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	**Professor Inez Myin-Germeys, 0000-0002-3731-4930**
Contributor name(s) (+ ORCID) & roles	**Dr Olivia Kirtley, 0000-0001-5879-4120, Co-Investigator** **Dr Ginette Lafit, 0000-0002-8227-128X, Co-Investigator**
Project number [1] & title	FWO G049023N; The who, how, and what of experience sampling - Building an evidence-based methodological foundation for experience-sampling research in different populations
Funder(s) GrantID [2]	FWO G049023N
Affiliation(s)	**X** KU Leuven ROR identifier KU Leuven: 05f950310

[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.

Please provide a short project description	More and more mental health research is moving out of the lab and into people's daily lives. One popular way of investigating people in their daily lives is the experience sampling method (ESM). In an ESM study, researchers give individuals short questionnaires about their daily experiences on a smartphone, multiple times per day and over several days. Even though many mental health researchers are using ESM, basic questions surrounding the design of ESM studies remain unanswered. For example, it is unclear how an ESM study should be set up and what questions can best be used. We also do not understand how these choices depend on the type of people that are being studied. This lack of methodological research makes it difficult to evaluate and compare results from ESM studies. In the proposed project, we will investigate how the number of ESM questionnaires per day influences the information collected across different groups of individuals, namely adolescents and vulnerable young adults. Next, we will find out how participants can best be paid to collect as much daily-life data as possible. Then we will test how well existing ESM questions allow us to measure daily life in different groups of people. Based on these studies, we will develop guidelines for researchers who want to design better ESM studies in the future. In the long run, our research will help to improve the quality of mental health research. The current project comprises five studies: two 'add on' studies, involving extending data collection for two ongoing ESM studies (Study 1: SCOUT-Clinical; and Study 2: SIGMA-X), a new qualitative study (Study 3), a new quantitative ESM study (Study 4), and a study reusing preexisting data (Study 5).

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 [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
Study 1 (SCOUT-Clinical)	'Add-on' data collection for ongoing study of suicidal thoughts and behaviours in young adults, including ESM, self-report questionnaire, and follow-up interview data	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images ☒ Sound ☒ Numerical ☒ Textual ☐ Model ☐ Software ☐ Other:	.csv; audio file format TBC	☐ < 1 GB ☒ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA Based on maximum possible observations, given protocol and sample size. ESM = 1800 observations; self-report questionnaires = 60 observations; interviews = 60 observations	NA
Study 2 (SIGMA-X)	'Add-on' data collection for	☒ Generate new data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images	.csv; audio file format TBC	☐ < 1 GB ☒ < 100 GB	NA

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

	ongoing study of mental health and well-being in youth from socioeconomically disadvantaged backgrounds including ESM, self-report questionnaire, and follow-up interview data	☐ Reuse existing data		☒ Sound ☒ Numerical ☒ Textual ☐ Model ☐ Software ☐ Other:		☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA Based on maximum possible observations, given protocol and sample size. ESM = 5400 observations; self-report questionnaires = 120 observations; interviews = 20 observations	
Study 3 (Qualitative study)	Response Process Evaluation study to gather responses to iterative self-report questionnaires about ESM items.	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images ☐ Sound ☒ Numerical ☒ Textual ☐ Model ☐ Software ☐ Other:	.csv	☐ < 1 GB ☒ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA Based on the maximum possible observations, given protocol and sample size. N= 64 responses to the iterative self-report	NA

						questionnaires.	
Study 4 (Quantitative ESM study)	New ESM study with student population to investigate effects of study design choices on ESM data quality and quantity, including ESM, self-report questionnaire, and follow-up interview data	☒ Generate new data ☐ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images ☒ Sound ☒ Numerical ☒ Textual ☐ Model ☐ Software ☐ Other:	.csv; audio file format TBC	☐ < 1 GB ☒ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA Based on maximum possible observations, given protocol and sample size. ESM = 22680 observations; self-report questionnaires = 180 observations; interviews = 180 observations	NA
Study 5 (Harvesting pre-exiting ESM data)	Pre-existing ESM data will be gathered from various sources to enable psychometric evaluation of ESM items	☐ Generate new data ☒ Reuse existing data	☒ Digital ☐ Physical	☐ Audiovisual ☐ Images ☐ Sound ☒ Numerical ☒ Textual ☐ Model ☐ Software ☐ Other:	. csv	☐ < 1 GB ☒ < 100 GB ☐ < 1 TB ☐ < 5 TB ☐ > 5 TB ☐ NA Unknown - this depends on the number of relevant	NA

						datasets we identify and the numbers of participants in these original studies. This will be determined later in the project.	

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.	Various sources of preexisting data will be used in Study 5, including from our own CCP database of studies, and potentially data from contributors to the ESM Item Repository (www.esmitemrepository.com) and EMOTE data-sharing platform (https://emotedatabase.com/).

