

## 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	<b>Lisa Ehlers, ORCID 0000-0001-8737-001X</b>
Contributor name(s) (+ ORCID) & roles	
Project number <sup>1</sup> & title	Understanding ADA2 deficiency on a cellular level
Funder(s) GrantID <sup>2</sup>	11E0123N
Affiliation(s)	<p><input checked="" type="checkbox"/> KU Leuven</p> <p><input type="checkbox"/> Universiteit Antwerpen</p> <p><input type="checkbox"/> Universiteit Gent</p> <p><input type="checkbox"/> Universiteit Hasselt</p> <p><input type="checkbox"/> Vrije Universiteit Brussel</p> <p><input type="checkbox"/> Other:</p> <p>Provide ROR<sup>3</sup> identifier when possible: 05f950310 (KU Leuven), 03qtxy027 (FWO)</p>
Please provide a short project description	<p>ADA2 deficiency (DADA2) is a monogenic disease caused by biallelic mutations in the ADA2 gene that presents with vasculopathy and haemato-immunological symptoms as well as an increased type I interferon signature. The cellular pathophysiology of DADA2 is poorly understood.</p> <p>In this project, I aim to characterise DADA2 on a cellular level and analyse cellular disturbances in the presence of mutant ADA2 in the context of the underlying genotype.</p> <p>To this end, I will (I.) create a thorough data set of DADA2 patients uniting information on genotype, in vitro characteristics and clinical manifestations of different ADA2 variants, (II.) characterise the subcellular localisation and trafficking of wild-type and mutant ADA2, and (III.) examine the role of proteotoxic stress and defective autophagy in DADA2 pathophysiology.</p>

<sup>1</sup> "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.

<sup>2</sup> 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.

<sup>3</sup> Research Organization Registry Community. <https://ror.org/>



## 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<sup>4</sup>.

Dataset Name	Description	New or Reused	Digital or Physical	ONLY FOR DIGITAL DATA	ONLY FOR DIGITAL DATA	ONLY FOR DIGITAL DATA	ONLY FOR PHYSICAL DATA
				Digital Data Type	Digital Data Format	Digital Data Volume (MB, GB, TB)	Physical Volume
WP1 : Characterisation of genetic ADA2 variants and their association with the clinical and immunological phenotype of ADA2 deficiency							
1.1 Clinical review of DADA2 patients	- systematic literature review summarising clinical characteristics of all published cases of DADA2 (excel file) - detailed review of the clinical characteristics including symptoms, treatment, disease flares of the UZ Leuven DADA2 cohort (excel file, patients anonymised using established lab codes) - results of the ADA2 enzyme assay (excel file) and ADA2	<input checked="" type="checkbox"/> Generate new data <input checked="" type="checkbox"/> Reuse existing data	<input checked="" type="checkbox"/> Digital <input type="checkbox"/> Physical	<input checked="" type="checkbox"/> Observational <input checked="" type="checkbox"/> Experimental <input checked="" type="checkbox"/> Compiled/aggregated data <input type="checkbox"/> Simulation data <input type="checkbox"/> Software <input type="checkbox"/> Other <input type="checkbox"/> NA	<input type="checkbox"/> .por <input type="checkbox"/> .xml <input type="checkbox"/> .tab <input type="checkbox"/> .csv <input type="checkbox"/> .pdf <input type="checkbox"/> .txt <input type="checkbox"/> .rtf <input type="checkbox"/> .dwg <input type="checkbox"/> .tab <input type="checkbox"/> .gml <input checked="" type="checkbox"/> other: .xlsx; .ab1 <input type="checkbox"/> NA	<input checked="" type="checkbox"/> < 100 MB <input type="checkbox"/> < 1 GB <input type="checkbox"/> < 100 GB <input type="checkbox"/> < 1 TB <input type="checkbox"/> < 5 TB <input type="checkbox"/> < 10 TB <input type="checkbox"/> < 50 TB <input type="checkbox"/> > 50 TB <input type="checkbox"/> NA	

<sup>4</sup> Add rows for each dataset you want to describe.

