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	**Ilse Vanhorebeek, ORCID 0000-0002-5261-5192**
Contributor name(s) (+ ORCID) & roles	**Greet Van den Berghe, ORCID 0000-0002-5320-1362, co-promotor**
Project number [1] & title	3M240697 Impaired muscle function years after critical illness: a Multi-Omics quest for the underlying mechanisms
Funder(s) GrantID [2]	Fonds Wetenschappelijk Onderzoek (FWO) - G017325N
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	Critically ill patients require vital organ support in an intensive care unit (ICU) to avoid imminent death. Among other complications, they are at high risk of developing muscle weakness, associated with adverse short-term outcomes. Survival improved with advances in intensive care but a substantial proportion of the patients show persistent physical impairments long after hospital discharge, compromising quality of life. Muscular alterations in critically ill patients are reminiscent of accelerated aging. Epigenetic changes are involved in muscle development and regeneration, accumulate with aging and are likely involved in permanent health effects of transient environmental influences. We hypothesize that long-term epigenetic changes that lead to altered RNA expression in muscle of former ICU patients as compared with matched controls may contribute to long-term adverse physical outcomes. Five years after ICU admission we will study the muscle transcriptome, investigate if abnormal long-term RNA expression may be explained by abnormal DNA methylation and investigate via an epigenetic clock if former ICU patients show accelerated biological aging, in relation with long-term physical outcome. Next, we will study time profiles of DNA methylation changes in muscle in ICU, in relation to morphological, molecular and functional changes during critical illness. Our research has great potential to provide new targets and a window of opportunity for future therapeutic intervention.

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
Objective 1	Long-term RNA expression after critical illness in relation with long-term physical outcomes: - Detailed clinical data from original EPaNIC study and follow-up - Muscle transcriptome raw and processed data; R code used for corresponding statistical analyses - Results from additional laboratory analyses - Stained muscle sections and corresponding photographs - Output of statistical	☒ Generate new data ☒ Reuse existing data	☒ Digital ☒ Physical	☐ Audiovisual ☒ Images ☐ Sound ☒ Numerical ☒ Textual ☐ Model ☐ Software ☐ Other:	.fmp .jmp .fastq .txt .Rmd .html .xlsx .csv .tiff .ppt .pdf .readme	☐ < 1 GB ☐ < 100 GB ☐ < 1 TB ☒ < 5 TB ☐ > 5 TB ☐ NA	- Muscle biopsies already available from 120 former critically ill patients and 31 controls, stored in cryotubes (up to ~300 mg, split over up to 6 tubes) - RNA extracted from those samples, stored in Eppendorf tubes - Ten tissue section boxes for glasses with stained muscle sections (capacity of 100 slides per box)

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

	analyses performed in JMP copied to powerpoint and converted to pdf						
Objective 2	Long-term DNA methylation after critical illness in relation with long-term RNA expression and physical outcomes: - Detailed clinical data from original EPaNIC study and follow-up - Muscle transcriptome processed data - Muscle genome-wide DNA methylation raw and processed data - R code used for "big data" statistical analyses - Output of statistical analyses performed in JMP copied to powerpoint and converted to pdf	☒ Generate new data ☒ Reuse existing data	☒ Digital ☒ Physical	☐ Audiovisual ☐ Images ☐ Sound ☒ Numerical ☒ Textual ☐ Model ☐ Software ☐ Other:	.fmp .jmp .idat .Rmd .html .xlsx .csv .ppt .pdf .readme	☐ < 1 GB ☐ < 100 GB ☒ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	- Muscle biopsies already available from 120 former critically ill patients and 31 controls, stored in cryotubes (up to ~300 mg, split over up to 6 tubes, same as for objective 1) - DNA extracted from those samples, stored in Eppendorf tubes
Objective 3	Biological aging of muscle after critical illness: - Detailed clinical data from original EPaNIC	☒ Generate new data ☒ Reuse existing data	☒ Digital ☒ Physical	☐ Audiovisual ☐ Images ☐ Sound ☒ Numerical ☒ Textual	.fmp .jmp .idat .Rmd .html	☐ < 1 GB ☒ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB	- Muscle biopsies already available from 120 former critically ill patients and 31 controls,

	study and follow-up - Muscle genome-wide DNA methylation raw and processed data - R code used for “big data” statistical analyses - Output of statistical analyses performed in JMP copied to powerpoint and converted to pdf			☐ Model ☐ Software ☐ Other:	.xlsx .csv .idat .ppt .pdf .readme	☐ NA	stored in cryotubes (up to ~300 mg, split over up to 6 tubes, same as for objective 1) - Additional 28 former patient and 20 control muscle biopsies, also already available - DNA extracted from those samples, stored in Eppendorf tubes
Objective 4	Time profiles of DNA methylation changes in muscle during the ICU stay, in relation to morphological, molecular, and functional changes during critical illnes - Detailed clinical data from CROSS trial - Muscle transcript-tome raw and processed data - Muscle genome-wide DNA methylation raw and processed data - R code used for “big	☒ Generate new data ☒ Reuse existing data	☒ Digital ☒ Physical	☐ Audiovisual ☒ Images ☐ Sound ☒ Numerical ☒ Textual ☐ Model ☐ Software ☐ Other:	.fmp .jmp .fastq .txt .Rmd .html .xlsx .csv .idat .tiff .ppt .pdf .readme	☐ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☒ > 5 TB ☐ NA	- Muscle biopsies already available from 153 critically ill patients and 20 controls, stored in cryotubes (up to ~300 mg, split over up to 6 tubes) - RNA and DNA extracted from those samples, stored in Eppendorf tubes - Ten tissue section boxes for glasses with stained

	data" statistical analyses - Results from additional laboratory analyses - Output of statistical analyses performed in JMP copied to powerpoint and converted to pdf						muscle sections (capacity of 100 slides per box)

