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	**Fonds Wetenschappelijk Onderzoek — Vlaanderen (FWO)**
Contributor name(s) (+ ORCID) & roles	**Aspirant fundamenteel onderzoek FWO Sofie Engelborghs (Dossier: 1172825N, ORCID-ID: 0009-0005-9029-8932, Role: PhD Researcher)**
Project number [1] & title	S68879 – ICARUS study
Funder(s) GrantID [2]	
Affiliation(s)	☑ KU Leuven ☐ Universiteit Antwerpen ☐ Universiteit Gent ☐ Universiteit Hasselt ☐ Vrije Universiteit Brussel ☐ Other: ROR identifier KU Leuven: 05f950310
Please provide a short project description	The present study will evaluate the role of both moderate load and overload training on the gut microbiome. The interplay by which exercise-induced alterations in the microbiome relate to changes in the skeletal muscle (gut-muscle axis), performance, and the overreaching phenotype will be explored. To this end, 60 recreationally active volunteers will be randomly allocated to eight weeks of moderate load training, overload training, or no training.

[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.

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**
Eligibility criteria	Preparticipation questionnaire	☒ Generate new data ☐ Reuse existing data	☒ Digital ☒ Physical	☐ Audiovisual ☐ Images ☐ Sound ☐ Numerical ☒ Textual ☐ Model ☐ Software ☐ Other:	(e)CRF: preparticipation questionnaire, .csv	☐ < 1 GB ☒ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	Preparticipation questionnaire
Sport medical history	Sport medical questionnaire	☒ Generate new data ☐ Reuse existing data	☒ Digital ☒ Physical	☐ Audiovisual ☐ Images ☐ Sound ☐ Numerical ☒ Textual ☐ Model ☐ Software ☐ Other:	(e)CRF: Sport medical questionnaire, .csv	☐ < 1 GB ☒ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	Sport medical questionnaire
Gut health	Biological samples: stool and urine samples Covariate questionnaires	☒ Generate new data ☐ Reuse existing data	☒ Digital ☒ Physical	☐ Audiovisual ☐ Images ☐ Sound ☐ Numerical ☒ Textual	Software-specific formats, with extracted numerical data saved as .xls and .csv (e)CRF: "Gastrointestinal and gut transit time	☐ < 1 GB ☐ < 100 GB ☒ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	± 480 stool samples in -80°C freezer ± 1440 urine samples in -20°C freezer. Gastrointestinal and gut

[3] Add rows for each dataset you want to describe.

				□ Model □ Software □ Other:	questionnaire", "medical questionnaire", dietary/bowel habit questionnaire, .csv		transit time questionnaire, medical questionnaire, dietary/bowel habit questionnaire
Muscle phenotype	Biological muscle sample	☒ Generate new data □ Reuse existing data	☒ Digital ☒ Physical	□ Audiovisual ☒ Images □ Sound ☒ Numerical □ Textual □ Model ☐ Software □ Other:	Software-specific formats, with extracted numerical data saved as .xls and histological images saved as .tiff	□ < 1 GB ☐ < 100 GB ☒ < 1 TB □ < 5 TB □ > 5 TB □ NA	± 480 Muscle samples in -80°C freezer
Exercise performance	Aerobic performance and anaerobic performance	☒ Generate new data □ Reuse existing data	☒ Digital ☒ Physical	□ Audiovisual □ Images □ Sound ☒ Numerical □ Textual □ Model ☐ Software □ Other:	(e)CRF: 30min TT and (e)CRF: 30-sec and 6-sec sprint test, .csv Software-specific formats, with extracted numerical data saved as .xls	□ < 1 GB ☒ < 100 GB □ < 1 TB □ < 5 TB □ > 5 TB □ NA	
Physical fitness	Cardiorespiratory fitness, Isometric knee-extensor muscle strength, single-leg critical power	☒ Generate new data □ Reuse existing data	☒ Digital ☒ Physical	□ Audiovisual □ Images □ Sound ☒ Numerical □ Textual □ Model ☐ Software □ Other:	(e)CRF: cardiorespiratory exercise test (CPET), (e)CRF: Promet (3" MVC + 5' all-out CPSL), .csv Software-specific formats, with extracted numerical data saved as .xls	□ < 1 GB ☒ < 100 GB □ < 1 TB □ < 5 TB □ > 5 TB □ NA	
Overreaching	Exercise performance, blood lactate concentration, autonomic nervous system, self-reported recovery-stress	☒ Generate new data □ Reuse existing data	☒ Digital ☒ Physical	□ Audiovisual □ Images □ Sound ☒ Numerical ☒ Textual □ Model ☐ Software	(e)CRF: 30min TT, (e)CRF: 30-sec sprint, (e)CRF: CPET, RESTQ questionnaire, POMS questionnaire, NASA-TLX questionnaire, WURSS-21 questionnaire,	□ < 1 GB ☒ < 100 GB □ < 1 TB □ < 5 TB □ > 5 TB □ NA	Blood samples will be stored in the -80°C freezer

