
DEVENIR: DEFINING VULVOVAGINAL CANDIDIASIS - ELEMENTS OF INFECTION AND REMEDY

A Data Management Plan created using DMPonline.be

Creator: Patrick Van Dijck  <https://orcid.org/0000-0002-1542-897X>

Affiliation: KU Leuven (KUL)

Funder: Fonds voor Wetenschappelijk Onderzoek - Research Foundation Flanders (FWO)

Template: FWO DMP (Flemish Standard DMP)

Principal Investigator: Patrick Van Dijck  <https://orcid.org/0000-0002-1542-897X>

Grant number / URL: S006424N

ID: 206234

Start date: 01-10-2023

End date: 30-09-2027

Project abstract:

Female specific health conditions are underrepresented in research. The taboo felt by women to talk about intimate wellbeing is fed by this gender gap in scientific knowledge. This project aims to meet these needs by studying one of the most prevalent female-specific infections, vulvovaginal candidiasis (VVC), and paving the way towards its efficient diagnosis and treatment. About 70% of women worldwide suffer from vaginal candidiasis at least once in their life. An episode of this fungal infection is accompanied by a burning sensation, pain, and a reduced mental wellbeing. Some women (about 5%) encounter such infections at least four times a year, referring to recurrent(R) VVC. The design of efficient diagnostic and therapeutic strategies for (R)VVC is hindered by a knowledge gap surrounding vaginal health. To meet the absolute need for more information, this project will characterize the role of the microbiome, metabolome, immune system, and pathogen characteristics in (R)VVC. To this end, a large sampling platform of women with/without VVC will be established in this project. We will identify the most important and clinically relevant micro-organisms, metabolites, and immune factors in VVC pathogenesis. VVC models will be developed and optimized, which will be used to validate the causality of the correlations identified in the cohort. Identified correlations will be proposed as biomarkers, and microbes, metabolites, and combinations, which effectively lower pathogenicity of *Candida* species, will be further investigated for therapeutic potential. State-of-the-art tools and know how in the consortium will allow us to unravel the involved molecular pathways and elucidate how these can be exploited to optimize therapeutic efficacy. Finally, the knowledge gathered in this project will be used to improve literacy of women on VVC using platforms established by the consortium as well as novel tools to be developed in this project.

Last modified: 08-04-2024

DEVENIR: DEFINING VULVOVAGINAL CANDIDIASIS - ELEMENTS OF INFECTION AND REMEDY

Application DMP

Questionnaire

Describe the datatypes (surveys, sequences, manuscripts, objects ...) the research will collect and/or generate and /or (re)use. (use up to 700 characters)

A collection of vaginal microbiomes will be obtained. Apart from the complete mixture, individual eukaryotic strains will be isolated and stored.

Excel files containing relevant information about these strains will be generated.

Images of microbial communities will be obtained and stored. All other data obtained during the project will be stored.

Manuscripts will be generated describing the experiments, interpretation and conclusions derived thereof

Isala survey (reviewed by DPO Koen Pepermans, ERC ethics review and UZA/UAntwerpen medisch ethisch committee)

- bacterial DNA sequences: WGS from isolates and 16S amplicon and (shallow) shotgun stored in UZA/UAntwerpen biobank
- fungal DNA sequences: ITS sequencing from isolates and possible WGS for some isolates, stored on server at KU Leuven and at Sciensano, Brussels.

Specify in which way the following provisions are in place in order to preserve the data during and at least 5 years after the end of the research? Motivate your answer. (use up to 700 characters)

1. KU Leuven: Luc Grauwels will be responsible for the transfer of data to Luna, the central KU Leuven storage system.

B. Storage capacity/repository

- during the research

At KU Leuven, Network File Storage, I: (personal) or J: (shared). Cost of 519€/TB/year for type 1 data and 1.137€/TB/year for Type 2 storage

- after the research

At KU Leuven, long-term data storage in ICTS data center (Archive K and L (LVS) drives, 173€/TB/year

Upon publication, data stored in Dryad repository

UAntwerp stores sequences to ENA (public databases) and NCBI. All samples will be registered in the Biobank using the SLIMS system. All fungal isolates will also be stored at Sciensano

What's the reason why you wish to deviate from the principle of preservation of data and of the minimum preservation term of 5 years? (max. 700 characters)

All data that can be stored will be stored on securely backed-up centrally managed file storage, as indicated above for at least 5 years after this project finishes, and most likely indefinitely. Following publication, the results associated with each study will also be deposited in the Dryad repository, where they will be preserved indefinitely.

