FWO DMP Template - Flemish Standard Data Management Plan

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	**Vittoria Marini (0000-0002-0504-8682)**
Contributor name(s) (+ ORCID) & roles	**Principal Invistigator: Prof. Maurilio Sampaolesi (0000-0002-2422-3757)** **Co-promotor: Dr. Yoke Chin Chai (0000-0001-9913-257X)**
Project number[1] & title	1S98723N, Spatially-Organized 3D Human Cardiac Bilayer Models for Duchenne Cardiomyopathy Modelling and Drug Discovery
Funder(s) GrantID[2]	FWO (#G066821N and #1S98723N)
Affiliation(s)	☒ KU Leuven ☐ Universiteit Antwerpen ☐ Universiteit Gent ☐ Universiteit Hasselt ☐ Vrije Universiteit Brussel ☐ Other: Provide ROR[3] identifier when possible:

[1] "Project number" refers to the institutional project number. This question is optional since not every institution has an internal project number different from the GrantID. 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.

[3] Research Organization Registry Community. https://ror.org/

Please provide a short project description	Duchenne Muscular Dystrophy (DMD) is the most common neuromuscular disorder characterized by the absence of DYSTROPHIN. Nowadays, cardiomyopathy represents the major cause of death in late-stage DMD patients. DMD patients often develop left ventricular non-compaction (LVNC) cardiomyopathy, due to alterations of the endocardium: prominent left ventricular trabeculae and deep intertrabecular recesses. The underlying mechanism of endocardium derangement in LVNC has been poorly investigated, largely due to the lack of human samples and preclinical models that reliably recapitulate the human disease hallmarks. Recently, we developed 3D cardiac organoids from DMD patient-derived induced pluripotent stem cells (iPSCs) with cardiomyopathic features. This PhD project aims at generating DMD cardiomyopathy 3D cardiac bilayer models with improved anatomical- and pathophysiological-relevant spatial complexity, reproducibility and predictivity. The spatially-organized DMD-LVNC 3D cardiac bilayer models will be obtained by combining dystrophic myocardial and endocardial layers derived from DMD-iPSCs using 3D bioprinting technology. By means of single-cell RNA sequencing and spatial transcriptomics, we aim to unravel the dysregulated pathways associated with LVNC pathogenesis at single-cell spatial precision. The newly generated DMD-LVNC 3D cardiac bilayer models will be used to screen NOX4 inhibitors as a proof-of-concept of their applicability and validity for drug screening purposes.

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[4].

				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
WP1-2: Insertion of iETV2 cassette in DMD-iPSC and DMD-Iso-iPSC lines and pluripotency studies.	Insertion of a doxycycline-inducible ETV2 overexpression gene cassette into the DMD patient-derived iPSC and the corresponding DMD isogenic lines available in our lab. Characterization and validation of pluripotency properties of the newly generated DMD-iETV2-hiPSC and DMD-Iso-iETV2-iPSC lines to ensure that any off-target affecting the pluripotent state	☒ 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	Physical data will be stored at -20°C (CRISPR/Cas9 related plasmids and constructs; DNA), at -80°C (RNA) and at -150°C (iETV2- hiPSCs; DMD-iETV2- hiPSCs).

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

	of the line be excluded from the study.						
WP3-4: Bioprinting of DMD-LVNC 3D cardiac bilayer models and transcriptomic analyses	The DMD-iETV2-iPSCs will be differentiated toward DMD-CMs and DMD endocardial-like cells. Afterwards, these two populations will be printed into multi-layered DMD-LVNC 3D cardiac bilayer constructs using the extrusion-based bioprinting process. The iETV2-hiPSC line will be used to generate the corresponding healthy models as controls. To identify the cell populations developing from the crosstalk	☒ 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	Physical data will be stored at -20°C (DNA), at -80°C (RNA). Freezer stocks of histological slides will be available upon request.

	between the dystrophic myocardium and endocardium within the DMD-LVNC 3D cardiac bilayer models, I will perform single cell RNA sequencing. Additionally, the spatial transcriptomic analysis will allow to visualize these populations on the 3D construct sections.						
WP5: NOX4 inhibitors drug screening on DMD-CM monolayer and DMD-LVNC 3D cardiac bilayer models	In collaboration with Prof. Andrea Mattevi (University of Pavia, IT), we will assess four NOX4 inhibitors (MC4762, MC4767, Molport11, Molport41) on my dystrophic	☒ 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:	☐ < 100 MB ☐ < 1 GB ☒ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ < 10 TB ☐ < 50 TB ☐ > 50 TB ☐ NA	Physical data will be stored at -20°C (DNA), at -80°C (RNA). Freezer stocks of histological slides will be available upon request.

	models. Firstly, the screening will be carried out on 2D cultures of DMD-CMs in order to select the most promising ones, which in turn will be tested on the DMD-LVNC 3D cardiac bilayer models. Following the identification of the optimal concentration of compounds, I will treat the dystrophic models with the inhibitors for 14 days and analyse the cell viability using XTT assay and propidium iodide/ Calcein-AM staining. Their effect on cell death and ROS levels as well as				☐ NA		

	the intracellular calcium handling and mitochondrial health will be evaluated.						

