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	Leen Delang, supervisor-spokesperson, 0000-0002-8874-675X
Contributor name(s) (+ ORCID) & roles	Marie Joossens, supervisor UGent, 0000-0001-8410-5528 Chris Callewaert, co-supervisor UGent, 0000-0001-7697-9188 Niels Verhulst, supervisor-spokesperson University of Zurich, 0000-0002-1106-9711
Project number [1] & title	Impact of skin bacteria on mosquito host finding and on host infection with mosquito-borne viruses
Funder(s) GrantID [2]	G0D0823N
Affiliation(s)	x KU Leuven ☐ Universiteit Antwerpen x Universiteit Gent ☐ Universiteit Hasselt ☐ Vrije Universiteit Brussel x Other: Universiteit Zurich ROR identifier KU Leuven: 05f950310
Please provide a short project description	Mosquitoes can transmit pathogens when they bite us, and these can cause diseases that are a substantial threat to the health of humans and animals worldwide. Some people are more attractive to mosquitoes than others, and this difference is related to how our skin smells. We have collected evidence that much of the smell from our skin, comes in fact from the beneficial bacteria that we have on our skin. Our recent results also show that certain bacteria, characteristic of our skin, can play a significant role in determining the severity of diseases transmitted through a mosquito bite. The goal of this research project is to significantly decrease human and animal attractiveness to mosquitoes by altering the composition of the bacteria on our skin. We will use innovative methodology to selectively kill bacteria that makes us attractive to mosquitoes, and consequently reduce pathogen transmission and disease.

2. Research Data Summary

[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.

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
		☐ Generate new data ☐ Reuse existing data	☐ Digital ☐ Physical	☐ Audiovisual ☐ Images ☐ Sound ☐ Numerical ☐ Textual ☐ Model ☐ Software ☐ Other:		☐ < 1 GB ☐ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA	
Protocols	Written protocols	New data	Digital	Other: descriptive	.docx .pdf	<100 MB	NA
Experiment measurements and observations	Measurements and observations (survival of mice, body weight, footpad swelling, trackin coordinates mosquitoes…)	New data	Digital	Experimental	.xls	<100 MB	NA
Biological samples	All biological samples resulting from	New data	Physical	NA	NA	NA	6000 samples

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

	experiments (cells, bacteria, viruses, mosquito saliva, mouse tissues, human skin biopsies, …)						
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 +	New data	Digital	Compiled	.xls	<100 MB	NA

	calculations						
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 cytometry data in FlowJo Software	New data	Digital	Software	.wsp	<100 MB	NA
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	New data	Digital	Software	.pzfx	<100 MB	NA

	with graphs resulting from analysis of data						
Graph images	Exported images from graphs made in Graphpad software	New data	Digital	Compiled	.tif	<1 GB	NA
Manuscripts	Manuscripts for publications of results	New data	Digital	Compiled	.pdf	<100 MB	NA
Bacteria profiling	Skin bacterial profiling by GC-MS analysis	New data	Digital	Experimental	.raw .cdf .cvs .xls	<1 GB	NA
Video tracking	Raw video files will be collected of mosquitoes flying in the behavioural laboratory setup.	New data	Digital	Audiovisual	.MP4	<20 GB	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.	Not applicable
Are there any ethical issues concerning the creation and/or use of the data (e.g. experiments on humans or animals, dual use)? If so, refer to specific datasets or data types when appropriate and provide the relevant ethical approval number.	☐ Yes, human subject data; provide SMEC or EC approval number: ☒ Yes, animal data; provide ECD reference number: P71-2019 ☐ Yes, dual use; provide approval number: ☐ No Additional information: 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).
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: NA
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

[4] See Glossary Flemish Standard Data Management Plan

Do existing 3rd party agreements restrict exploitation or dissemination of the data you (re)use (e.g. Material/Data transfer agreements, research collaboration agreements)? If so, please explain to what data they relate and what restrictions are in place.	☐ Yes ☒ No If yes, please explain: 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). *RDM guidance on documentation and metadata*.	Daily lab work (protocols, calculations, results,…) will be documented in an online OneNote lab book 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 together with the image. For flow cytometry, FCS files containing metadata will be generated and stored. Video tracking: Both tracking and video files will be stored together with a metadata file describing the experimental setting such as time, date, treatment, location and mosquito species. GC-MS: Each folder for each experiment with .raw files will have a metadata file describing the experimental setting such as time, date, treatment, location and GC-MS settings.

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: Teams folder online

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) 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. All datafiles from the University of Zurich will be saved on backed-up UZH servers and external high-capacity hard drives locked in a cupboard.
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 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.
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*	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). Data is only stored on UZH servers and backup disks located in a cupboard only accessible by the researchers working on the project.

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 funding 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...). *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):
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, 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: 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. Video tracking and GC-MS: All data shall be deposited in a non-commercial, public repository (zenodo). All file formats will comply with the repository. Video files will be kept on external hard drives and will be available upon request.
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:

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) Video tracking data and GC-MS data collected by the University of Zurich shall be deposited in a non-commercial, public repository (zenodo). All file formats will comply with the repository. Video files will be kept on external hard drives and will be available upon request.
When will the data be made available?	☒ Upon publication of research results ☐ Specific date (specify) ☐ Other (specify)
Which data usage licenses are you going to provide? If none, please explain why. *A DATA USAGE LICENSE INDICATES WHETHER THE DATA CAN BE REUSED OR NOT AND UNDER WHAT CONDITIONS. IF NO LICENCE IS GRANTED, THE DATA ARE IN A GREY ZONE AND CANNOT BE LEGALLY REUSED. DO NOTE THAT YOU MAY ONLY RELEASE DATA UNDER A LICENCE CHOSEN BY YOURSELF IF IT DOES NOT ALREADY FALL UNDER ANOTHER LICENCE THAT MIGHT PROHIBIT THAT.* *Check the RDR guidance on licences for data and software sources code or consult the License selector tool to help you choose.*	☐ CC-BY 4.0 (data) ☒ Data Transfer Agreement (restricted data) ☐ MIT licence (code) ☐ GNU GPL-3.0 (code) ☐ Other (specify)
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?	Data will be deposited in a non-commercial, public repositories; therefore, no costs are expected. Key findings of this project will be made available through publication in peer-reviewed journals, which will cost between 2000-5000 € per publication (depending on the journal). We expect to publish several manuscripts, expecting a cost of 10.000 to 15.000 € in total. This is accounted for in the project budget.

7. Responsibilities	
Who will manage data documentation and metadata during the research project?	Leen Delang, Marie Joossens, Chris Callewaert, and Niels Verhulst will be responsible for data and metadata documentation and preservation.
Who will manage data storage and backup during the research project?	Leen Delang, Marie Joossens, Chris Callewaert, and Niels Verhulst will be responsible for data storage. The IT departments of KU Leuven, UG and University of Zurich will be responsible for maintenance and back up of the servers.
Who will manage data preservation and sharing?	Leen Delang, Marie Joossens, Chris Callewaert, and Niels Verhulst will share responsibility for ensuring data preservation and sharing.
Who will update and implement this DMP?	Leen Delang, Marie Joossens, Chris Callewaert, and Niels Verhulst will share responsibility for updating and implementing this DMP.