	state, hormonal status, dietary energy intake, mood disturbances, self-reported physical and mental work load Self-reported symptoms of upper respiratory tract infections, sleep quality and quantity			□ Other:	Software-specific formats, with extracted numerical data saved as .xls , .csv		
Anthropometrics	Body weight, body height, body composition (DXA scan)	☒ Generate new data □ Reuse existing data	☒ Digital ☒ Physical	□ Audiovisual □ Images □ Sound ☒ Numerical □ Textual □ Model ☐ Software □ Other:	(e)CRF: anthropometrics Software-specific formats, with extracted numerical data saved as .pdf , .csv	□ < 1 GB ☒ < 100 GB □ < 1 TB □ < 5 TB □ > 5 TB □ NA	
Vascular health	Endothelial function via tissue reoxygenation rate	☒ Generate new data □ Reuse existing data	☒ Digital ☒ Physical	□ Audiovisual □ Images □ Sound ☒ Numerical □ Textual □ Model ☐ Software □ Other:	(e)CRF: Occlusion test Software-specific formats, with extracted numerical data saved as .xls, .csv	□ < 1 GB ☒ < 100 GB □ < 1 TB □ < 5 TB □ > 5 TB □ NA	
Glucose homeostasis	MCH	☒ Generate new data □ Reuse existing data	☒ Digital ☒ Physical	□ Audiovisual □ Images □ Sound ☒ Numerical □ Textual □ Model ☐ Software	(e)CRF: OGTT, .csv Software-specific formats, with extracted numerical data saved as .xls,	□ < 1 GB ☒ < 100 GB □ < 1 TB □ < 5 TB □ > 5 TB □ NA	

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.	N.A.
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: S68879 - B3222024001516 ☐ Yes, animal data; provide ECD reference number: ☐ Yes, dual use; provide approval number: ☐ No Additional information:

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: S68879 - G-2024-7846 – Personal data for organizing the research: name, phone number & email address. This data will not be included in the analysis and will be stored separately from the research data. – Personal data for research purposes (these will be pseudonymized): o Demographic data: age, gender. o Use of medication, dietary supplements, drugs, alcohol. o Body composition. o Training history and current physical activity levels. o Conditions considered exclusion factors for participation in the study: chronic intestinal diseases, cardiovascular diseases, musculoskeletal disorders.
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:
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:

[4] See Glossary Flemish Standard Data Management Plan

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*.	**Data Organization**: All data will be systematically structured and stored in appropriately organized folders. **Code for Complex Data**: MATLAB code has been developed or will be created to handle and process complex data files efficiently. These files will be foreseen of the necessary descriptions to use the scripts correctly. **Standard Operating Procedures (SOPs)**: SOPs have been or will be established for each dataset to ensure consistency and clarity in data handling. A standardized **case report form (CRF)** will be completed during data collection, containing researchers notes, remarks concerning data quality, contextual information, deviations from the protocol, etc. These CRFs will be kept on paper, in the same folder as the research data that are collected on paper. Paper CRFs will be transcribed to REDCap. Given the use of REDCap, a data dictionary/codebook will be available. A README file (KU Leuven template) will be added a project level.
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: At project level • The RDR metadata format will be followed (see Data sharing & reuse) At data level • OME-XML standard for microscopic images If no, please specify (where appropriate per dataset or data type) which metadata will be created:

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: REDCap Data collected on paper: paper copies of the descriptive data and questionnaires will be stored in a secured locker at the Department of Movement Sciences, Building Gymnasium, of the KU Leuven. Only authorized personnel will have access to this locked storage room as they will need to be granted access by the PI (Stefan De Smet).
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 ☐ No If yes, please specify: 1. KU Leuven network drive, specifically L-drive. Our research group has a L-drive with a capacity of 5 TB for active research data. As the estimated size of the dataset is 1 TB sufficient storage and backup capacity is available. 2. REDCap. REDCap is hosted on central ICTS webservices and provides unlimited capacity.