GUIDANCE:

DATA CAN BE DIGITAL OR PHYSICAL (FOR EXAMPLE BIOBANK, BIOLOGICAL SAMPLES, …). DATA TYPE: DATA ARE OFTEN GROUPED BY TYPE (OBSERVATIONAL, EXPERIMENTAL ETC.), FORMAT AND/OR COLLECTION/GENERATION METHOD.

EXAMPLES OF DATA TYPES: OBSERVATIONAL (E.G. SURVEY RESULTS, SENSOR READINGS, SENSORY OBSERVATIONS); EXPERIMENTAL (E.G. MICROSCOPY, SPECTROSCOPY, CHROMATOGRAMS, GENE SEQUENCES); COMPILED/AGGREGATED DATA[5] (E.G. TEXT & DATA MINING, DERIVED VARIABLES, 3D MODELLING); SIMULATION DATA (E.G. CLIMATE MODELS); SOFTWARE, ETC.

EXAMPLES OF DATA FORMATS: TABULAR DATA (.POR,. SPSS, STRUCTURED TEXT OR MARK-UP FILE XML, .TAB, .CSV), TEXTUAL DATA (.RTF, .XML, .TXT), GEOSPATIAL DATA (.DWG,. GML, ..), IMAGE DATA, AUDIO DATA, VIDEO DATA, DOCUMENTATION & COMPUTATIONAL SCRIPT.

DIGITAL DATA VOLUME: PLEASE ESTIMATE THE UPPER LIMIT OF THE VOLUME OF THE DATA PER DATASET OR DATA TYPE.

PHYSICAL VOLUME: PLEASE ESTIMATE THE PHYSICAL VOLUME OF THE RESEARCH MATERIALS (FOR EXAMPLE THE NUMBER OF RELEVANT BIOLOGICAL SAMPLES THAT NEED TO BE STORED AND PRESERVED DURING THE PROJECT AND/OR AFTER).

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.	Some iPSC lines described in WP 1-5 have been previously generated and reported in the following papers: doi: 10.1016/j.stemcr.2021.12.019 doi: 10.1038/s41398-019-0535-1

[5] These data are generated by combining multiple existing datasets.

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, please describe these issues further and refer to specific datasets or data types when appropriate.	☒ Yes, human subject data ☐ Yes, animal data ☐ Yes, dual use ☐ No If yes, please describe: We will perform experiments on human-derived induced pluripotent stem cells. The research will be performed under normal laboratory safety rules. All necessary safety measures for laboratory will be taken. For use of patient derived stem cells in the research, we will follow the guidelines and rules from the HSE Department (Health, Safety and Environment). The use of human samples from healthy control donors and DMD subjects for experimental purposes and protocols in the present study was approved by the Ethics Committee of the University Hospitals Leuven (S66794 and S65190)
Will you process personal data[6]? If so, briefly describe the kind of personal data you will use. Please refer to specific datasets or data types when appropriate. If available, add the reference to your file in your host institution's privacy register.	☐ Yes ☒ No If yes: - Short description of the kind of personal data that will be used: - Privacy Registry Reference:

[6] See Glossary Flemish Standard Data Management Plan

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: The 3D models proposed in this project will be used as a platform for screening new drug candidates aimed to treat cardiomyopathies related to DMD. The incorporation of multiple cell types, the standardization in size and volume as well as the ordered architecture thanks to the bioprinting process, and the bioink biocompatibility make this model easily applicable in routine industrial drug testing and high throughput screening by overcoming the common limitations for 3D models in the pharmaceutical field.
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).	Documentation and metadata linked to each experiment will be documented by research groups in hard copy lab notebooks for this project. This includes the research design, protocol, context of data collection, data collection methods, quality control procedures, processing and analysis procedures. **Digital data:** -Experimental design and protocol (.docx file) -Steps involved in data analysis and relevant analysis scripts (R, ImageJ) -Raw data (specific file format according to data type) -Analysed data (specific file format according to data type) **Physical data:** Samples used for the experiments will be documented and stored for up to three years after the end of the project. Storage will be in fixative or in freezers depending on the kind of sample. Immunohistological stained slides will be stored in appropriate boxes in a dry place or freezer.
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: We will adopt a single, well-defined file-folder structure and file-naming rules. Every data folder will be accompanied by appropriate metadata files consisting of a readme.txt with info on nomenclature, file format, software and adopted data standards.

