

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	Margaux Evenepoel
FWO Project Number & Title	11N1222N Gut-brain interactions in Autism: A multi-modal investigation of the brain, bacteria and behavior
Affiliation	<input checked="" type="checkbox"/> KU Leuven <input type="checkbox"/> Universiteit Antwerpen <input type="checkbox"/> Universiteit Gent <input type="checkbox"/> Universiteit Hasselt <input type="checkbox"/> Vrije Universiteit Brussel <input type="checkbox"/> Other:
2. Data description	
Will you generate/collect new data and/or make use of existing data?	<input type="checkbox"/> Generate new data <input checked="" type="checkbox"/> Reuse existing data

<p>Describe the origin, type and format of the data (per dataset) and its (estimated) volume</p> <p><i>If you reuse existing data, specify the source of these data.</i></p> <p><i>Distinguish data types (the kind of content) from data formats (the technical format).</i></p>	<ul style="list-style-type: none"> - 80 children with an official Autism Spectrum Disorder (ASD) diagnosis - 40 children from the general population, without any neuropsychiatric diagnosis <p>(1) <u>Personal data</u>: captured via electronic patient record in online Klinisch WerkStation (KWS) environment Type of data: Surveys of all 140 participants and their primary caregiver Source: Electronic surveys via electronic patient record (myNexuzhealth) in KWS platform UZ Leuven File format: KWS Excel export (.xlsx) Volume: 10 MB</p> <p>(2) <u>Magnetic resonance imaging (MRI) of the brain</u>: performed via a Philips 3 Tesla Intera scanner with standard head coil stationed at NMR-unit of the Department of Radiology, University Hospital Leuven Type of data: Anatomical MR image Source: T1-weighted spin-echo imaging to assess brain structure (MPRAGE) File format: NiFTI file (.nii) Volume: 1 scan: 20 MB/scan Type of data: Resting-state functional MR image Source: BOLD contrast imaging to assess brain activity at rest (WP 1) File format: NiFTI file (.nii) Volume: 1 run: 230 MB/run</p> <p>(3) <u>Biological (stool) samples for characterization of gut microbiome</u> Type of data: Microbiome compositions Source: Fecal DNA extraction, DNA amplicon sequencing and taxonomical classification</p>
---	--

3. Ethical and legal issues

<p>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.</p> <p><i>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.</i></p>	<p><input checked="" type="checkbox"/> Yes <input type="checkbox"/> No</p> <p>If yes:</p> <ul style="list-style-type: none"> - Privacy Registry Reference: G-2021-3480, approved on April 16, 2021. - All personal data will be collected via electronic surveys and accessed via the secured electronic patient record (mynexushealth) within the KWS platform at UZ Leuven: <ol style="list-style-type: none"> 1) Ordinary personal data: name of the children and their parents, parents' contact details > only acquired for communicating with the participants throughout the course of the study, will not be extracted from KWS for long-term storage. 2) Sensitive personal data: age, gender, ASD symptom severity, standardized questionnaires input, IQ, data related to parents' socio-economic status (SES; educational level, profession category, wage) - Participant's contact details (name, address, e-mail address) and study ID codes are stored in a secured database with restricted access (PI, study clinician, and study coordinator). Samples, data, and results of analyses are de-identified, and stored and processed in coded form. Coded data and results of sample analyses are stored in secondary independent secured databases, accessible by researchers of collaborating labs only. These databases will not contain data that would allow participant identification without decoding.
<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, add the reference to the formal approval by the relevant ethical review committee(s).</p>	<p><input checked="" type="checkbox"/> Yes <input type="checkbox"/> No</p> <p>If yes:</p> <ul style="list-style-type: none"> - Approval for the research protocol and Informed Consent Forms was granted by the UZ/KU Leuven Ethical Committee (approval number: S61358) on September 27, 2018. PRET application number: G-2021-3480, approved on April 16, 2021.

<p>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?</p>	<p><input type="checkbox"/> Yes <input checked="" type="checkbox"/> No If yes, please comment:</p>
<p>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?</p>	<p><input type="checkbox"/> Yes <input checked="" type="checkbox"/> No If yes, please comment:</p>

4. Documentation and metadata

What documentation will be provided to enable understanding and reuse of the data collected/generated in this project?

