Data Management Plan

Improving behavioral treatment for children with disruptive behavior

*[PROJECT ACRONYM (if available)**Error! Unknown document property name.Error! Unknown document property name.Error! Unknown document property name.**]*

[OPEN 6H.65724 Data Management Plan Manual NEXT TO THIS TEMPLATE;

THE MANUAL CONTAINS EXPLANATORY INFORMATION AND EXAMPLES OF WHAT INFORMATION SHOULD BE INCLUDED IN EACH STEP OF THE DMP.

DO YOU HAVE ANY QUESTIONS OR NEED SUPPORT?

PLEASE CONTACT THE DIGITAL COMPETENCE CENTER AT RESEARCHSUPPORT@UMCG.NL

WE STRONGLY ADVISE THAT YOU SUBMIT YOUR DMP TO THE DIGITAL COMPETENCE CENTER FOR REVIEW (RESEARCHSUPPORT@UMCG.NL)]

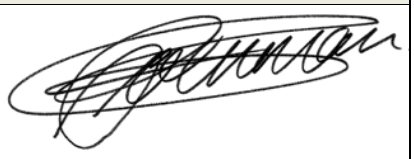
Date	**Version**	
03-01-2024	1	
Completed by		
Role in the study	Main applicant, co-PI	
Full name	Annabeth P. Groenman	
Department	Accare child and adolescent psychiatry	
Authorized by		
Role in the study	Project leader	Signature 
Full name	Tycho Dekkers	
Department	Accare child and adolescent psychiatry	Signature date

Table of contents

Data Management Plan 3
- Step 1: Details about your research 4
  - Research project summary 4
  - Data management roles and responsibilities 7
- Step 2: Describing your data 8
  - a. Tooling and infrastructure 8
  - b. Research data characteristics 8
  - c. In research involving personal data: data privacy and security measures 10
  - d. Resources for data management 10
- Step 3: Data storage and archiving 11
  - a. Storage and back up 11
  - b. Main security risks and measures 11
  - c. In research involving personal data: data security and protection measures 12
  - d. Preserving data 12
- Step 4: Data documentation and metadata 13
  - a. Metadata and documentation 13
  - b. Data quality control measures 14
- Step 5: Data availability, access and re-use 14
  - a. Intellectual property rights and data ownership 14
  - b. Data availability 15
  - c. Findability of research data: Unique and persistent identifiers 15

Data Management Plan

The University Medical Center Groningen (UMCG) was the first University Hospital in the Netherlands with a comprehensive ISO9001 certificate and is compliant covering all of its strategic activities: Healthcare, Research and Education. A hallmark of its ambition to achieve and maintain the highest quality possible in all of its endeavours. In this context, and to preserve scientific integrity, the Board of Directors of the UMCG has drawn up a Research Code. The UMCG Research Code provides researchers with guidelines for conducting medical research in a correct and ethical fashion. Everyone who conducts research at the UMCG and all UMCG employees who conduct research elsewhere in the world, are expected to be familiar with the Code and to act in accordance with its rules of conduct. The UMCG Toolbox has been developed to support and provide guidance to researchers in medical sciences research.

As far as ethics are concerned, the UMCG has a long-term policy (since 2007) on research integrity covering rules, regulations and best practices with respect to the use of patient data. Data protection and privacy is coordinated by the data protection officer (privacy@umcg.nl), which supervises that research with personal data conducted within the UMCG is performed in accordance with the General Data Protection Regulation (GDPR).

The UMCG and the RUG have adopted the FAIR data principles, and partnered on Research IT and the data management domain. In addition to the governance of the overall quality of research data, the re-use and intellectual property (IP) of data in particular, is governed by various agreements between the UMCG and RUG.

Moreover, the UMCG and the RUG have established Digital Competence Centers (DCCs) that provide advice about best practices with respect to research data management and guidance about applicable infrastructures and services that meet your project's needs. The UMCG DCC can help you in making your research data FAIR and compliant with applicable laws and regulations throughout the entire research data lifecycle (contact: researchsupport@umcg.nl).