	sequencing (.ab1)						
1.2 In-depth analysis of genotype-phenotype correlations	- type I interferon signature from whole blood, PBMCs, EBV-LCLs, fibroblasts (qPCR files .edt / .eds) - ADA2 mRNA expression (qPCR files .edt / .eds) - measurements from whole blood are completed longitudinally and data from previous time points are already available	<input checked="" type="checkbox"/> Generate new data <input type="checkbox"/> Reuse existing data	<input checked="" type="checkbox"/> Digital <input type="checkbox"/> Physical	<input type="checkbox"/> Observational <input checked="" type="checkbox"/> Experimental <input type="checkbox"/> Compiled/aggregated data <input type="checkbox"/> Simulation data <input type="checkbox"/> Software <input type="checkbox"/> Other <input type="checkbox"/> NA	<input type="checkbox"/> .por <input type="checkbox"/> .xml <input type="checkbox"/> .tab <input type="checkbox"/> .csv <input type="checkbox"/> .pdf <input type="checkbox"/> .txt <input type="checkbox"/> .rtf <input type="checkbox"/> .dwg <input type="checkbox"/> .tab <input type="checkbox"/> .gml <input checked="" type="checkbox"/> other: .edt/.eds <input type="checkbox"/> NA	<input type="checkbox"/> < 100 MB <input checked="" type="checkbox"/> < 1 GB <input type="checkbox"/> < 100 GB <input type="checkbox"/> < 1 TB <input type="checkbox"/> < 5 TB <input type="checkbox"/> < 10 TB <input type="checkbox"/> < 50 TB <input type="checkbox"/> > 50 TB <input type="checkbox"/> NA	
	-whole blood samples collected in PAXgene tube from DADA2 patients and healthy controls -PBMC, EBV-LCL and fibroblasts cryosamples from DADA2 patients and healthy controls -sample collection has been ongoing since 2014, both stored and newly prepared samples will be used	<input checked="" type="checkbox"/> Generate new data <input checked="" type="checkbox"/> Reuse existing data	<input type="checkbox"/> Digital <input checked="" type="checkbox"/> Physical	<input type="checkbox"/> Observational <input checked="" type="checkbox"/> Experimental <input type="checkbox"/> Compiled/aggregated data <input type="checkbox"/> Simulation data <input type="checkbox"/> Software <input type="checkbox"/> Other <input type="checkbox"/> NA	<input type="checkbox"/> .por <input type="checkbox"/> .xml <input type="checkbox"/> .tab <input type="checkbox"/> .csv <input type="checkbox"/> .pdf <input type="checkbox"/> .txt <input type="checkbox"/> .rtf <input type="checkbox"/> .dwg <input type="checkbox"/> .tab <input type="checkbox"/> .gml <input type="checkbox"/> other: .edt/.eds <input checked="" type="checkbox"/> NA	<input type="checkbox"/> < 100 MB <input type="checkbox"/> < 1 GB <input type="checkbox"/> < 100 GB <input type="checkbox"/> < 1 TB <input type="checkbox"/> < 5 TB <input type="checkbox"/> < 10 TB <input type="checkbox"/> < 50 TB <input type="checkbox"/> > 50 TB <input checked="" type="checkbox"/> NA	- PAXgene tubes, approx. 100 samples, stored at -20°C temporarily until RNA extraction - RNA samples (1.5 mL Eppendorf tubes), approx. 500 samples, stored at -80°C in boxes comprising 81 samples each

