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	Sam Verwimp, 0000-0002-9957-0307
Contributor name(s) (+ ORCID) & roles	Leen Delang, 0000-0002-8874-675X, promotor Chris Callewaert, 0000-0001-7697-9188, co-promotor
Project number[1] & title	Impact of skin bacteria on the infection and disease of mosquito-borne Ross River virus
Funder(s) GrantID[2]	11D5923N
Affiliation(s)	[x] KU Leuven ☐ Universiteit Antwerpen ☐ Universiteit Gent ☐ Universiteit Hasselt ☐ Vrije Universiteit Brussel ☐ Other: Provide ROR[3] identifier when possible: 05f950310

[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	Mosquito-borne viruses are posing an enormous threat to public health. Despite their severe disease burden and worldwide spread, no treatment or prophylaxis is available, indicating the need for new antiviral strategies. The host skin is the first site of viral infection and replication upon an infectious mosquito bite, which is crucial for further viral dissemination in the body and for disease. It is currently poorly understood whether host bacteria, present at the bite site in the skin, affect mosquito-borne virus and disease in the host. Our preliminary work shows that topical antibiotic treatment of the skin at the site of infection resulted in more severe Zika virus-induced disease in AG129 mice. However, whether bacteria were directly involved in the phenotype, still remains to be defined. This project aims to further characterize this effect in an immunocompetent mouse model. To this end, we will unravel the effect of skin bacteria on the replication, dissemination and pathogenesis of Ross River virus, an emerging mosquito-borne virus of growing concern, in immunocompetent mice. Furthermore, we will verify whether skin bacteria can affect the virus infection in ex vivo human skin. The outcomes of this research proposal will majorly contribute to the knowledge on the initial and crucial stage of the mosquito-borne virus infection, providing fundamental steps towards new strategies to fight mosquito-borne virus diseases.

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
		☐ 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	
Protocols	Written protocols	New data	Digital	Other: descriptive	.docx .pdf	<100 MB	NA
Experiment measurements and observations	Measurements and observations (survival of	New data	Digital	Experimental	.xls	<100 MB	NA

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

	mice, body weight, footpad swelling,…)						
Biological samples	All biological samples resulting from experiments (cells, bacteria, viruses, mosquito saliva, mouse tissues, human skin biopsies, …)	New data	Physical	NA	NA	NA	6000 samples
qPCR raw data	Files resulting from QuantStudio RT-PCR System	New data	Digital	Experimental	.eds	<1 GB	NA
qPCR results	Excel files exported from QuantStudio Analysis Software with calculations of viral RNA levels	New data	Digital	Compiled	.xls	<100 MB	NA

Plaque assay images	Images taken from plaque assay plates	New data	Digital	Experimental	.png	<100 MB	NA
Plaque assay results	Excel files with raw data and analysis + calculations	New data	Digital	Compiled	.xls	<100 MB	NA
Blood agar plating images	Images taken from blood agar plates	New data	Digital	Experimental	.png	<100 MB	NA
Blood agar plating results	Excel files with raw data and analysis + calculations	New data	Digital	Compiled	.xls	<100 MB	NA
Intravital imaging	Images resulting from intravital imaging	New data	Digital	Experimental	.tif	<100 GB	NA
Flow cytometry raw data	Files resulting from LSRFortessa Flow Cytometer system	New data	Digital	Software	.fcs	<100 GB	NA
Flow cytometry analysis	Files for processing and visualizing flow	New data	Digital	Software	.wsp	<100 MB	NA

	cytometry data in FlowJo Software						
Flow cytometry results	Excel files for calculations resulting from analysis in FlowJo Software	New data	Digital	Compiled	.xls	<100 MB	NA
Flow cytometry layouts	Layout of graphs made in FlowJo Software	New data	Digital	Compiled	.pdf	<100 MB	NA
Sequencing raw data	Genetic sequences	New data	Digital	Experimental	.fastq	<100 MB	NA
Sequencing results	Analysis of results from sequencing + calculations	New data	Digital	Compiled	.xls	<100 MB	NA
Graphs	Graphpad software files with graphs resulting from analysis of data	New data	Digital	Software	.pzfx	<100 MB	NA
Graph images	Exported images from	New data	Digital	Compiled	.tif	<1 GB	NA

	graphs made in Graphpad software						
Manuscripts	Manuscripts for publications of results	New data	Digital	Compiled	.pdf	<100 MB	NA

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.	NA

[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: All experimental work is/will be approved by the relevant ethical committees. For the mouse experiments, we already obtained approval by the Ethical Committee for Animal Experimentation (KU/UZ Leuven) (P071/2019). For the human skin biopsy experiments in WP3, we will submit an application with the Committee for Medical Ethics (VUB/UZ Brussel) in order to obtain ethical approval prior to performing the experiments.
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:
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: NA

[6] 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: NA
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: NA

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).	Daily labwork (protocols, calculations, results,…) will be documented in an online OneNote labbook which is continuously being backed up by KU Leuven servers. Additionally, original files with raw data and files with analysed data will be labelled and stored on servers controlled and backed up by the KU Leuven IT department.

