FWO DMP Template

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.

1. General Information	
Name applicant	**Fabienne Ver Donck**
FWO Project Number & Title	aspirant beurs 1115222N: Genomics and disease modelling for inherited platelet disorders.
Affiliation	☒ KU Leuven ☐ Universiteit Antwerpen ☐ Universiteit Gent ☐ Universiteit Hasselt ☐ Vrije Universiteit Brussel ☐ Other:
2. Data description	
Will you generate/collect new data and/or make use of existing data?	☒ Generate new data ☒ Reuse existing data

Describe the origin, type and format of the data (per dataset) and its (estimated) volume

*If you **reuse** existing data, specify the **source** of these data.*
*Distinguish data **types** (the kind of content) from data **formats** (the technical format).*

Data Origin	**Type**	**Format**	**Estimated volume**
Whole genome sequencing of 342 patients with bleeding, platelet and trombotic disorders, recruited from the UZ Leuven. These data are stored on protected server of University of Cambridge-UCAM, Freson K (promotor) has access to this server.	WGS data	bam	20 TB
RNAseq data of platelets, monocytes, neutrophils and T-cell for patients with an inherited platelet disorder and healthy controls	RNAseq data	Fastq, bam, count, txt, xlsx	1 TB
Flow Cytometry and FACS sort files (FlowJo and equipment specific files)	Flow cytometry data	Fcs, pdf, jpg, tiff	1 GB
Imaging and quantification files from BioTek Citation device (KU Leuven servers).	Gen5 program files	xml, tiff, wmv, mp4	2 MB/image
Confocal microscopy	Images	Jpg, tiff	Approx. 1MB/image
Data analysis and manuscript preparation	Text documents, images	Xml, docx, pdf, xlsx, pdf, jpg, tiff, pptx	5 GB

3. Ethical and legal issues

Will you use personal data? If so, shortly describe the kind of personal data you will use AND add the reference to your file in your host institution's privacy register. *In case your host institution does not (yet) have a privacy register, a reference is not yet required of course; please add the reference once the privacy register is in place in your host institution.*	☒ Yes ☐ No If yes: - Privacy Registry Reference: - Short description of the kind of personal data that will be used: Whole genome sequencing, RNA sequencing and personal metadata (age, gender, condition, blood counts) will be collected
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, add the reference to the formal approval by the relevant ethical review committee(s).	☒ Yes ☐ No If yes: - Reference to ethical committee approval: S63666
Does your work possibly result in research data with potential for tech transfer and valorisation? Will IP restrictions be claimed for the data you created? If so, for what data and which restrictions will be asserted?	☒ Yes ☐ No If yes, please comment: If data will be obtained of interest for valorisation, IP restriction will be claimed. It is not clear from the start what novel genetic targets relevant for megakaryopoeisis and platelet formation/function can be identified.
Do existing 3rd party agreements restrict dissemination or exploitation of the data you (re)use? If so, to what data do they relate and what restrictions are in place?	☒ Yes ☐ No If yes, please comment: The megakaryocyte model itself "imMKCL" cannot be protected by IP restrictions as this model was developed by our collaborators and we have signed an MTA (with dr K Eto, Kyoto university). If this model is needed for the overall IP restriction, a joined application is a possibility (as stated in the MTA)

4. Documentation and metadata	
What documentation will be provided to enable understanding and reuse of the data collected/generated in this project?	**Generated sequencing data will be uploaded to EGA in combination with related metadata (e.g. age, gender, case/control status, sequencing platform/library… etc.) to be accessible to the public. Flow cytometry and sorting: information on gating strategy for cell identification and sorting will be saved in electronic files with details on antibody concentrations and protocols for cell preparation and staining will be described in detail in lab books.**
Will a metadata standard be used? If so, describe in detail which standard will be used. If not, state in detail which metadata will be created to make the data easy/easier to find and reuse.	☒ Yes ☐ No If yes, please specify: The metadata standards of EGA will be used for submission of sequencing data, as can be consulted on https://ega-archive.org/submission/sequence/unaligned