	- EBV-LCL cell lines are created by the CME (UZ Leuven), an additional cryostock of each cell line is stored in their facilities						- cryosamples (approx. 100 samples) are stored in 1.8 mL cryotubes in the liquid nitrogen tank of the lab
1.3 Differences in gene expression profiles in ADA2-deficient immune cells	-bulk RNA-seq files from whole blood from healthy controls and DADA2 patients - bulk RNA-seq files from whole blood from healthy controls and DADA2 patients - analysed data stored in excel format -scRNA-seq data sets published by Watanabe et al. (doi: 10.1002/JLB.3HI0220-119RR) and Wu et al. (doi: 10.1002/JLB.5A0621-314R) deposited on GEO: GSE142444 and GSE168163	<input checked="" type="checkbox"/> Generate new data <input checked="" type="checkbox"/> Reuse existing data	<input checked="" type="checkbox"/> Digital <input type="checkbox"/> Physical	<input type="checkbox"/> Observational <input checked="" type="checkbox"/> Experimental <input type="checkbox"/> Compiled/aggregated data <input type="checkbox"/> Simulation data <input type="checkbox"/> Software <input type="checkbox"/> Other <input type="checkbox"/> NA	<input type="checkbox"/> .por <input type="checkbox"/> .xml <input type="checkbox"/> .tab <input type="checkbox"/> .csv <input type="checkbox"/> .pdf <input type="checkbox"/> .txt <input type="checkbox"/> .rtf <input type="checkbox"/> .dwg <input type="checkbox"/> .tab <input type="checkbox"/> .gml <input checked="" type="checkbox"/> other: .fastq.gz; .xlsx <input type="checkbox"/> NA	<input type="checkbox"/> < 100 MB <input type="checkbox"/> < 1 GB <input checked="" type="checkbox"/> < 100 GB <input type="checkbox"/> < 1 TB <input type="checkbox"/> < 5 TB <input type="checkbox"/> < 10 TB <input type="checkbox"/> < 50 TB <input type="checkbox"/> > 50 TB <input type="checkbox"/> NA	
	-Trizol samples from untreated vs. treated CD14+ monocytes from	<input checked="" type="checkbox"/> Generate new data <input checked="" type="checkbox"/> Reuse	<input type="checkbox"/> Digital <input checked="" type="checkbox"/> Physical	<input type="checkbox"/> Observational <input checked="" type="checkbox"/> Experimental <input type="checkbox"/> Compiled/	<input type="checkbox"/> .por <input type="checkbox"/> .xml <input type="checkbox"/> .tab	<input type="checkbox"/> < 100 MB <input type="checkbox"/> < 1 GB <input type="checkbox"/> < 100 GB	-Trizol samples (approx. 100) are stored in 1.5 mL

	healthy controls and DADA2 patients -for whole blood samples see 1.2	existing data		aggregated data <input type="checkbox"/> Simulation data <input type="checkbox"/> Software <input type="checkbox"/> Other <input type="checkbox"/> NA	<input type="checkbox"/> .csv <input type="checkbox"/> .pdf <input type="checkbox"/> .txt <input type="checkbox"/> .rtf <input type="checkbox"/> .dwg <input type="checkbox"/> .tab <input type="checkbox"/> .gml <input type="checkbox"/> other: <input checked="" type="checkbox"/> NA	<input type="checkbox"/> < 1 TB <input type="checkbox"/> < 5 TB <input type="checkbox"/> < 10 TB <input type="checkbox"/> < 50 TB <input type="checkbox"/> > 50 TB <input checked="" type="checkbox"/> NA	Eppendorf tubes at -80°C
WP2: Subcellular localisation, trafficking and function of ADA2							
2.1 Subcellular localisation of ADA2	- microscopy slides - whole cell extracts & subcellular fractions	<input checked="" type="checkbox"/> Generate new data <input type="checkbox"/> Reuse existing data	<input type="checkbox"/> Digital <input checked="" type="checkbox"/> Physical	<input type="checkbox"/> Observational <input checked="" type="checkbox"/> Experimental <input type="checkbox"/> Compiled/aggregated data <input type="checkbox"/> Simulation data <input type="checkbox"/> Software <input type="checkbox"/> Other <input type="checkbox"/> NA	<input type="checkbox"/> .por <input type="checkbox"/> .xml <input type="checkbox"/> .tab <input type="checkbox"/> .csv <input type="checkbox"/> .pdf <input type="checkbox"/> .txt <input type="checkbox"/> .rtf <input type="checkbox"/> .dwg <input type="checkbox"/> .tab <input type="checkbox"/> .gml <input type="checkbox"/> other: <input checked="" type="checkbox"/> NA	<input type="checkbox"/> < 100 MB <input type="checkbox"/> < 1 GB <input type="checkbox"/> < 100 GB <input type="checkbox"/> < 1 TB <input type="checkbox"/> < 5 TB <input type="checkbox"/> < 10 TB <input type="checkbox"/> < 50 TB <input type="checkbox"/> > 50 TB <input checked="" type="checkbox"/> NA	-approx. 100 slides, stored in slide box at room temperature - approx. 100 samples temporarily stored at -80°C before western blotting (1-2 boxes)
	- microscopy images: EVOS microscope for optimisation, confocal microscopy for final images (raw files in tif format; png files after editing) - ImageStream files (.rif or .fcs)	<input checked="" type="checkbox"/> Generate new data <input type="checkbox"/> Reuse existing data	<input checked="" type="checkbox"/> Digital <input type="checkbox"/> Physical	<input type="checkbox"/> Observational <input checked="" type="checkbox"/> Experimental <input type="checkbox"/> Compiled/aggregated data <input type="checkbox"/> Simulation data <input type="checkbox"/> Software <input type="checkbox"/> Other <input type="checkbox"/> NA	<input type="checkbox"/> .por <input type="checkbox"/> .xml <input type="checkbox"/> .tab <input type="checkbox"/> .csv <input type="checkbox"/> .pdf <input type="checkbox"/> .txt <input type="checkbox"/> .rtf <input type="checkbox"/> .dwg <input type="checkbox"/> .tab	<input type="checkbox"/> < 100 MB <input type="checkbox"/> < 1 GB <input checked="" type="checkbox"/> < 100 GB <input type="checkbox"/> < 1 TB <input type="checkbox"/> < 5 TB <input type="checkbox"/> < 10 TB <input type="checkbox"/> < 50 TB <input type="checkbox"/> > 50 TB <input type="checkbox"/> NA	