Step 1: Details about your research

This step comprises mainly of administrative information and data management roles and responsibilities of your research. It is good to fill in the following tables completely and to include the names of the research project team member(s) involved. In collaborative projects, explain the co-ordination of data management responsibilities across partners.

Research project summary	
Project full title:	*Improving behavioral treatment for children with disruptive behavior*
Project acronym:	
UMCG Research Register number:	*18902*
Research protocol register number:	*Will be added later*
Subsidising party(ies) and grant number:	*ZonMw* # 07440212310027
Project data contact:	*Annabeth Groenman, https://orcid.org/0000-0002-8394-6605, Accare, center for child and adolescent psychiatry. Research Institute child development and education*
Principal investigator(s):	*Tycho Dekkers, https://orcid.org/0000-0001-8572-5606, Accare, center for child and adolescent psychiatry, Marjolein Luman, https://orcid.org/0000-0002-1539-2831, VU university Amsterdam. Annabeth Groenman, https://orcid.org/0000-0002-8394-6605, Accare, center for child and adolescent psychiatry. Research Institute child development and education*
Participating researcher(s) and/or organisation(s):	*J Joli Luijckx joli.luijckx@balansdifitaal.nl Balans* *dr. MK Marie Deserno m.k.deserno@uva.nl Universiteit van Amsterdam* *prof. dr. BJ Barbara van den Hoofdakker b.van.den.hoofdakker@accare.nl Accare* *prof. dr. PJ Pieter Hoekstra p.hoekstra@accare.nl Accare* *dr. PHO Patty Leijten p.h.o.leijten@uva.nl Universiteit van Amsterdam* *M Marijn Nijboer m.nijboer@accare.nl Accare* *dr. R. Rianne Hornstra r.hornstra@accare.nl Accare* *prof. dr. S Saskia van der Oord saskia.vanderoord@kuleuven.be KU Leuven*

Research data management support staff consulted and date[1]:	*UMCG DCC* *26 January 2024*	
Project duration:	Start: *1 April 2024*	End: *1 April 2028*

Short abstract of the research project (max. 500 words):

Achtergrond en relevantie

Ondersteuning van ouders via een gedragstherapeutische training (oudertraining) is de meest effectieve niet-medicamenteuze behandeling voor aandachtsdeficiëntie/hyperactiviteitsstoornis (ADHD) bij kinderen in de basisschoolleeftijd. In een oudertraining leren ouders technieken om gewenst gedrag bij hun kind te stimuleren, ongewenst gedrag te verminderen en de relatie met hun kind te verbeteren. Deze technieken vormen de kernelementen van oudertraining. Oudertraining werkt zeker niet voor alle gezinnen even goed en de effecten zijn klein tot middelgroot. Er is dus veel ruimte voorverbetering, zowel in het personaliseren van de behandeling als in het verhogen van de effectiviteit.

Probleemstelling en doelstellingen

In onze eerdere studies vonden we twee groepen veelbelovende kernelementen van oudertraining, namelijk antecedente technieken die gericht zijn op het structureren van de situatie voorafgaand aan gedrag, en consequente technieken die gericht zijn op het reageren op gedrag. Echter, inzicht in welke kernelementen we moeten inzetten bij welke gezinnen hebben we nog niet. Ook hebben we nog weinig begrip van de werkingsmechanismen van de succesvolle elementen: leiden deze vooral tot gedragsverandering bij ouders, of juist tot veranderingen in hun cognities en/of emoties? Kennis over deze kennishiaten kan ertoe leiden dat oudertraining beter op gezinnen afgestemd kan worden, verder kan worden geoptimaliseerd en effectiever is.

Dit project bestaat uit drie deelprojecten met de volgende doelstellingen:

Deelproject 1. Bepalen welke kernelementen meer of minder effectief zijn voor welke gezinnen.

Deelproject 2. Specificeren van de werkingsmechanismen van de meest veelbelovende kernelementen door inzicht te krijgen in de sequentie van veranderingen die plaatsvinden bij ouders en kinderen direct na het ontvangen van deze kernelementen.

Deelproject 3. In kaart brengen van ervaringen van kinderen van wie de ouders deel hebben genomen aan een kortdurende oudertraining.