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*	1. KU Leuven network drive, specifically L-drive. The KU Leuven network drives are incorporated within secured KU Leuven environments, are password-protected (including smartphone-based multi-factor identification) and are only accessible by registered collaborating researchers. Only the PI can request access to the network drive for study personnel. 2. REDCap. When using KU Leuven REDCap, physical access to the data centres is logged and restricted to authorized KU Leuven Information Technology (IT) personnel, using badge identification. At the clinical database level only study team members, monitors and auditors/inspectors for whom the Coordinating or Principal Investigator (as applicable) has requested project-specific access, are granted data access. Upon successful training completion each user is centrally assigned a user role, associated with predefined system/data privileges, in accordance with CR DM-WI-001. The gatekeeper for UZL REDCap is UZL CTC (ctc.datamanagement@uzleuven.be). 3. Data collected on paper. Data collected on paper (e.g. informed consents) will be stored in a locked cabinet in a locked room at the department of movement sciences. During data collection the cabinet will only be accessible to study personnel.
What are the expected costs for data storage and backup during the research project? How will these costs be covered?	1. KU Leuven network drive, specifically L-drive. The L-drive costs € 475.7 / 5 TB / year. Our dataset is estimated at maximum 5 TB and the project will run for 4 years, resulting in a total cost of € 1902. The department of Movement Sciences provides our research group with an L-drive of 5TB. As such, costs will be covered by the department. In case of insufficient storage (as the drive is shared by several projects), the drive can be extended. Additional costs could be covered by the FWO bench fee. 2. REDCap. A REDCap project costs €80/year. As the project will run for 3 years, costs are estimated at € 240. This will be covered by the fwo bench fee. 3. Data collected on paper. No costs are attached to storage of data collected on paper.

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)
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): Research data collected on paper, as well as informed consent forms will be stored in the local storage facility at the department of movement sciences. Research data and informed consent forms will be kept is separate folders in a locked cabinet in the locked storage facility, only accessible to the PI.
What are the expected costs for data preservation during the expected retention period? How will these costs be covered?	1. Digital data: Current costs for the K-Drive are € 11.38/100GB/year. Given the expected size of the database of 1 TB = 1000 GB, costs for long-term storage are estimated at € 114 /year. 2. Paper files: No costs are attached to archiving of data collected on paper.

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: All digital data will be made available in a restricted access repository after the last publication.
If access is restricted, please specify who will be able to access the data and under what conditions.	Scientific researchers will have to motivate why they want access to the data. Access to the dataset will be granted for dedicated research questions only. Only data that is essential to answering the proposed research questions will be shared, not the full dataset (unless required to answer the proposed research questions). In their application, requesters are required to provide a research description which includes a description of the research project that the data will be used for and clear information about why these specific data are being requested. We will always ask to give credit to the original data creators when the data is being used by other researchers.
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: We work with confidential data (e.g. sex, age, several subjective perceptions,…). Participants have to consent to data sharing in the informed consent forms. If they do not consent, their data will not be shared. Furthermore, the consent form specifies that data will only be shared for research that is approved by an ethical committee.

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): after publication of the last article when the PI (Stefan De Smet) gives approval.
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)
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
What are the expected costs for data sharing? How will these costs be covered?	RDR is free for KU Leuven personnel, hence, no costs are expected for data sharing.

7. Responsibilities	
Who will manage data documentation and metadata during the research project?	The PhD researcher (Sofie Engelborghs) will be responsible for data documentation & metadata, under supervision of the PI (Stefan De Smet).
Who will manage data storage and backup during the research project?	Data management, storage and back up will be performed by the PhD researcher (Sofie Engelborghs), under supervision of the PI (Stefan De Smet).
Who will manage data preservation and sharing?	The PI (Stefan De Smet) will be responsible for ensuring data preservation and sharing.
Who will update and implement this DMP?	The PhD researcher (Sofie Engelborghs) will be responsible for updating this DMP. The PI (Stefan De Smet) bears the end responsibility for updating and implementing this DMP.