use of human samples (intestinal biopsy-derived organoid cultures). It has formal approval by the UZ Leuven ethical committee: S53684.
Will you process personal data[4]? If so, please refer to specific datasets or data types when appropriate and provide the KU Leuven or UZ Leuven privacy register number (G or S number).	☒ Yes (provide PRET G-number or EC S-number below) ☐ No Additional information: Personal data relating to the donors of organoid cultures (e.g. disease status, age, gender) are pseudonymized with a patient ID. The file linking the patient ID with personal identifiers will only be accessible to authorized individuals and stored in a restricted access, secure environment managed by the KU Leuven/UZ Leuven ICT facility. Personal data collection is covered by the Ethical approval of 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 ☐ No If yes, please comment: After finishing the project and before dissemination of the results, we will look carefully if any valorization is in place. In case of any potential, the invention will be evaluated and may be IP protected with the support of the Intellectual Property Unit of KU Leuven Research & Development (LRD). Furthermore, we regularly interact with various stakeholders, including the food industry, which are highly relevant when it comes to the subsequent steps of valorisation of the findings in the current project

[4] See Glossary Flemish Standard Data Management Plan

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 to what data they relate and what restrictions are in place.	☐ Yes ☒ No If yes, please explain:
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 to what data they relate and which restrictions will be asserted.	☐ Yes ☒ No If yes, please explain:

3. 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). *RDM guidance on documentation and metadata*.	Standard experimental procedures (SOPs) will be fully documented (as PDF) and saved on the KUL Shared J-drive assigned to our group. The methodology and protocols 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. Data folders containing the raw and processed data are being stored on our KUL J-drive and K-drive. Each individual file with experimental data will contain information on the study design, the origin of the samples, and all necessary information for an independent analyst to use or reuse the data accurately and efficiently.

Will a metadata standard be used to make it easier to **find and reuse the data**? If so, please specify which metadata standard will be used. If not, please specify which metadata will be created to make the data easier to find and reuse. *REPOSITORIES COULD ASK TO DELIVER METADATA IN A CERTAIN FORMAT, WITH SPECIFIED ONTOLOGIES AND VOCABULARIES, I.E. STANDARD LISTS WITH UNIQUE IDENTIFIERS.*	☐ Yes ☒ No If yes, please specify (where appropriate per dataset or data type) which metadata standard will be used: If no, please specify (where appropriate per dataset or data type) which metadata will be created: 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. - 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 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 - MSD 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).

4. Data Storage & Back-up during the Research Project

Where will the data be stored? *Consult the interactive KU Leuven storage guide to find the most suitable storage solution for your data.*	☒ Shared network drive (J-drive) ☐ Personal network drive (I-drive) ☒ OneDrive (KU Leuven) ☐ Sharepoint online ☐ Sharepoint on-premis ☐ Large Volume Storage ☐ Digital Vault ☒ Other: 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. Physical lab books will be kept in the lab at all time. The physical samples (cells, supernatant, RNA, cDNA…) will be stored in our biobank at TARGID for potential later use. The necessary storage room for the expected amount of biological material is already ascertained.
How will the data be backed up? *WHAT STORAGE AND BACKUP PROCEDURES WILL BE IN PLACE TO PREVENT DATA LOSS?*	☒ Standard back-up provided by KU Leuven ICTS for my storage solution ☐ Personal back-ups I make (specify) ☐ Other (specify)
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 we expect to be sufficient for this project. Larger data files containing raw data can in addition be stored on the K-drive of KU Leuven. For storage of physical samples we have currently sufficient storage space in fridges and freezers. ☐ No If no, please specify:

How will you ensure that the data are securely stored and not accessed or modified by unauthorized persons? *CLEARLY DESCRIBE THE MEASURES (IN TERMS OF PHYSICAL SECURITY, NETWORK SECURITY, AND SECURITY OF COMPUTER SYSTEMS AND FILES) THAT WILL BE TAKEN TO ENSURE THAT STORED AND TRANSFERRED DATA ARE SAFE.* *Guidance on security for research data*	Access to KU Leuven administered drives is conditioned by KU Leuven security groups. These drives are password protected and can only be accessed by people from the research group. Data concerning patient information stored in excel files will be password protected and only the responsible researchers will have access. Hard copies of paper lab notebooks are kept in locked cabinets in the lab of the PIs.
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. When needed, the KU Leuven J- and K-drive are expandable in blocks and funding to cover the costs is available in our group. In case extra freezers are needed for storage of samples, there is funding available in our group to cover these costs.

