PhD: Protective factors for suicidal ideation and behaviour in young people
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
SIGMA (S61395) ESM-study	Wave 1 (N = 1913): longitudinal study of adolescents (aged 11-20y). 22 schools from across Flanders participated.	**Reuse existing data**	**Digital**	**Observational** Quantitative self-report questionnaires collected and stored using RedCap. The questionnaires cover topics such as demographics, social functioning, attachment, well-being, (mental) health problems etc. Quantitative self-report ESM data: collected with an app and uploaded in redcap. The ESM questionnaire contains items about affective state, context, mental health symptoms	**.csv**	**<5TB**	NA

HOPE focus groups (S66032)	Interviews + focus groups with 4 panels each consisting of N=8 adolescents (aged 16-25y)	**Reuse existing data + generate new data** Data collection ongoing: 2 of the 4 panels have been completed, with two more panels to be carried out. Data has not been used for analysis before.	**Digital** **Physical (offline questionnaires)**	**Observational** Sound; Textual Qualitative data: Audio recordings of the focus groups and interviews. Transcripts of these audio recordings (spoken language transcribed to written language)	.mp3; .txt	**<100GB**	+/- 100 pages
BOUNCE BACK - online survey	Large online international open-text survey (N=2000) with adults.	**generate new data**	**Digital**	**Observational** Quantitative self-report questionnaires collected and stored using RedCap + Qualitative responses (written language) on open-text questions	.csv .txt	**<100GB**	NA
ESM-study on protective factors	ESM-study (10 questionnaires per day for 14 days at semi-random time points) with N=50, consisting of 25 adolescents (aged 12-18) and 25 young adults (aged 18- 25).	**generate new data**	**Digital**	**Observational** Quantitative self-report questionnaires collected and stored using RedCap. The questionnaires cover topics such as demographics, social functioning, attachment, well-being, (mental) health problems etc. Quantitative self-report ESM data: collected with an app and uploaded in redcap. The ESM questionnaire contains items about affective state, context, mental health symptoms	.csv	**<100GB**	NA

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:

I will reuse the data from the SIGMA study (wave 1): A codebook of the self-report questionnaire and ESM data has been created (https://sigmaleuven.shinyapps.io/Interactive-Codebook/). The protocol and sample description paper for SIGMA Wave 1 is available as a preprint: Kirtley et al., 2021; https://psyarxiv.com/jp2fk. Details about the questionnaires used in SIGMA and any relevant info about adaptations is provided in the protocol and sample description paper (and accompanying supplementary materials). The HOPE data collection is ongoing: two of the four panels have already taken place and were conducted by other members of the research team. I will be conducting the other 2 panels. The data is currently stored with our data manager on the J-drive. Apart from being transcribed, no other work has been done with the data. No DOIs exist.

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

The studies in this project involving new data collection will be submitted to the appropriate ethical committee (EC UZ Leuven) for ethical approval prior to starting data collection.

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.

- Yes

All studies included in this project will contain general (e.g. demographic information such as age, gender, ethnicity) and mental health related personal data. The S-number of the SIGMA and the HOPE study can be found in the table above. S-numbers of the Bounce Back study and the ESM study on protective factors will be provided once obtained. The data will always be pseudonymised. Information about participants' name, e-mail address etc. (if available) will never be part of datasets exported for data analysis.

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.

- No

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.

- No

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.

- No

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).

For all studies in this project that use Redcap* for data collection and storage, the Redcap codebook will be available to provide information on all variables included in the studies, including variable names, corresponding questionnaire items, response format and type (e.g. open text, multiple choice, etc.). In addition, readme files will be created to provide additional information on study details including study design, study population, ESM protocol and protocol violations. If study-specific scoring or cleaning rules apply, these will also be explained in the readme file.

For the HOPE study, where data will be stored on the secured KU Leuven J-drive, a readme file will be created to provide additional information on study details.

I will create a pre-registration for each study (post-registration for those studies where data already exist) documenting the study design, sampling, measures, variables and data analysis plan on the Open Science Framework (OSF).

* REDCap is a secure web application for building and managing online surveys and databases hosted on KU Leuven servers (Harris et al, 2009)

PA Harris, R Taylor, R Thielke, J Payne, N Gonzalez, JG. Conde, Research electronic data capture (REDCap) – A metadata-driven methodology and workflow process for providing translational research informatics support, *J Biomed Inform. 2009 Apr;42(2):377-81.* PA Harris, R Taylor, BL Minor, V Elliott, M Fernandez, L O'Neal, L McLeod, G Delacqua, F Delacqua, J Kirby, SN Duda, REDCap Consortium, The REDCap consortium: Building an international community of software partners, *J Biomed Inform. 2019 May 9 [doi: 10.1016/j.jbi.2019.103208]*

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.

- Yes

All data that are stored in REDCap uses standardized metadata.