Plan van aanpak/onderzoeksdesign

In deelproject 1 zullen we twee reeds door ons opgebouwde databases koppelen: een individuele participant data meta-analyse (IPDMA) database, met data van 23 studies (N=1,972 kinderen) die de effectiviteit van oudertrainingsprogramma's voor kinderen met ADHD onderzochten en een database met gecodeerde kernelementen, met daarin gedetailleerde informatie over de dosering van een drietal groepen succesvolle kernelementen (antecedente, en positieve en negatieve consequente technieken) van de oudertrainingsprogramma's in de IPDMA database.

[1] *For projects funded by NWO or ZonMw, you are expected to consult with data management support staff. Plans that have not been consulted with institutional data management support staff will not be considered by NWO or ZonMw. Contact the UMCG Digital Competence Center (researchsupport@umcg.nl).*

In deelproject 2 maken we gebruik van een Multiple Single Case Experimental Design (MSced) met intensieve tijdreeksanalyses. We zullen 12 gezinnen intensief volgen voor en na het ontvangen van succesvolle kernelementen. We brengen daarbij met behulp van een online applicatie veranderingen in gedrag, emoties en cognities van ouders en veranderingen in gedrag van het kind in kaart en analyseren hoe deze met elkaar samenhangen. In deelproject 3 zullen we kinderen interviewen aan de hand van semigestructureerde interviews van ongeveer 30 minuten.

Objecten/eenheden van het project

In deelproject 1 maken we gebruik van bestaande data. In deelproject 2 includeren we twaalf gezinnen van kinderen met ADHD van 6-10 jaar.

In deelproject 1 analyseren we kernelementen en gezinskenmerken van de geïncludeerde oudertrainingsprogramma's. In deelproject 2 krijgen ouders een korte gerichte training die bestaat uit enkele succesvolle kernelementen.

In deelproject 3 includeren we ca. 10 kinderen in de leeftijd van 8 t/m 11 jaar. Middels interviews brengen we de ervaringen van kinderen in kaart.

Primary or Overarching aim:	*Oudertraining voor ADHD verder personaliseren en optimaliseren.*
Secondary or Sub-studies aim(s):	*(1) welke veelbelovende kernelementen werken het best voor welke gezinnen?* *(2) wat zijn de werkingsmechanismen van veelbelovende kernelementen?* *(3) wat zijn de ervaringen van kinderen nadat hun ouders deelnemen aan een kortdurende oudertraining?*

Data management roles and responsibilities	
Data management plan:	*Annabeth Groenman, main applicant, co pi*
Data collection:	*In wp1 worden geen nieuwe data verzameld* *Marijn Nijboer onder begeleiding van Tycho Dekkers, Marjolein Luman en Annabeth Groenman is verantwoordelijk voor dataverzameling in deelproject 2 en 3*
Data quality control:	*Marijn Nijboer, de nog aan te stellen postdoc onderzoeker, Tycho Dekkers, Marjolein Luman en Annabeth Groenman*
Data processing:	*Marijn Nijboer, de nog aan te stellen postdoc onderzoeker, Tycho Dekkers, Marjolein Luman en Annabeth Groenman*
Data analysis:	*Marijn Nijboer, de nog aan te stellen postdoc onderzoeker, Tycho Dekkers, Marjolein Luman en Annabeth Groenman*
Data archiving:	*Marijn Nijboer, de nog aan te stellen postdoc onderzoeker, Tycho Dekkers, Marjolein Luman en Annabeth Groenman*
Data access:	*Annabeth Groenman, Tycho Dekkers, Barbara van den Hoofdakker*

Step 2: Describing your data

This step is targeted at getting a clear overview, inventorying the types of data involved in you research project, which tools/infrastructures will be used for collecting these data, whether privacy sensitive data will be collected and how will these be protected, etc. The next steps (3-5) of the DMP will refer to the data described in here, in Step 2.

Please answer all sub-questions that are applicable to your research.

Questions:

a. ***Tooling and infrastructure*** *to collect and process new data and/or existing data re-used.*

- What data are you using for this research? Will you re-use any existing data?
- How will data be created/collected for the research project? With which software, tool and methodology?
- What data is considered 'raw' and what data 'processed' in your research project?
- How will you clean/pre-process and process the 'raw' data? With which software, tool and methodology?