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: ☐ Yes, dual use; provide approval number: ☐ No Additional information: Study 1: The SCOUT-Clinical study has already received ethical approval from the KU Leuven/UZ Leuven Medical Ethics Committee (METC; S66261). We are currently preparing an amendment to enable us to collect data for the add-on study for this project. Study 2: The SIGMA-X study has received ethical approval from the KU Leuven/UZ Leuven Medical Ethics Committee (METC; S61395). We are currently preparing an amendment to enable us to collect data for the add-on study for this project. Study 3: I am currently preparing an ethics application for this qualitative study, which will be submitted to SMEC at KU Leuven. Study 4: I am currently preparing an ethics application for this quantitative study, which will be submitted to SMEC at KU Leuven. Study 5: When suitable datasets are identified for reuse, we will prepare ethical applications according to the reuse requirements for these specific datasets, either via SMEC, KU Leuven/UZ Medical Ethics Committee, or the ethical committee of the data-generator's institution.

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: Personal data (name, email address, telephone number, (postal) address) will be collected from the 'add-on' samples for SCOUT-Clinical and SIGMA-X (Studies 1 and 2 in the current project), because these are longitudinal studies and these data are required enable participant follow up. Personal data (name, email address, telephone number, (postal) address) will also be collected in Studies 3 and 4. In the context of the current project, personal data will only be used to enable us to reimburse participants for taking part in the study.
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:
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: For Study 5, we will use preexisting ESM data. For some of these data we will not be the owner, therefore we will follow the relevant data transfer agreements, research collaboration agreements, where necessary, which will likely include not disseminating data for this study.
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:

[4] See Glossary Flemish Standard Data Management Plan

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*.	**Study 1:** Codebooks of the self-report questionnaires and ESM data from the SCOUT-Clinical study will be created using REDCap. Protocols for the study will be documented in a Word document. A users' guide for the dataset will be made available after data collection has concluded. **Study 2:** A codebook of the self-report questionnaire and ESM data from SIGMA-X will be created using REDCap. A users' guide to the dataset will be produced. Details about the questionnaires used in SIGMA and any relevant info about adaptations will be provided in the protocol. **Study 3:** A codebook of the quantitative and qualitative variables for this study will be produced using REDCap. Protocols for the study will be documented in a Word document. A users' guide for both datasets will be made available after data collection has concluded. **Study 4:** Codebooks of the self-report questionnaires and ESM data from this study will be created using REDCap. Protocols for the study will be documented in a Word document. A users' guide for the datasets will be made available after data collection has concluded. **Study 5:** Where codebooks for preexisting data do not exist, we will produce them. The protocol for this study will be documented in a Word document. As, in some cases, we will not be the owners of the data, we will create a users' guide for the datasets only for internal use during this study.
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: All data from all studies are stored in REDCap, which uses standardized metadata. If no, please specify (where appropriate per dataset or data type) which metadata will be created:

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: All data are stored in REDCap, hosted by secure KU Leuven servers.
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) All data are stored in REDCap, hosted by secure KU Leuven servers, with automatic daily backup. Manual backup of data will occur every week and these backups are stored on access-controlled, secure drives, hosted on secure KU Leuven servers.

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 Study 1: SCOUT-Clinical data will be stored and backed-up in REDCap. A REDCap project (and storage space) has been purchased for these studies. Study 2: SIGMA-X data will be stored and backed up in REDCap. A REDCap project (and storage space) has been purchased for these studies. Study 3: Data will be stored and backed-up in REDCap. A REDCap project (and storage space) will be purchased for these studies. Study 4: Data will be stored and backed-up in REDCap. A REDCap project (and storage space) will be purchased for these studies. Study 5: Data will be stored and backed-up in REDCap. A REDCap project (and storage space) will be purchased for these studies. If no, please specify:
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 are stored in REDCap, hosted on secure KU Leuven servers. Access to REDCap is limited by the Data Manager and is only permissible for primary researchers on the project. Identifying information is stored separately within REDCap and cannot be linked to participants' data.

What are the expected costs for data storage and backup during the research project? How will these costs be covered?	Two REDCap projects per study will be required for Studies 1 and 2, and one each for Studies 3, 4, and 5. Each project costs 80 Euros per year to store. Costs will therefore be 14 x (80 x 7) = 7,840 (4 years for duration of project + 10 years post-project storage). Study 1: The SCOUT-Clinical study is covered by an FWO Senior Postdoctoral Fellowship to Dr Olivia Kirtley (CI on the current project and incoming Assistant Professor at the CCP). Following the end of the fellowship, these costs will be covered by the CCP. Study 2: SIGMA-X data storage costs are covered by an FWO Odysseus grant to Professor Inez-Myin Germeys, Director of the Center for Contextual Psychiatry at KU Leuven. Study 3: Data storage and back-up costs will be covered by the FWO grant for this project. Following the end of the project, these costs will be covered by the CCP. Study 4: Data storage and back-up costs will be covered by the FWO grant for this project. Following the end of the project, these costs will be covered by the CCP. Study 5: Data storage and back-up costs will be covered by the FWO grant for this project. Following the end of the project, these costs will be covered by the CCP.