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: For microscopy images of the intravital imaging experiments: the Leica LASX software generates metadata files for all images, which will be stored stored together with the image. For flow cytometry, FCS files containing metadata will be generated and stored. If no, please specify (where appropriate per dataset or data type) which metadata will be created: NA

4. Data Storage & Back-up during the Research Project	
Where will the data be stored?	All data will be stored on drives controlled and backed up by KU Leuven. Data with small volumes will be stored on the J drive, in a subfolder that can only be accessed by personnel of the Delang lab. Data with large volumes (microscopy images) will be stored on the L drive of our team. The separate K-drive of KU Leuven will be used for long term storage of files and data from finished projects. Biological samples from experiments will be stored in freezers and registered in https://freezerpro.rega

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 KU Leuven central servers (J/K/L drives). A back-up of the data on these drives will automatically be generated two times per day. Additionally, data will be mirrored and stored on a cloud-based service offered by KU Leuven (OneDrive), which is synced every 10 minutes.
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: All data with small volumes will be stored on the J drive controlled by KU Leuven. There is sufficient storage space foreseen (1.4 Tb) and this is constantly monitored by KU Leuven IT services. Data with larger volumes (microscopy images, FASTQ files) will be stored on a specifically allocated L drive of KU Leuven on which sufficient storage space is foreseen (10 TB) and which is also constantly monitored by KU Leuven IT services. The separate K-drive of KU Leuven will be used for long term storage of files and data from finished projects (200 GB). If needed, capacity of these KU Leuven drives can be increased at any time. If no, please specify: NA

[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 data will be stored on a KU Leuven backed up server, for which access is only granted to the Delang lab members. This access is controlled by the head of our research group (Leen Delang).
What are the expected costs for data storage and backup during the research project? How will these costs be covered?	The costs of a KU Leuven server storage are: 415.2 euros/year for the J drive (1.4 TB), 1138.4 euros/year for the L drive (10 TB) and 22.768 euros/year for the K drive. The costs for data storage and backups are concerning the whole research group and are not specific for this project. Hence, the costs will be divided over all funding available by our group including the bench fee available by this project.

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...).	All generated data of this project will be stored in a folder on the network drive specifically designated for long term storage (K drive), which is controlled and backed up by KU Leuven. Data will be retained for at least 10 years, conform the KU Leuven RDM policy. Biological samples (RNA, tissues,…) will be stored in freezers (-80°C) until publication of the results. Relevant samples (virus stocks, cell lines) which can be reused in other projects will be preserved in freezers as long as possible.
Where will these data be archived (stored and curated for the long-term)?	The data, associated metadata and electronical labbooks will be stored on the K drive of KU Leuven with automatic back-up procedures for at least 10 years, conform the KU Leuven RDM policy.

What are the expected costs for data preservation during the expected retention period? How will these costs be covered?	Costs to preserve data on the K drive will depend on the storage size at a specific moment in time as this can always be increased/decreased on demand, but are estimated at 11.4 euros/100 GB. This is paid annually and concern the whole research group. The costs will be divided over all funding available by 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, in an Open Access repository ☒ Yes, in a restricted access repository (after approval, institutional access only, …) ☐ No (closed access) ☒ Other, please specify: by mail The key findings of this project will be made available through publication in peer-reviewed journals. Upon publication, relevant raw data and experimental details will be made available in the KU Leuven data repository. Additionally, data might be made available upon reasonable request by mail.
If access is restricted, please specify who will be able to access the data and under what conditions.	Data will be available on the KU Leuven research data repository (after publication) or by mail on individual basis to potential collaborators or interested researchers upon reasonable request, which will be assessed by the head of our research group Prof. Delang.
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: NA
Where will the data be made available? If already known, please provide a repository per dataset or data type.	Data will be deposited on the KU Leuven research data repository or by mail upon request.

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'.*	At the end of the research project and after publication of the research results, data will be deposited on the KU Leuven research data repository.
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]	Data will not be protected by license as it concerns fundamental research and has no application potential. After publication of the respective data, data will be shared without restrictions.
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?	Costs will be controlled by the research group and divided over all available funding and discussed with collaborators.

7. Responsibilities	
Who will manage data documentation and metadata during the research project?	The grant holder, Sam Verwimp, will be responsible for data and metadata documentation and preservation.
Who will manage data storage and backup during the research project?	The grant holder, Sam Verwimp, will be responsible for data collection, correct documentation and storage onto the KU Leuven servers. The KU Leuven IT department will be responsible for maintenance and back up of the servers.
Who will manage data preservation and sharing?	The grant holder, Sam Verwimp, and the promotor and head of the research group (Prof. Leen Delang) will share responsibility for ensuring data preservation and sharing.
Who will update and implement this DMP?	The grant holder, Sam Verwimp, and the promotor and head of the research group (Prof. Leen Delang) will share responsibility for updating and implementing this DMP.