A batch of the strains will be stored in duplicate in two separate refrigerators in the KU Leuven-MCB and ure isolates are also supplied to Sciensano to be stored in the BCCM collection.

Are there issues concerning research data indicated in the ethics questionnaire of this application form? Which specific security measures do those data require? (use up to 700 characters)

Yes, the research involves personal data collection and thus the specific rules of GDPR will be followed. A procedure has been worked out with Koen Pepermans (DPO) and ERC. It includes pseudonymization of the participants, a password protected website and a dedicated software (Qualtrics) for the questionnaires.

Which other issues related to the data management are relevant to mention? (use up to 700 characters)

Not applicable

DEVENIR: DEFINING VULVOVAGINAL CANDIDIASIS - ELEMENTS OF INFECTION AND REMEDY

FWO DMP (Flemish Standard DMP)

1. 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.

				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
		<i>Please choose from the following options:</i> <ul style="list-style-type: none"> • Generate new data • Reuse existing data 	<i>Please choose from the following options:</i> <ul style="list-style-type: none"> • Digital • Physical 	<i>Please choose from the following options:</i> <ul style="list-style-type: none"> • Observational • Experimental • Compiled/aggregated data • Simulation data • Software • Other • NA 	<i>Please choose from the following options:</i> <ul style="list-style-type: none"> • .por, .xml, .tab, .csv, .pdf, .txt, .rtf, .dwg, .gml, ... • NA 	<i>Please choose from the following options:</i> <ul style="list-style-type: none"> • <100MB • <1GB • <100GB • <1TB • <5TB • <10TB • <50TB • >50TB • NA 	
vaginal swab	vaginal swabs will be taken from women (with VVC, RVVC and healthy)	generate new data	physical	Experimental	.txt; xml; cvs; pdf	NA	several drawers in -80 ° freezer
strain characterisation	bacterial and fungal species will be isolated and identified	generate new data	physical and digital	Experimental	.txt; sequencing data; pdf	<1TB	several drawers in -80 °C freezer
metabolites	metabolite analysis of vaginal swabs	generate new data	digital	Experimental	.txt; pdf;	<1TB	
strain characterisation data	drug susceptibility profiling, anti-Candida activity of yeast isolates;	generate new data	digital	Experimental	.txt; .csv; .pdf	<100Mb	

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 do not reuse previously or published data

Are there any ethical issues concerning the creation and/or use of the data (e.g. experiments on humans or animals, dual use)? Describe these issues in the comment section. Please refer to specific datasets or data types when appropriate.

- Yes, human subject data
- Yes, animal data

Vaginal swabs from women with vulvovaginal infections and the recurrent version of this will be taken.

Animal experiments to show the efficacy of yeast probiotics (or bacterial probiotics) on *Candida albicans* and *Candida glabrata* vaginal infections will be performed

Will you process personal data? If so, briefly describe the kind of personal data you will use in the comment section. Please refer to specific datasets or data types when appropriate.

- 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

In case we identify *Saccharomyces cerevisiae* strains in vaginal swabs of healthy women and we can show that these strains prevent fungal infections, such strains could be used therapeutically. One of the members of the users committee is already supporting a Baekeland student to further explore this possibility.

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 in the comment section to what data they relate and what restrictions are in place.

- Yes

A consortium agreement has been made which stipulates which partner is owner or co-owner of the data. This agreement is available

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 in the comment section to what data they relate and which restrictions will be asserted.

- Yes

A consortium agreement has been made which stipulates which partner is owner or co-owner of the data. This agreement is available

2. 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).

We have a lab guide with standardized protocols that are being done in the lab. In addition, the results will be published and materials and methods will be described in detail so that the experienced researcher can repeat the analysis.