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): The Center for Contextual Psychiatry (where this research is taking place) operates a data checkout system to facilitate open science practices, including preregistration and Registered Repots - Data curation for Open Science (DROPS; Kirtley, Lafit, Wampers, & Myin-Germeys, 2020). Full datasets are available only to the CCP data manager, Dr Martien Wampers, and abstract submission and variable access requests operate via a series of linked questionnaires in REDCap. The DROPS system can be accessed here: https://redcap.gbiomed.kuleuven.be/surveys/?s=WDYAFAHWK4. Interested researchers will be provided with the link to the DROPS system in order to apply for access.
What are the expected costs for data preservation during the expected retention period? How will these costs be covered?	**See section:** What are the expected costs for data storage and backup during the research project? How will these costs be covered?

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 full datasets for studies 1 – 4 will be available in REDCap via the DROPS system — the Center for Contextual Psychiatry's in-house data checkout system (Kirtley, 2022; Kirtley et al., 2020). As Study 5 may include data for which we are not the data owners, these data will not be made available and we will follow the sharing requirements/restrictions of the data owners.

If access is restricted, please specify who will be able to access the data and under what conditions.	Researchers wishing to access the data from Studies 1 – 4 in DROPS are able to submit a request to access specific sets of variables relevant for their analyses, but a full dataset will not be released, i.e., including variables not relevant for specific analyses. This is to facilitate open science practices including preregistration and Registered Reports, which require restriction of data and knowledge of the data until after preregistration or Stage 1 manuscript acceptance (Registered Reports). The DROPS system is in place for all researchers, irrespective of their internal or external status. External applicants will also be required to sign a data agreement, to ensure that data are used only for research (and not for commercial purposes) and that KU Leuven ethical and privacy standards are adhered to. Kirtley, O. J. (2022). Advancing credibility in longitudinal research by implementing open science practices: Opportunities, practical examples, and challenges [Note]. *Infant and Child Development*, *31*(1), Article e2302. https://doi.org/ARTN e2302 10.1002/icd.2302 Kirtley, O. J., Lafit, G., Wampers, M., & Myin-Germeys, I. (2020, 07/12/20 - 08/12/20). *Establishing a low-threshold data checkout system using REDCap to facilitate preregistration and Registered Reports for pre-existing data* CSPD 2020: Sharing Psychological Research Data: Best Practices and New Developments, https://www.conference-service.com/CSPD2020/xpage.html?xpage=226&lang=en

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: In cases where researchers wish to access data in order to conduct secondary data analysis, access will not be granted where the proposed work substantially overlaps with in progress work of the immediate research team. Data are only available for research use and commercial use is strictly prohibited. For Study 5, which involves harvesting preexisting data for reuse, we will follow the sharing requirements/restrictions of the data owners.
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) The Center for Contextual Psychiatry (where this research is taking place) operates a data checkout system to facilitate open science practices, including preregistration and Registered Repots - Data curation for Open Science (DROPS; Kirtley, Lafit, Wampers, & Myin-Germeys, 2020). Full datasets are available only to the CCP data manager, Dr Martien Wampers, and abstract submission and variable access requests operate via a series of linked questionnaires in REDCap. The DROPS system can be accessed here: https://redcap.gbiomed.kuleuven.be/surveys/?s=WDYAFAHWK4. Interested researchers will be provided with the link to the DROPS system in order to apply for access.

When will the data be made available?	☐ Upon publication of research results ☐ Specific date (specify) ☒ Other (specify) When research results are published, data relevant to those analyses will be requestable via DROPS for purposes of verification and reanalysis. The full range of variables within the datasets will be requestable via DROPS within five years of the project ending, in order to allow the immediate research team sufficient time to publish from these data.
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?	The DROPS system operates via REDCap, therefore there is a yearly cost of 80 Euros per REDCap project associated with maintaining this system. This is covered by the FWO grant supporting the current project, an FWO Senior Postdoctoral Fellowship to Dr Olivia Kirtley, and an FWO Odysseus grant to Professor Inez Myin-Germeys, Director of the Center for Contextual Psychiatry at KU Leuven — the center in which this research is taking place.

7. Responsibilities	
Who will manage data documentation and metadata during the research project?	The PI, Prof. dr Inez Myin-Germeys and CIs, Dr Olivia Kirtley and Dr Ginette Lafit, will be responsible for this along with Dr Martien Wampers, Data Manager at the Center for Contextual Psychiatry (CCP), where this research is taking place.
Who will manage data storage and backup during the research project?	Dr Martien Wampers, Data Manager at the Center for Contextual Psychiatry (CCP), where this research is taking place.
Who will manage data preservation and sharing?	The PI, Prof. dr Inez Myin-Germeys and CIs, Dr Olivia Kirtley and Dr Ginette will be responsible for this along with Dr Martien Wampers, Data Manager at the Center for Contextual Psychiatry (CCP), where this research is taking place.
Who will update and implement this DMP?	The PI, Prof. dr Inez Myin-Germeys and CIs, Dr Olivia Kirtley and Dr Ginette bear the end responsibility of updating & implementing this DMP.