WP1 will make use of existing IPDMA database from the same research group.

WP 2 will collect data about child disruptive behavior (using an inhouse list of problem behaviors), parental adaptive and maladaptive parenting behavior, parental competence, and expressed emotions. Data will be collected twice daily for 42 days using m-path. This data is considered raw data. SPSS in combination with R studio will be used to clean and process this raw data. In WP3, the audiotaped interviews are considered raw data. Transcribed data will be analysed by thematic analysis using Atlas.ti.

b. ***Research data characteristics****: data types, formats and volumes that will be collected and produced.*

Data	Source	Data collection system	Kind of data collected and file format	Data preprocessing system and file format	Data processing system and file format	Data volume in total
WP2: Demographic data	Parent reported questionnaire data	Qualtrics	Numeric (.csv)	Spss/R	Spss/R	Expected data volume is small (< 1 gb for total dataset)
WP2: Parent reported data on child behavior and parenting, expressed emotions and parenting cognitions	Parent reported questionnaire data	At baseline Qualtrics, further measurement via m-path	Numeric (.csv)	Spss/R	Spss/R	Expected data volume is small (< 1 gb for total dataset)
WP3: Audiofiles	Child interviews	Atlas.ti	Qualitative	atlas	atlas	Expected data volume is small

						(< 1 gb for total dataset)

c. ***<u>In research involving personal data</u>: data privacy and security measures*** *to ensure compliance with legislation on personal data.*

- Does the informed consent form ask permission for preservation of personal data and sharing for verification and re-use?
- Is the collection of personal data minimized? What personal data do you explicitly require for your research?
- How and when will you de-identify personal data during the research?
- How will you prevent data breaches in every step of your research (e.g. multi-factor authentication, encryption, access management)?

In wp1, the original researchers are responsible for gathering informed consents of their participants.

in WP2 and WP3 we will ask participating parents and children > 12 years for permission to collect, use and reuse data. We will only collect those personal data that are necessary to answer our hypotheses (data of birth, sex of parent and child, mental health status, SES). The key between contact information and participant ID will be stored in a password protected excel file. Only the main researchers will have contact to this information. All other data will be pseudonymized.

d. ***Resources for data management*** *and ensuring data will be FAIR (Findable, Accessible, Interoperable, Re-usable).*

- How will the necessary resources be covered? Resources may include storage and/or archiving costs, devices, software licenses, data manager time, costs of preparing data for storage and archiving, repository charges.

Cost related to storing data will most likely be small. Cost related to software licenses are either covered by the ZonMw grant (software related to data collection), or Accare (SPSS).

Step 3: Data storage and archiving

This step aims to document how the data will be stored and backed up during the research, and how this will be done after the research is finished for the long term (archiving).

Please answer all sub-questions that are applicable to your research.

Questions:

a. ***Storage and back up*** *of data, metadata and documentation <u>during</u> the research.*

- What is the estimated size of the dataset?
- Where and how will 'raw' and 'processed' data, metadata and documentation be stored and backed up during research?
- Which storage medium will be used to store research data, metadata and documentation?
  - If you are using a storage medium other than, or in addition to the ones provided by the UMCG, does the selected storage medium fit the requirements of the Corporate Information Security Officer (CISO) of the UMCG? Please explain your choice.

The estimated size of the data in this project is small <1 gb.

Data	Project	Personal information	Storage location	Storage time
Demographic information	WP2	Yes	Pseudonymized Accare r-drive	15 years
Disruptive behaviors	WP2	no	Pseudonymized Accare r-drive	15 years
Parenting behaviors	WP2	no	Pseudonymized Accare r-drive	15 years
Parenting self-efficacy	WP2	no	Pseudonymized Accare r-drive	15 years
Parental attitude towards their child	WP2	no	Pseudonymized Accare r-drive	15 years
datafiles	WP1	yes	Rug Y-drive Y:\staff\umc\Psy\PAINT\PAINT-IPDMA	15 years
Raw audio files	WP3	yes	Pseudonymized Accare r-drive	Delete after transcription
Transcribed audio files	WP3	Yes	Pseudonymized Accare r-drive	15 years

b. ***Main security risks and measures*** *that will be taken to mitigate these risks.*

- How will access control be managed <u>during</u> the conduct of the research project (i.e. data collection and analysis)? Who will have access and who will authorize access?
- How will data be recovered in the event of an incident (e.g. stolen or lost laptop or external hard drive, or deleted personal cloud environment after your contract expires)?