All generated data is stored on the KU Leuven server, which contains protected project directories to which only researchers involved have access. In addition, every researcher has a personal directory on the KU Leuven server for safe data storage and a OneDrive and Teams

directory. Data is never stored on personal or work devices to prevent data loss upon technical failure

Will a metadata standard be used to make it easier to find and reuse the data? If so, please specify (where appropriate per dataset or data type) which metadata standard will be used. If not, please specify (where appropriate per dataset or data type) which metadata will be created to make the data easier to find and reuse.

- No

3. Data storage & back-up during the research project

Where will the data be stored?

All data is stored on the KU Leuven shared server and additional online directories (e.g. personal directory on KU Leuven server, Teams, OneDrive) as described before.

How will the data be backed up?

Standard back-up provided by KU Leuven ICTS

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

We have large volumes available on both I and J drives. Upon payment, we can also get more space if necessary

How will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?

The data will be stored in the university's secure environment

What are the expected costs for data storage and backup during the research project? How will these costs be covered?

The expected cost is 500 euro/TB/year. We do not expect to go over that cost. We add this type of cost in our consumable money for any grant application

4. 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 digital as well as physical material will be stored for 10 years. The physical material (yeast strains, Candida isolates, bacterial isolates) will be stored in a - 80°C freezer and all strains information are well-documented in our filemaker databases

Where will these data be archived (stored and curated for the long-term)?

The data will be archived on the large volume storage and the shared network drive (J-drive).

What are the expected costs for data preservation during the expected retention period? How will these costs be covered?

In view of the expected size of the database (probably less than 1 TB), the estimated cost will be below 500 euro to set up the database and an annual fee of 50 euro for support

5. Data sharing and reuse

Will the data (or part of the data) be made available for reuse after/during the project? In the comment section please explain per dataset or data type which data will be made available.

- Yes, in an Open Access repository

The physical material can be requested and will be sent to the other lab, unless a specific strain may be used for commercialisation. The digital information will be available upon request and/or may be available. Sciensano will WGS the fungal strains and probably provide this information to any customer.

If access is restricted, please specify who will be able to access the data and under what conditions.

Basically, access is not restricted, unless a company takes a patent on one of the strains for a specific application (e.g. prevention/cure vulvovaginitis).

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 in the comment section per dataset or data type where appropriate.

- Yes, Intellectual Property Rights

The U Antwerp will mainly focus on the bacterial isolates, the KU Leuven on the fungal isolates. In case we find that a mix of both types of strains is required for an optimal treatment, then negotiations, as agreed upon in the consortium agreement, need to take place

Where will the data be made available? If already known, please provide a repository per dataset or data type.

The U Antwerp has the ISALA platform, but the last work package is completely on awareness of vaginal infections, so several channels (for specialists or for the public at large) will be informed about the data.

The data will also be made available through publications and at international conferences. The publications will be the source of information where interested researchers can trace back a strain and ask this strain from the lab.

All digital data will be made available through the Research data repository.

When will the data be made available?

Upon publication (or patenting) of the research results

Which data usage licenses are you going to provide? If none, please explain why.

Data from the project that can be shared will be made available under a Creative Commons Attribution licence (CC-BY 4.0), so that the users have to give credit to the original data creators

Do you intend to add a PID/DOI/accession number to your dataset(s)? If already available, you have the option to provide it in the comment section.

- Yes

All publications will receive a DOI.

A PID will be added upon deposit of data in a repository.

What are the expected costs for data sharing? How will these costs be covered?

Sharing data via NCBI is free of charge. Data present in manuscripts will cost the cost of such publications, which may go up to more than 7000 euro for papers in the Nature group (currently 10500 euro for Nature Microbiology). If other researchers request strains, we normally ask them to provide us with a DHL or FedEx number.

Costs associated with a patent application (10000 euro) will be paid but the aim here is to find a company that takes a license on such patent

6. Responsibilities

Who will manage data documentation and metadata during the research project?

Patrick Van Dijck (KU Leuven); Paul Vandecruys (KU Leuven); Sarah Lever (U Antwerp); Mark Gresnigt (HKI Jena); Gilbert Donders (UZA)

Who will manage data storage and backup during the research project?

Luc Grauwels (KU Leuven)

Who will manage data preservation and sharing?

Patrick Van Dijck (KU Leuven); Sarah Lever (U Antwerp); Mark Gresnigt (HKI Jena)

Who will update and implement this DMP?

Patrick Van Dijck (KU Leuven)