5. Data storage & backup during the FWO project	
Where will the data be stored?	PC owned by the group
How will the data be backed up?	Double backup on: KU Leuven OneDrive, external SSD drive
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: external SSD capacity of 2 TB, KU Leuven OneDrive capacity of 2 TB (can be further increased) ☐ No If no, please specify:
What are the expected costs for data storage and backup during the project? How will these costs be covered? *Although FWO has no earmarked budget at its disposal to support correct research data management, FWO allows for part of **the allocated project budget** to be used to cover the cost incurred.*	**Data storage costs: OneDrive is included for all employees of KU Leuven (no additional storage cost), J drive storage (add. Backup) costs 52 euro per year.**

Data security: how will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?	**Since we will be working with sensitive personal data that will only be anonymised at the end of the project, the data will be stored in the university's secure environment for private data and only shared with other researchers involved in the project. Sequencing data stored on the external SSD card will be encrypted and protected with a password.**

6. Data preservation after the end of the FWO project FWO expects that data generated during the project are retained for a period of minimally 5 years after the end of the project, in as far as legal and contractual agreements allow.	
Which data will be retained for the expected 5 ear period after the end of the project? In case only a selection of the data can/will be preserved, clearly state the reasons for this (legal or contractual restrictions, physical preservation issues, ...).	**Data that will no longer be used for analyses will be transfer from our active KULeuven servers to the archive KULeuven-servers after 10 years. No data will be disposed.**
Where will these data be archived (= stored for the long term)?	**The data will be stored on the university's central archive servers (with automatic backup procedures) for at least 10 years, conform the KU Leuven RDM policy.**
What are the expected costs for data preservation during these 5 years? How will the costs be covered? *Although FWO has no earmarked budget at its disposal to support correct research data management, FWO allows for part of **the allocated project budget** to be used to cover the cost incurred.*	**Permanent storage after the project for at least 10 years after the end of the project: (Archive storage K-drive) €5,69 per 100GB per year.**

7. Data sharing and reuse

Are there any factors restricting or preventing the sharing of (some of) the data (e.g. as defined in an agreement with a 3rd party, legal restrictions)?	☒ Yes ☐ No If yes, please specify: **Patients and healthy controls have consented for data storage and the release of anonymous sequencing data in the public domain when used for publications. At any moment they can withdraw from the study and their data will no longer be used. Medical data will be recorded for the patients. All medical and genetic data will only be stored on either UZLeuven or KU Leuven controlled servers to ensure optimal data security. No local copies will be kept during or after the research is completed.**
Which data will be made available after the end of the project?	**Raw sequencing data will be submitted to EGA, analysis data will be published in manuscripts.**
Where/how will the data be made available for reuse?	☐ In an Open Access repository ☐ In a restricted access repository ☐ Upon request by mail ☒ Other (specify): EGA, where access to data is granted based on applications to a data access committee that oversees the dataset.
When will the data be made available?	**After publication of the research results in a peer-reviewed journal**
Who will be able to access the data and under what conditions?	**Researchers can request access through the data access committee. Access is granted based on motivation in the application and will be limited to disease specific research.**
What are the expected costs for data sharing? How will these costs be covered? *Although FWO has no earmarked budget at its disposal to support correct research data management, FWO allows for part of **the allocated project budget** to be used to cover the cost incurred.*	**Data will be released after publication of a manuscript and data sharing costs will be covered by the publication cost.**

8. Responsibilities

Who will be responsible for the data documentation & metadata?	**PI Kathleen Freson, PhD student Fabienne Ver Donck**
Who will be responsible for data storage & back up during the project?	**PIs Kathleen Freson, Veerle Labarque** **Phd student Fabienne Ver Donck** **Technician Chantal Thys**
Who will be responsible for ensuring data preservation and sharing?	**PI Kathleen Freson**
Who bears the end responsibility for updating & implementing this DMP? *Default response: The PI bears the overall responsibility for updating & implementing this DMP*	**The PI bears the end responsibility of updating & implementing this DMP.**