<p><b>GUIDANCE:</b></p> <p>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.</p> <p>EXAMPLES OF DATA TYPES: OBSERVATIONAL (E.G. SURVEY RESULTS, SENSOR READINGS, SENSORY OBSERVATIONS); EXPERIMENTAL (E.G. MICROSCOPY, SPECTROSCOPY, CHROMATOGRAMS, GENE SEQUENCES); COMPILED/AGGREGATED DATA<sup>5</sup> (E.G. TEXT &amp; DATA MINING, DERIVED VARIABLES, 3D MODELLING); SIMULATION DATA (E.G. CLIMATE MODELS); SOFTWARE, ETC.</p> <p>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 &amp; COMPUTATIONAL SCRIPT.</p> <p>DIGITAL DATA VOLUME: PLEASE ESTIMATE THE UPPER LIMIT OF THE VOLUME OF THE DATA PER DATASET OR DATA TYPE.</p> <p>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).</p>	
<p>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.</p>	<p>scRNA-seq data sets published by Watanabe et al. (doi: 10.1002/JLB.3HI0220-119RR) and Wu et al. (doi: 10.1002/JLB.5A0621-314R) deposited on GEO: GSE142444 and GSE168163</p>

---

<sup>5</sup> These data are generated by combining multiple existing datasets.



<p>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.</p>	<p><input checked="" type="checkbox"/> Yes, human subject data  <input type="checkbox"/> Yes, animal data  <input type="checkbox"/> Yes, dual use  <input type="checkbox"/> No</p> <p>If yes, please describe:  The in-depth description of the clinical phenotype (WP 1.1) requires the collection of sensitive patient data. All patients are given a lab code and their data will be analysed and stored anonymously. The patients are pseudonymized so that the patient's identity can be traced back with the applicable key. Patient samples and patient-derived cell lines used in WP1-3 will equally be labelled with the respective lab code for the purpose of anonymization. Enrolled patients are thoroughly informed about this study project and we have signed informed consent from every patient/legal guardian. Approval has been granted by the ethical committee (S61245 / S63077 / S63807).</p>
<p>Will you process personal data<sup>6</sup>? 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.</p>	<p><input checked="" type="checkbox"/> Yes  <input type="checkbox"/> No</p> <p>If yes:</p> <ul style="list-style-type: none"> <li>- Short description of the kind of personal data that will be used: We will collect sensitive patient data including gender, age, genotype, disease onset, symptoms, flares, treatment and family history available in KWS/LWS as well as self-reported during consultation service. All data will be anonymized for our analyses. We will collect and use patient derived material (as listed above to be used in WP 1.2, 1.3, 2.2, 3.1, 3.3, 3.4) within our biobank/study number S61245 / S63077 / S63807 for molecular diagnosis, in depth immunophenotyping, in vitro immune cell compartment evaluation, pathway analysis</li> <li>- Privacy Registry Reference: GDPR questionnaire of UZ Leuven was completed for our projects with Biobank/Studynumber S61245 / S63077 / S63807.</li> </ul>