(1) Survey data

- Metadata (e.g. timestamp, electronic instructions) are automatically captured during survey completion in KWS.
- A standard operating procedure will be written to describe how to export participant-related information from KWS to the e-CRF in RedCap.
- Using RedCap, a Data Dictionary Codebook will be generated containing variable-level information for all captured information: Variable / Field name, Field Label (including question text) and Field Attributes (including Field Type, Validation, Choices, Calculations etc.)

(2) Magnetic Resonance Imaging (MRI) data:

- A standardized electronic worksheet (Word file) will be completed during each data collection session containing all relevant meta-data, researchers notes, remarks concerning data quality, contextual information, deviations from the protocol etc. These worksheets will be kept in the same folder as the actual research data.
- Names of the specific files will contain information related to their scanning modality and time when generated (e.g. _STRUCT, _REST, _TASK).
- Raw MRI volumes will be preprocessed via fMRIPrep: a standardized fMRI data preprocessing pipeline implemented within an Open Source container structure, with extensive online documentation: <https://fmripred.org/en/stable/index.html>
- A Jupyter Notebook will be compiled to describe the performed analyses.

(3) Biological (stool) samples for characterization of gut microbiome:

- Meta-data regarding time stamps and instruction for sample collection will be available.
- Questionnaire data relating to ratings of stool sample composition (Bristol stool sample scale) and life style assessment (medication use, dietary profile, incidence of gastrointestinal symptoms) will be linked to the biological samples.

<p>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.</p>	<p><input checked="" type="checkbox"/> Yes <input type="checkbox"/> No If yes, please specify:</p> <p>(2) Magnetic Resonance Imaging (MRI) metadata standards: NIFTI and BIDS</p> <ul style="list-style-type: none"> - MRI volumes collected during the various scans will be stored according to the Neuroimaging Informatics Technology Initiative (NIFTI) file format. This format consists of a header file to store all meta-information regarding acquisition parameters and procedures, and an image file including the actual neuroimaging data. - In the NIFTI image file, the first three dimensions are reserved to describe the acquired volumes/images across three spatial dimensions (x, y and z), while the fourth dimension is reserved to define the time points (t). - All images collected, as well as their derivatives created during preprocessing, will be stored and organized according to the standardized BIDS format: https://bids.neuroimaging.io/
--	---

5. Data storage & backup during the FWO project

Where will the data be stored?	<p>(1) Personal data</p> <ul style="list-style-type: none"> - Source data: captured and stored within the electronic patient record of the secure UZ Leuven KWS environment for private data. <p>We will store personal data using individual codes in an electronic format that does not allow connecting the data to individuals and document all data collected throughout the course of the research in a case report form (CRF) via a secure web-based application (Research Electronic Data Capture, REDCap).</p> <p>(2) Neural MRI data:</p> <ul style="list-style-type: none"> - Source data will be exported immediately after collection from their respective research instruments and stored on the project's central storage facility, a shared folder on the password-protected L:drive for Large Volume Storage within the KU Leuven Internet File Access environment. For active use, copies from the master data on the L:drive can be made and kept on the personal devices of the involved researchers. <p>(3) Biological (stool) samples for characterization of gut microbiome:</p> <ul style="list-style-type: none"> - Stool samples will be stored in secure freezers at the UZ KU Leuven Biobank.
How will the data be backed up?	The data will be stored on RedCap and a central server (L:drive) with automatic daily back-up procedures.
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.	<p><input checked="" type="checkbox"/> Yes</p> <p><input type="checkbox"/> No</p> <p>If no, please specify:</p> <p>RedCap is hosted on central ICTS webservices and provides unlimited capacity. The minimum for large volume storage provided by the KU Leuven ICTS-hosted L:drive is 5 TB. It is expected this volume is sufficient for the current project. A disaster recovery (mirror) copy of the data is included in this fee.</p>