Study specific folder can only be accessed by authorized personnel, who are involved in this study. Access to the study specific folder (WP1) at the RUG is arranged and maintained by Jaap Jansen (manager research) of the department psychiatry at the UMCG. Access to the study specific folder (WP2&3) is arranged and maintained by a research support person at Accare.

After the research project has been completed (i.e., data collection, data analysis and submission of research article completed) all the digital data will be transferred to a study specific folder for long-term storage on the RUG/Accare network. This folder can only be accessed by the Principal Investigator and by authorized personnel after approval of the Principal Investigator.

This data storage environment allows for careful access management and is only accessible with username and password and is compliant with RUG policies. Data storage is backed-up automatically every day on the servers of the RUG/Accare. This study will not produce non-digital data.

c. ***In research involving personal data, data security and protection measures*** *taken in order to take care of sensitive data during the research.*

- Where and how will personal data, the code list and coded data be stored?
- Who is responsible for authorizing access to the sensitive data?
- Which institutional data protection policies are in place?

See 3b.

d. ***Preserving data****: How will data for preservation be selected, and where and for how long will data be preserved for the long-term (i.e. data archive) after the research is finished.*

- Can you preserve all data resulting from the project or must some of the data be destroyed (e.g. due to contractual, legal or regulatory purposes)? If not all data will be kept, please explain why.
- Where will data be archived/preserved long-term? Include the link once available.
- For how long will you need to archive the data?

See 3b.

Step 4: Data documentation and metadata

This is a crucial step as 90% of making data FAIR, is about creating metadata. Good documentation of all research steps and procedures used and clear metadata will help other researchers understand your data for re-use and also yourself years from now. How can you ensure that your research data can be (re-)used by other researchers (in the near or distant future)? What methods/procedures, files and information are needed for other researchers to understand your research data in order to be able to (re-)use and/or reproduce it?

Please answer all sub-questions that are applicable to your research.

Questions:

a. ***Metadata and documentation*** *(for example, the methodology to collect and process data, and the way of organizing data) that will accompany your data to help other researchers re-use it*.

- What metadata and documentation (such as information describing how the dataset was generated and how to use it) are needed to enable re-use?
- What standards apply to your research discipline or field?
    - Which community metadata standards are applicable to your data?
    - Which terminology will be used for recording your data?
- How will the information required to enable the re-use of data be captured and where will it be recorded?
- How will data be organised during the project? Note that consistent, well-ordered research data will be easier to find, understand, and re-use
    - Which file naming strategy (or convention –if applicable to your data) will you follow?
    - Is there a version control system in place?
    - Is there a folder structure in place?

The protocol for the IPD is registered at PROSPERO (https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=355664). The protocol for We aim to preregister analyses plans for all papers, either at the open Science Framework or AsPredicted.

WP1 Data for each original study will be stored in a folder containing the first author of the published manuscript on this work. If multiple studies from a single author will be included, a subfolder with the year will be created. Within a study folder, there are two subfolders (i.e., raw, processed). The folder raw, will only contain the original data, the folder processed, will contain the processed data and code associated with this. Individual data sets will be merged when completed. Where possible we will follow the PRISMA guidelines for meta-analyses.

WP2
The protocol for the study will be registered at AsPredicted where other researchers can locate our procedures. We aim to preregister planned analyses for all papers, either at the Open Science Framework or AsPredicted.

All questionnaire data and data collected via the mobile application will be stored in one, pseudonymized file, that is accompanied by a ‘readme’ text file that contains a code book explaining the meaning of all variables. The file contains separate pages (“tabs”) for all different

measures at all different timepoints. A separate logbook-file will be created documenting all decisions that will be made during the process from raw to published data.