<sup>6</sup> See Glossary Flemish Standard Data Management Plan

<p>Does your work have potential for commercial valorization (e.g. tech transfer, for example spin-offs, commercial exploitation, ...)?</p> <p>If so, please comment per dataset or data type where appropriate.</p>	<p><input type="checkbox"/> Yes</p> <p><input checked="" type="checkbox"/> No</p> <p>If yes, please comment:</p>
<p>Do existing 3rd party agreements restrict exploitation or dissemination of the data you (re)use (e.g. Material/Data transfer agreements, research collaboration agreements)?</p> <p>If so, please explain to what data they relate and what restrictions are in place.</p>	<p><input checked="" type="checkbox"/> Yes</p> <p><input type="checkbox"/> No</p> <p>If yes, please explain: MDTA is set up with VIB Genetics Core facility for RNA-seq. There are however no restrictions on the data dissemination / exploitation beside prior notification given the third party.</p>
<p>Are there any other legal issues, such as intellectual property rights and ownership, to be managed related to the data you (re)use?</p> <p>If so, please explain to what data they relate and which restrictions will be asserted.</p>	<p><input type="checkbox"/> Yes</p> <p><input checked="" type="checkbox"/> No</p> <p>If yes, please explain:</p>

### 3. Documentation and Metadata

<p>Clearly describe what approach will be followed to capture the accompanying information necessary to keep <b>data understandable and usable</b>, 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).</p>	<p>Every lab member of the inborn errors of immunity lab has an account of the electronic lab book provider <i>LabArchives</i>. LabArchives is used for detailed documentation of all experiments including date of experiment, samples used, patient-derived material (anonymised using lab codes), reference and lot numbers of used reagents. The names of the corresponding files containing the experiments' results will be listed for every performed experiment to ensure traceability. Detailed sample documentation is available in the lab's shared folder on the KU Leuven shared drive (J: GBW-0496_Inborn_Errors_Immunolgy) including lab codes and localisation in the respective fridges, freezers and tanks. A uniform labelling system is in place for patient-derived material. For specific experiments, label legend are provided in the researchers' lab books. Standard operating procedures are in place for all protocols regularly used in the lab. They have been verified by to independent researchers and are stored in the lab's shared LabArchives notebook. A back-up of the protocols is kept on the J: drive. Code used to analyse transcriptomics and proteomics data is documented by the lab's bioinformatician and stored on an internal lab website to ensure lasting transparency.</p>
<p>Will a metadata standard be used to make it easier to <b>find and reuse the data</b>?</p> <p>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.</p> <p><i>REPOSITORIES COULD ASK TO DELIVER METADATA IN A CERTAIN FORMAT, WITH SPECIFIED ONTOLOGIES AND VOCABULARIES, I.E. STANDARD LISTS WITH UNIQUE IDENTIFIERS.</i></p>	<p><input type="checkbox"/> Yes  <input checked="" type="checkbox"/> No</p> <p>If yes, please specify (where appropriate per dataset or data type) which metadata standard will be used:</p> <p>If no, please specify (where appropriate per dataset or data type) which metadata will be created:  For every analysed sample the following information are documented in the researcher's LabArchives notebook and in excel sheets on the shared lab folder (J: drive): lab code, sample type, date, stimulating condition. Detailed information on the experimental procedures used to generate and process the samples will be available in the LabArchives entry of the respective experiment or by reference to the used SOP where applicable. For samples sent out for analysis, excel sheets containing the sample information and used sample identifiers are kept in the shared J: folder.</p> <p>Mass spectrometry data will be deposited for open access to the ProteomeXchange Consortium via the PRIDE partner repository.</p>

#### 4. Data Storage & Back-up during the Research Project

Where will the data be stored?

Observational and experimental data is stored in respective project folders on the shared drive (J:) of the inborn errors of immunity lab (GBW-0496\_Inborn\_Errors\_Immunolgy). The shared folder contains project folders with subfolder dedicated to the different work packages and individual experiments. Large data sets (i.e. bulk RNA-sequencing, proteomics) are stored on the Large storage (L:). Additionally, OneDrive is used to synchronise data between different computers and instruments in the laboratory to allow for easy data exchange.