4. Data Storage & Back-up during the Research Project	
Where will the data be stored?	The host institute provides a secure data storage system (KU Leuven LUNA servers) with automated onsite back-up and mirroring. Every person has storage capacity of 2 TB with regular backup system (OneDrive) so the data will be stored there for active use and copies can be made and kept on personal devices. For active use of the data during the project, OneDrive will ensure data transfer between computers, and will also be stored on the KU Leuven LUNA Large Volume Storage space. Biological samples will be taken, and stored in labelled fridges, freezers and closets in the lab. The inventory of all locations is shared on the KU Leuven LUNA Shared drive.
How will the data be backed up? *WHAT STORAGE AND BACKUP PROCEDURES WILL BE IN PLACE TO PREVENT DATA LOSS? DESCRIBE THE LOCATIONS, STORAGE MEDIA AND PROCEDURES THAT WILL BE USED FOR STORING AND BACKING UP DIGITAL AND NON-DIGITAL DATA DURING RESEARCH.*[7] *REFER TO INSTITUTION-SPECIFIC POLICIES REGARDING BACKUP PROCEDURES WHEN APPROPRIATE.*	The data will be stored on the university's central servers (OneDrive) with automatic daily backup procedures.
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 concisely: storage capacity of 2 TB with regular backup system If no, please specify:

[7] Source: Ghent University Generic DMP Evaluation Rubric: https://osf.io/2z5g3/

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.* [7]	All storage spaces (OneDrive and J drive) are hosted in the KU Leuven ICTS data center, with a mirror in the second ICTS center. They are accessible only with my KU Leuven credentials, that provides disaster recovery and additional back-up capacity with guaranteeing long-term data availability.
What are the expected costs for data storage and backup during the research project? How will these costs be covered?	J-drive (shared drive): 503.66 euro/TB/year OneDrive: every student/researcher gets 2 TB for free I-drive (personal drive): 503.66 euro/TB/year Digital vault for private data: windows server 1302 euro/year The costs will be covered by part of the allocated project budget.

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...).	Digital data will be retained for the expected 5 year period and for most publications we expect that we will make the data publicly available on data repositories. Laboratory notebooks and electronic data will be stored by the PI, Prof. Maurilio Sampaolesi. Sequencing data will be submitted to public databases (EBI-ENA/NCBI-SRA), where they will be permanently archived to preserve access to the public. Physical data: Freezer stocks of histological slides will be available upon request. After the conclusion of the project samples will be stored for up to three years after the end of the project. Storage will be in fixative or in freezers depending on the kind of sample.
Where will these data be archived (stored and curated for the long-term)?	Digital data: KU Leuven LUNA servers Physical data: Notebooks and publications: Laboratory repository

What are the expected costs for data preservation during the expected retention period? How will these costs be covered?	KU Leuven generally recommends in its RDM policy to keep **relevant research data** generated during research projects for a minimum of 10 years for reproducibility, verification and potential reuse: Digital vault for private date (windows server) 1302 euro/year Large volume storage is intended for long-term storage of l**arge volumes of research data** in a **cost-efficient** manner: 104.42 euro/TB/year 9 (to be purchased in blocks of 5 TB) This cost will be covered by a general lab budget.

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, in an Open Access repository ☐ Yes, in a restricted access repository (after approval, institutional access only, …) ☐ No (closed access) ☐ Other, please specify: Sequencing data will be submitted to public databases (EBI-ENA/NCBI-SRA), where they will be permanently archived to preserve access to the public. Written progress reports, thesis will be stored for internal purposes and can be accessed by KU Leuven researcher upon request. For most publications we expect that we will make the data publicly available on data repositories.
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 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.	In an Open Access repository
When will the data be made available? *THIS COULD BE A SPECIFIC DATE (DD/MM/YYYY) OR AN INDICATION SUCH AS 'UPON PUBLICATION OF RESEARCH RESULTS'.*	Upon publication of the research results
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.* *EXAMPLE ANSWER: E.G. "DATA FROM THE PROJECT THAT CAN BE SHARED WILL BE MADE AVAILABLE UNDER A CREATIVE COMMONS ATTRIBUTION LICENSE (CC-BY 4.0), SO THAT USERS HAVE TO GIVE CREDIT TO THE ORIGINAL DATA CREATORS."* [8]	All the studies will result in papers subjected to a peer review process. All accepted articles will be published under a CC-BY 4.0 license and also those articles that meet the open access requirements of funding agencies and would be submitted to PubMed. Indeed the first publication I got at the beginning of my PhD program is in *Frontiers Cell and developmental Biology* (doi: 10.3389/fcell.2022.878311) and all Frontiers articles from July 2012 onwards are published with open access under the Creative Commons CC-BY license (the current version is CC-BY, version 4.0).
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 ☐ No If yes:

[8] Source: Ghent University Generic DMP Evaluation Rubric: https://osf.io/2z5g3/

What are the expected costs for data sharing? How will these costs be covered?	The expected cost for data sharing will be low, since the use of OneDrive is free for KU Leuven members up to 1TB.

7. Responsibilities	
Who will manage data documentation and metadata during the research project?	Responsibility for generating data and ensuring data preservation and sharing is with the supervisors (Maurilio Sampaolesi, Yoke Chin Chai) and myself.
Who will manage data storage and backup during the research project?	Supervisors (Maurilio Sampaolesi, Yoke Chin Chai) and myself.
Who will manage data preservation and sharing?	Supervisors (Maurilio Sampaolesi, Yoke Chin Chai) and myself.
Who will update and implement this DMP?	Supervisors (Maurilio Sampaolesi, Yoke Chin Chai) and myself.