Other data such as the interview audio files and transcripts will be securely archived and not covered by a specific metadata file other than the README file with descriptive summaries of the data as described above.

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

Where will the data be stored?

The data from self-report questionnaires, the ESM questionnaires and the open-text questions of the studies mentioned in table 1 are stored and managed in REDCap (Research Electronic Data Capture). REDCap is a secure web application designed for building and managing online surveys and databases, hosted on KUL servers. Access to the redcap projects containing the research data is limited to the data manager and the study PI's.

ESM data will be collected using a smartphone app and subsequently uploaded into REDCap to link self-report and ESM data within the same study participant.

The qualitative data collected with the interviews and focus groups are stored using central storage facilities of the faculty of biomedical sciences (J-drive) with automated backup. This includes both the audio data and the transcribed data. Access to the j-drive folders is restricted.

Offline copies (questionnaires) and informed consents will be separately archived in a locked room for the expected 10 year period after the end of the project.

How will the data be backed up?

Standard back-up provided by KU Leuven ICTS for my storage solution

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

SIGMA data are stored and backed up in REDCap for which we pay for adequete storage space via KU Leuven.
Future data will also be stored and backed-up in REDCap. A REDCap project (and storage space) will be purchased for these studies.

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

Access to data stored in both REDCap and the J-drive is restricted to the data manager and the PIs.
Pseudonymized data of the SIGMA study can be requested for specific research questions by submitting an abstract through an automated abstract submission system (DROPS), developed at Center for Contextual Psychiatry.
https://sigmaleuven.shinyapps.io/DROPS_User_Guide/#Which_data_can_be_requested_through_DROPS

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

For the existing datasets collected and managed via REDCap, there is an annual fee of 80 euro per project.
The qualitative data from the HOPE-study will be stored on the j-drive, with average costs of approximately 105 euros per year per 100 GB. Given the limited number of included participants (N = 16), the required storage capacity and consequently the cost will be limited, utilizing a portion of a 100 GB drive that will be purchased.These expenses are funded through the grants for these studies.
For the new ESM-study on protective factors, data collection will occur via REDCap (self-report questionnaires) and via mPath (for collecting ESM data). For the use of REDCap, there will be an annual fee of 80 euros for the period during which the data need to be retained. For the use of mPath, 1300 euros (excluding VAT) will be paid. The anticipated size of this data is less than 100 GB, so the maximum cost will be 105 euros per year. These expenses will be covered by the project funding.

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 data will be preserved for 10 years according to KU Leuven RDM policy.

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

As long as research with the existing or newly collected data is conducted on a regular basis, storage in REDCap and on the j-drive will be retained. Afterwards, the information from REDCap, including all documentation required for the long-term usability of the data (codebook, study protocol, etc.), will be exported from REDCap and stored on the j-drive. Unless the volume of data storage on another drive is more cost-effective, in which case another central storage facility within KUL will be used.
Offline copies (questionnaires) and informed consents will be separately archived in a locked room for the expected 10 year period after the end of the project.

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

This will depend on the available storage options at that time. For the next 5 years, the costs will be as described above.

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 a restricted access repository (after approval, institutional access only, …)

Pseudonymized data of the SIGMA study can already be requested for specific research questions by submitting an abstract through an automated abstract submission system (DROPS), developed at Center for Contextual Psychiatry (Kirtley, 2022; Kirtley et al., 2020).
All datasets produced by this project may be made available to other researchers, although decisions on this are still being made. If data are made available, researchers can submit an abstract through an automated data check out system (DROPS). Upon approval of the abstract, the researchers will receive a dataset containing the data necessary to address their research questions.
Researchers wishing to access the data are able to submit 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 to data and knowledge of the data until after preregistration or Stage 1 manuscript acceptance (Registered Reports).
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

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

Anyone can submit an abstract via the automated abstract submission system (DROPS), developed at Center for Contextual Psychiatry) to request data. If the abstract is approved by the PI or authorized CI and a data transfer agreement is signed, the researcher can receive a CSV file containing the data necessary to test their hypotheses as well as necessary documentation such as codebook, readme files etc.
Data access will be granted subject to the relevant ethical, legal, and privacy regulations.

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.

- No

NA

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

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. Interested researchers will be provided with the link to the DROPS system in order to apply for access.
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

When will the data be made available?

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 dataset 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.

Not yet known

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

Not yet known

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 associated with maintaining this system. This is covered by an FWO Odysseus grant to Professor Inez Myin-Germeys, Director of the Center for Contextual Psychiatry at KU Leuven, in which this research is taking place.

6. Responsibilities

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

Eveline Van Raes (PhD candidate) will be responsible for this along with her supervisor prof. dr. Olivia Kirtley and 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?

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?

Eveline Van Raes (PhD candidate) will be responsible, with the help of Dr Martien Wampers, Data Manager at the Center for Contextual Psychiatry (CCP), where this research is taking place.