For storage of physical samples, there is storage capacity in the lab (4 fridges, 5 freezers at -20°C, 1 freezer at -80°C, 1 liquid nitrogen tank).

<p><b>How will the data be backed up?</b></p> <p><i>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.<sup>7</sup></i></p> <p><i>REFER TO INSTITUTION-SPECIFIC POLICIES REGARDING BACKUP PROCEDURES WHEN APPROPRIATE.</i></p>	<p>Data backup is provided by the KU Leuven IT department. Backups of project data stored via the KU Leuven network are generated using snapshot technology. The following backup regimes are in place for the respective folders:</p> <p><b>KU Leuven shared drives (J:)</b></p> <ul style="list-style-type: none"> <li>- Backups are made using “snapshot” technology, which is the online storage of incremental data changes.</li> <li>- An hourly backup (at 8 AM, 12 PM, 4 PM and 8 PM) the last 6 of which are stored on our servers</li> <li>- A daily backup, at midnight, the last 6 of which are stored on our servers</li> <li>- A weekly backup, Saturday night at midnight, the last 12 of which are stored on our servers</li> </ul> <p><b>KU Leuven Large Storage (L:)</b></p> <p>Automatic version management of the files. Version management is done using "snapshot" technology, where the previous versions of the changed files are kept online in a snapshot on the same storage system.</p> <ul style="list-style-type: none"> <li>- by default, 1 snapshot is taken daily and is kept for 14 days. So you can go back to previous versions of the file up to 14 days.</li> <li>- end users can restore older files themselves from within their Windows PC via the "previous versions   previous versions" functionality.</li> </ul> <p>For non-digital data, loss of samples is prevented by connection of the lab’s large freezing devices to the KU Leuven surveillance system.</p>
--	---

<sup>7</sup> Source: Ghent University Generic DMP Evaluation Rubric: <https://osf.io/2z5g3/>

<p>Is there currently sufficient storage &amp; 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.</p>	<p><input checked="" type="checkbox"/> Yes  <input type="checkbox"/> No</p> <p>If yes, please specify concisely:  OneDrive: 2 TB / researcher  Shared drive (J:): 1 TB  Large storage (L:): 5 TB</p> <p>If no, please specify:</p>
<p>How will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?</p> <p><i>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. <sup>7</sup></i></p>	<p>Data security on KU Leuven servers including OneDrive, shared drives and large storage options is guaranteed by ICTS to be suitable for public, confidential and strictly confidential data. The OneDrive folders are additionally protected by multifactor authentication with the KU Leuven Authenticator app. For additional data security, all patient information are saved in an anonymized manner using the laboratory's lab code system.</p>
<p>What are the expected costs for data storage and backup during the research project? How will these costs be covered?</p>	<p>KU Leuven provides free access to OneDrive for staff and students.</p> <p>The inborn errors of immunity lab has purchased the following storage options on the KU Leuven network:  Large storage: € 569 / year for 5 TB storage  Shared drive: € 519 / year for 1 TB storage</p> <p>Costs for data storage and backup are covered by the ERC Grant "MORE2ADA2".</p>

## 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...).	Raw data will be stored for more than 5 years. For storage of large data sets (e.g transcriptomics analyses) large storage access is available in the lab. Analysed datasets including analytical excel and Graphpad Prism files are stored by the individual researchers. After project completion (PhD), excess data is removed and valuable data remains under the name the fellow in the dedicated folder on the KU Leuven shared drive for a minimum of 5 years to allow re-analysis of all published work if needed.
Where will these data be archived (stored and curated for the long-term)?	Data will remain on the KU Leuven shared drive (J:) and large storage (L:) as specified above.
What are the expected costs for data preservation during the expected retention period? How will these costs be covered?	Costs for storage drives of the KU Leuven network are as follows:  The inborn errors of immunity lab has purchased the following storage options on the KU Leuven network: Large storage: € 569 / year for 5 TB storage Shared drive: € 519 / year for 1 TB storage  Costs for data storage and backup are covered by the ERC Grant "MORE2ADA2".