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.	We will reuse clinical data that we have collected during the large randomized EPaNIC trial and its long-term follow-up (datasets for objectives 1 to 3) and during the large observational CROSS trial (dataset for objective 4) performed by our research group. These data are stored on UZ Leuven servers. The present objectives focus on new analyses to be performed on skeletal muscle biopsies that we have collected during those studies and are stored in the biobank. The data that will be reused are participants' characteristics, clinical information and physical outcomes.
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: S50404: objectives 1 to 3 (EPaNIC RCT and its long-term follow-up) S58533: objective 4 (CROSS trial) ☐ Yes, animal data; provide ECD reference number: ☐ Yes, dual use; provide approval number: ☐ No
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) S50404: objectives 1 to 3 (EPaNIC RCT and its long-term follow-up) S58533: objective 4 (CROSS trial) ☐ No

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: Although not the primary objective, in case study observations could lead to intellectual property rights, patent applications will be drafted together with the KU Leuven Technology Transfer office.
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

[4] See Glossary Flemish Standard Data Management Plan

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*.	For the objectives in this project, we will study skeletal muscle biopsies from a subgroup of patients and controls included in the EPaNIC/EPaNIC follow-up and CROSS trials. All data that have been collected during these clinical studies are stored in large structured Filemaker databases to which newly obtained data are fed. The study protocols describe data collection and definition of variables, standing operating procedures are in place to describe data collection, and for more complex data a definition is provided as info label in the database. All these documents are stored electronically in the structured study master file. All stored data can be queried by our clinical data manager to retrieve specific participants or samples. Our clinical data manager provides requested data as exports in Excel format. These Excel files are then read in JMP or R in which the statistical analyses of the data are performed. We will keep a separate registry documenting the names and locations for raw and processed data exports as used for every step in the project.
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: All participant characteristics can be used as metadata in the structured Filemaker case report form to retrieve specific participants or samples. Metadata of laboratory analyses will be provided as readme, word or excel files, that contain all settings and technical descriptions of the performed analyses and the resulting data.

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 for raw data of transcriptome/methylome analyses and tissue images ☐ Digital Vault ☒ Other: Shared network drive UZ Leuven server for clinical master databases
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): Standard back-up provided by UZ Leuven IT for clinical master databases
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 The storage volume made available by UZ/KU Leuven is appropriate for storage. Standard back-up is provided.
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*	The clinical master databases are user ID/password protected, with logged access control at network, directory and database level. Biological samples are stored in a registered biobank, only accessible to authorized people, with a log record of all sample handlings. The databases are stored on secure servers within UZ/KU Leuven, maintained by the IT department and maximally protected by firewalls and login procedures with daily backups. The biobank has standard procedures to protect adequate storage.
What are the expected costs for data storage and backup during the research project? How will these costs be covered?	UZ Leuven servers: data preservation is currently free of costs. KU Leuven large volume storage: Estimated costs for storage of new data generated in the project~€1500. These costs will be covered by budgets of the Laboratory of Intensive Care Medicine.

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 (for the data generated in this project) ☒ 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 (participants' characteristics, clinical information and physical outcomes that have been collected during the clinical studies and that will be reused in this project) ☐ 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): UZ Leuven servers
What are the expected costs for data preservation during the expected retention period? How will these costs be covered?	- KU Leuven RDR: currently free of costs. - Large Volume Storage: ~€5000, will be covered by budgets of the Laboratory of Intensive Care Medicine. - UZ Leuven servers: data preservation is currently free of costs.

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:
If access is restricted, please specify who will be able to access the data and under what conditions.	Data sharing will be considered only on a collaborative basis with the principal investigators of the project, after evaluation of the proposed study protocol and statistical analysis plan.
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: The clinical databases contain sensitive and personal information of the study participants. Even though the data are pseudonymized, a theoretical possibility remains that a patient could be identified based for instance on a combination of demographic characteristics, admission date and admission diagnosis. Therefore, it is of utmost importance to only share data that are necessary to answer a specific research question, and only under a data transfer and confidentiality agreement.

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) As mentioned, the clinical data will not be made available unconditionally in the public space, due to ethical and privacy restrictions.
When will the data be made available?	☐ Upon publication of research results ☐ Specific date (specify) ☒ Other (specify) Data sharing will be considered only on a collaborative basis with the principal investigators, after evaluation of the proposed study protocol and statistical analysis plan, and after signing a data transfer and confidentiality agreement.
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?	We do not expect any costs for data transfer, but in case they would occur, they will be covered by the requesting parties.

7. Responsibilities	
Who will manage data documentation and metadata during the research project?	The PhD student affiliated with the project and the clinical database manager of the research group will manage data documentation and metadata, under supervision of the principal investigators (Ilse Vanhorebeek and Greet Van den Berghe).
Who will manage data storage and backup during the research project?	The PhD student affiliated with the project and the clinical database manager of the research group will manage data documentation and metadata, under supervision of the principal investigators (Ilse Vanhorebeek and Greet Van den Berghe).
Who will manage data preservation and sharing?	Ilse Vanhorebeek and Greet Van den Berghe
Who will update and implement this DMP?	Ilse Vanhorebeek