<p>What are the expected costs for data storage and backup during the project? How will these costs be covered?</p> <p><i>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.</i></p>	<p>The price to set-up a RedCap projects is € 80 per year. Data storage on L:drive storage will result in a cost of € 569,2 per year (for max. 5 TB of data). Costs for data storage will be covered by personal funds of the involved PIs (Kaat Alaerts and Bart Boets).</p>
<p>Data security: how will you ensure that the data are securely stored and not accessed or modified by unauthorized persons?</p>	<ul style="list-style-type: none"> - All included storage facilities (KWS, RedCap, L:drive) are incorporated within secured KU / UZ Leuven environments, are password-protected (including smartphone-based multi-factor identification) and are only accessible by registered collaborating researchers. - Access of study personnel to the UZ Leuven KWS environment will be revoked after study completion. - All data files will be collected, processed and stored in a de-identified format by means of subject ID codes (i.e. pseudonymization). These datafiles will not contain information that would allow participant identification. - Personal data collected on paper (e.g. informed consent forms) are stored in a locked cabinet onsite (during data collection: accessible only to study personnel; after data collection: accessible solely by PI of the study). - Stool samples will be stored in secure freezers at the UZ KU Leuven Biobank.

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 year 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, ...).	All generated research data will be archived for minimal 10 years after study completion conform the FWO and KU Leuven RDM policy. Biological samples will be stored for a maximum of 30 years, or until collected biological material (e.g. saliva) is fully depleted by use in the analyses.
Where will these data be archived (= stored for the long term)?	The generated research data, the accompanying metadata and all documentation necessary to reuse the data will be transferred to the K:drive for long-term data archiving (managed by KU Leuven ICTS with automatic back-up procedures). Stool samples will be stored in secure freezers at the UZ KU Leuven Biobank.
What are the expected costs for data preservation during these 5 years? How will the costs be covered? <i>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.</i>	Pricing for data storage on the K:drive includes € 11,38 per 100 GB (with 50% of the cost covered by Group Biomedical Sciences). In view of the expected size of the database (including raw and preprocessed data), estimated cost of long-term data storage will be € 56,9 per year for 500 MB.

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 3 rd party, legal restrictions)?	<input type="checkbox"/> Yes <input checked="" type="checkbox"/> No If yes, please specify:
---	---

Which data will be made available after the end of the project?	Pending new data releases, we plan to make the de-identified, de-faced and anonymized raw MRI data collected during resting state (WP 1) available via the Autism Brain Imaging Data Exchange (ABIDE; http://fcon_1000.projects.nitrc.org/indi/abide) online repository.
Where/how will the data be made available for reuse?	<input type="checkbox"/> In an Open Access repository <input checked="" type="checkbox"/> In a restricted access repository <input type="checkbox"/> Upon request by mail <input type="checkbox"/> Other (specify): Pending new data releases, we plan to make the de-identified, de-faced and anonymized raw MRI data collected during resting state (WP 1) available via the Autism Brain Imaging Data Exchange (ABIDE; http://fcon_1000.projects.nitrc.org/indi/abide) online repository
When will the data be made available?	Upon publication of the research results
Who will be able to access the data and under what conditions?	Access to the dataset will be monitored via the NeuroImaging Tools & Resources Collaboratory (NITRC) Image Repository and will be considered after a request is submitted explaining the planned reuse. Only uses for research purposes will be allowed and commercial reuse will be excluded.
What are the expected costs for data sharing? How will these costs be covered? <i>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.</i>	There will be no expected costs for data sharing.

8. Responsibilities

Who will be responsible for the data documentation & metadata?	Applicant Margaux Evenepoel (in cooperation with other collaborating researchers)
Who will be responsible for data storage & back up during the project?	Back-up and immediate storage: all research personnel Long-term storage: PIs Kaat Alaerts, Bart Boets
Who will be responsible for ensuring data preservation and sharing?	PIs Kaat Alaerts, Bart Boets

<p>Who bears the end responsibility for updating & implementing this DMP?</p> <p><i>Default response: The PI bears the overall responsibility for updating & implementing this DMP</i></p>	<p>Applicant Margaux Evenepoel will, together with other team members, be responsible for day-to-day data management as well as the implementation of this DMP.</p> <p>PI Kaat Alaerts (together with co-PIs Bart Boets, Marie Joossens and Jellina Prinsen) will bear the end responsibility for overall data management, storage and preservation.</p>
--	--