## 6. Data Sharing and Reuse

<p>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.</p> <p><i>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 &amp; RESTRICTED ACCESS. FOR MORE INFORMATION: <a href="https://wiki.surfnet.nl/display/standards/info-eu-repo/#infoeu-repo-accessrights">https://wiki.surfnet.nl/display/standards/info-eu-repo/#infoeu-repo-accessrights</a></i></p>	<p><input checked="" type="checkbox"/> Yes, in an Open Access repository  <input type="checkbox"/> Yes, in a restricted access repository (after approval, institutional access only, ...)  <input type="checkbox"/> No (closed access)  <input type="checkbox"/> Other, please specify:</p> <p>Part of the raw data will be made available in the publishing process as required by good scientific practice. This includes in particular – but not exclusively – RNA-sequencing, proteomics and metabolomics data sets. Data will be deposited for open access to the ProteomeXchange Consortium via the PRIDE partner repository (proteomics), the GEO genomics data repository (RNA-sequencing) or Zenodo, an open repository developed under the European OpenAIRE program and operated by CERN. All data sets will be published in an anonymised format.</p>
<p>If access is restricted, please specify who will be able to access the data and under what conditions.</p>	<p>Raw data that will not be made publicly available in the publishing process will be available upon request. Access will be managed by Prof. Isabelle Meyts (PI) and Dr. Leen Moens (senior postdoc).</p>
<p>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.</p>	<p><input checked="" type="checkbox"/> Yes, privacy aspects  <input type="checkbox"/> Yes, intellectual property rights  <input type="checkbox"/> Yes, ethical aspects  <input type="checkbox"/> Yes, aspects of dual use  <input type="checkbox"/> Yes, other  <input type="checkbox"/> No</p> <p>If yes, please specify: All patient information will only be handled and published in an anonymised form to comply with privacy requirements.</p>



<p>Where will the data be made available? If already known, please provide a repository per dataset or data type.</p>	<p>Data will be deposited for open access to the ProteomeXchange Consortium via the PRIDE partner repository (proteomics), the GEO genomics data repository (RNA-sequencing) or Zenodo, an open repository developed under the European OpenAIRE program and operated by CERN. All data sets will be published in an anonymised format.</p>
<p>When will the data be made available?</p> <p><i>THIS COULD BE A SPECIFIC DATE (DD/MM/YYYY) OR AN INDICATION SUCH AS 'UPON PUBLICATION OF RESEARCH RESULTS'.</i></p>	<p>Data will be made available upon publication of research results.</p>
<p>Which data usage licenses are you going to provide? If none, please explain why.</p> <p><i>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.</i></p> <p><i>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." <sup>8</sup></i></p>	<p>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.</p>

---

<sup>8</sup> Source: Ghent University Generic DMP Evaluation Rubric: <https://osf.io/2z5g3/>

<p>Do you intend to add a PID/DOI/accession number to your dataset(s)? If already available, please provide it here.</p> <p><i>INDICATE WHETHER YOU INTEND TO ADD A PERSISTENT AND UNIQUE IDENTIFIER IN ORDER TO IDENTIFY AND RETRIEVE THE DATA.</i></p>	<p><input checked="" type="checkbox"/> Yes  <input type="checkbox"/> No  If yes:</p> <p>Accession numbers and identifiers will be assigned to the data sets upon publication and provided in the published manuscript for data transparency.</p>
<p>What are the expected costs for data sharing? How will these costs be covered?</p>	<p>Usage of the above-mentioned public servers for data sharing is free. Costs for open access publication will be covered by the FWO bench fee or within the ERC grant MORE2ADA2.</p>

## 7. Responsibilities

Who will manage data documentation and metadata during the research project?	Lisa Ehlers (PhD Student), Leen Moens (senior post doc), Prof. Isabelle Meyts (PI)
Who will manage data storage and backup during the research project?	Prof. Isabelle Meyts (PI) and Leen Moens (senior post doc)
Who will manage data preservation and sharing?	Prof. Isabelle Meyts (PI) and Leen Moens (senior post doc)
Who will update and implement this DMP?	Lisa Ehlers (PhD Student) and Prof. Isabelle Meyts (PI)