Old versions of all these documents will be stored, document names will contain their date to ensure version control. File naming will be done in a consistent way, and the structure of the data folder will be explained in a readme file.

Marijn Nijboer and the research assistant continuously update the data folders. The supervisor (Tycho Dekkers) will perform monthly checks to make sure data storage is in line with the data management plan.

WP3

Interview questions and qualitative methods will be described in osf.io. Data cannot be reused and will not be accessible to other researchers.

b. ***Data quality control measures****.*

- How will consistency and quality of data be ensured at each stage of the data life cycle?

For both projects, researchers continuously update the data folders. The supervisors will perform monthly checks to make sure data storage is in line with our data management plan.

Step 5: Data availability, access and re-use

This step documents what will happen with the research data once the project is finished.

Below each main question, a number of sub-questions are listed for Step 5 of the DMP. Please answer all sub-questions that are applicable to your research.

Questions:

a. ***Intellectual property rights and data ownership****.*

- Do the outcomes of the research need to be protected (e.g. with a patent, trademark or copyright)? How will intellectual property rights and data ownership be managed?
- What legislation and codes of conduct are applicable to the research project, and how will compliance with these be ensured?
- Who will be the data owner and who will manage access rights to data after the research?
- What will the terms of re-use and access conditions be?
- Are there any restrictions/agreements in place on the re-use of third-party data (e.g. conditions for the re-use of existing data obtained from Lifelines)? If so, please explain.
- In multi-partners projects and/or with multiple data owners, have the matters of rights to control access to data been defined in a consortium agreement?

WP1

Access permissions are controlled by the research project's principal investigator. The raw data containing identifiable information will be kept strictly separate from the processed data and can only be accessed by the supervisors and PhD student. Making data available for reuse will be discussed among the collaborators group on a paper-by-paper basis. If all collaborators agree, and if informed consent allows for this, anonymized data and scripts used for analyses will be made available on the open science framework.

WP2 & 3.

Access permissions are controlled by the research project's principal investigator. The raw data containing identifiable information will be kept strictly separate from the processed data and can only be accessed by the supervisors, research assistant and junior researcher. The processed, pseudonymised data will be made available restricted access for re-use. Requests for re-use of data will be evaluated by the steering committee who will check whether the research question falls within the scope of the informed consent.

b. ***Data availability****.*

- Do you think your data and samples are relevant for re-use? For what purposes could the data be re-used by other researchers in your department, in your research field, or in other contexts?
- Can the data be shared or made publicly available for re-use? If data can be made available, indicate what data will be considered for re-use and how will it be made available.
- How will you make your data available for re-use?
    - Will you be using a repository to make your data available? Which data repository will you use? Under which license?
    - When will the data be available for re-use? Does access to data need to be postponed/restricted due to intellectual property rights, patents or publishing?
    - What are the conditions under which you will share the data (e.g. open or restricted access, and what are the "terms of use")?
    - Will you need to adapt your data in order to make it available (e.g. repository)?
    - How will the tools or software (scripts, codes or algorithms) developed during your project be made available for other potential researchers to access, interpret and (re-)use the data?
- If data cannot or will not be made available in a repository, how will data be findable?

On a project by project basis we will decide whether it is possible to make data available for re-use. When we will make data available, we will store the deidentified file on OSF so that the file will get a DOI. We will link to the data in the published paper.

c. ***Findability of research data: Unique and persistent identifiers*** *to data.*

- Will a persistent identifier for the data will be pursued?

Data stored on OSF receive a DOI.

Research Data Management SOPs

Step 2: Describing your data

- SOP UMCG Guidelines for Processing Personal Data in Research
- SOP Processing Data UMCG
- SOP Data De-identification Guidelines
- SOP Data Management Infrastructure and Data Storage System

Step 3: Data storage and archiving

- SOP Data Management Infrastructure and Data Storage System
- SOP Preserve Data

Step 4: Data documentation and metadata

- SOP Metadata for datasets and variables in Research
- SOP Processing Data UMCG

Step 5: Data availability, access, and re-use

- SOP Data Transport
- SOP Data Management Infrastructure and Data Storage System
- SOP Preserve Data