5. 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...). *Guidance on data preservation*	☒ All data will be preserved for 10 years according to KU Leuven RDM policy ☐ All data will be preserved for 25 years according to CTC recommendations for clinical trials with medicinal products for human use and for clinical experiments on humans ☒ Certain data cannot be kept for 10 years (explain) The cell cultures/organoids will not be stored for 10 years since this is biologically not possible. Data obtained through the experiments and microscopy slides will be stored.

Where will these data be archived (stored and curated for the long-term)? *Dedicated data repositories are often the best place to preserve your data. Data not suitable for preservation in a repository can be stored using a KU Leuven storage solution, consult the interactive KU Leuven storage guide.*	☐ KU Leuven RDR ☐ Large Volume Storage (longterm for large volumes) ☒ Shared network drive (J-drive) ☒ Other (specifiy): Paper lab books will be stored in locked cabinets in the lab of the PI.
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 additional storage space is necessary, the costs can be covered by funding within our research group.

6. Data Sharing and Reuse	
Will the data (or part of the data) be made available for reuse after/during the project? Please explain per dataset or data type which data will be made available. *NOTE THAT 'AVAILABLE' DOES NOT NECESSARILY MEAN THAT THE DATA SET BECOMES OPENLY AVAILABLE, CONDITIONS FOR ACCESS AND USE MAY APPLY. AVAILABILITY IN THIS QUESTION THUS ENTAILS BOTH OPEN & RESTRICTED ACCESS. FOR MORE INFORMATION: HTTPS://WIKI.SURFNET.NL/DISPLAY/STANDARDS/INFO-EU-REPO/#INFOEUREPO-ACCESSRIGHTS*	☐ Yes, as open data ☐ Yes, as embargoed data (temporary restriction) ☒ Yes, as restricted data (upon approval, or institutional access only) ☐ No (closed access) ☐ Other, please specify: 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 transfer 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 per dataset or data type where appropriate.	☐ Yes, privacy aspects ☐ Yes, intellectual property rights ☐ Yes, ethical aspects ☐ Yes, aspects of dual use ☐ Yes, other ☒ No If yes, please specify:
Where will the data be made available? If already known, please provide a repository per dataset or data type.	☒ KU Leuven RDR ☐ Other data repository (specify) ☐ Other (specify)
When will the data be made available?	☒ Upon publication of research results ☐ Specific date (specify) ☐ Other (specify)

Which data usage licenses are you going to provide? If none, please explain why. *A DATA USAGE LICENSE INDICATES WHETHER THE DATA CAN BE REUSED OR NOT AND UNDER WHAT CONDITIONS. IF NO LICENCE IS GRANTED, THE DATA ARE IN A GREY ZONE AND CANNOT BE LEGALLY REUSED. DO NOTE THAT YOU MAY ONLY RELEASE DATA UNDER A LICENCE CHOSEN BY YOURSELF IF IT DOES NOT ALREADY FALL UNDER ANOTHER LICENCE THAT MIGHT PROHIBIT THAT.* *Check the RDR guidance on licences for data and software sources code or consult the License selector tool to help you choose.*	☐ CC-BY 4.0 (data) ☒ Data Transfer Agreement (restricted data) ☐ MIT licence (code) ☐ GNU GPL-3.0 (code) ☐ Other (specify) 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, please provide it here. *INDICATE WHETHER YOU INTEND TO ADD A PERSISTENT AND UNIQUE IDENTIFIER IN ORDER TO IDENTIFY AND RETRIEVE THE DATA.*	☒ Yes, a PID will be added upon deposit in a data repository ☐ My dataset already has a PID ☐ No 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?	For digital data, the filesender of Belnet is for free. If shipment of material is required by another research group, after approval the costs of drafting of MTA/DTA and shipment itself will be covered by the requesting party.

7. Responsibilities	
Who will manage data documentation and metadata during the research project?	The researcher (Eva Vissers) and PI (João Sabino)
Who will manage data storage and backup during the research project?	The researcher (Eva Vissers) and PI (João Sabino)

Who will manage data preservation and sharing?	The PI (João Sabino)
Who will update and implement this DMP?	The PI (João Sabino) bears the end responsibility of updating & implementing this DMP.