Plan Overview

A Data Management Plan created using DMPonline

Title: Scan2Go

Creator: Dennis Klomp

Principal Investigator: Dennis Klomp, Esther Bron

Affiliation: UMC Utrecht

Template: UMC Utrecht DMP

Project abstract:

Most people experiencing cognitive decline want a timely and accurate diagnosis. They also want to know what to expect (prognosis) and seek treatment to prevent further cognitive decline, particularly since emerging treatments are considered most effective in early disease stages.

MRI is a powerful tool for diagnosis, prognostication, and assessment of effects of novel treatments. It is, however, relatively expensive as it requires experts for scanner operating and diagnostic reporting, which is subject to interpretation. From patients' perspective, MRI can be unpleasant and results difficult to understand. Our objective is to make MRI more accessible for timely dementia assessment at large by reducing costs ten-fold, improving patient friendliness, and facilitating automatic, comprehensible reporting.

We will design an easy-to-use autonomous MRI dedicated for diagnosing neurodegeneration and cerebrovascular injury consistent with (early-stage) dementia. By bringing together recent advances in hardware technology we will simplify operation and minimize anxiety so that people can scan themselves. Moreover, we will use artificial intelligence(AI) to interpret the images, with clearly understandable communication of results(Amsterdam-UMC).

Innovations include embedding high-conductive RF-coil arrays(TeslaDC) to facilitate automated patient couch(Inno-metaal) and metal detection for safe(UMCU) and autonomous MRI in Helium-free magnets(Philips), using quantified fast 3D acquisition(UMCU) with gradients(Futura) driven(Prodrive) above the human-hearing frequency to minimize anxiety, and AI(ErasmusMC) automated diagnostic reporting(Quantib) directly operated by patients outside the hospital(QuaRijn).

Synchronized to international programs and supported by industry-partners, we will deliver the demonstrator, ready for implementation/industrialization, that can provide sustainable diagnosis of early-stage cognitive impairments at large.

ID: 127287

Start date: 31-07-2023

End date: 30-04-2027

Last modified: 29-06-2023

Grant number / URL: can2Go_HH_04102022

Copyright information:

Scan2Go

1. General features

1.1. Please fill in the table below. When not applicable (yet), please fill in N/A.

DMP template version	30 (don't change)
ABR number *(only for human-related research)*	N/A
METC number *(only for human-related research)*	N/A
DEC number *(only for animal-related research)*	N/A
Acronym/short study title	Scan2Go
Name Research Folder	RFS
Name Division	Beeld&Oncologie
Name Department	highfield
Partner Organization	Erasmus AMC and Quarijn
Start date study	1-5-2023
Planned end date study	1-5-2027
Name of datamanager consulted*	Jacco van der Laan
Check date by datamanager	

1.2 Select the specifics that are applicable for your research.

- WMO
- Observational study
- Clinical study
- Multicenter study
- Prospective study
- Use of Questionnaires

2. Data Collection

2.1 Give a short description of the research data.

Subjects	Volume	Data Source	Data Capture Tool	File Type	Format	Storage space
human	500	eCRF	castor	text	.csv	5MB
human	500	devices	RDP	image	.dcm	500GB

Subjects	Volume	Data Source	Data Capture Tool	File Type	Format	Storage space
human	500	EPD (HiX)	Castor	Quantitative	.csv	0-10GB
human	500	eCRF/questionnaire	Castor	Quantitative	.csv	0-10GB
human	500	PACS (MRI)	RIA - CTP - Secured XNAT Health RI (Imaging database)	image	.dcm	500GB

All clinical and demographic data from participants will be collected from the EPD and entered in Castor. Questionnaires are being sent from Castor and also stored in Castor.

The MRI data from UMCU will be pseudonymized via the Research Imaging Architecture and when pseudonymized transfered to the central secured XNAT Health RI Imaging database where all imaging data from all centers will be collected.

2.2 Do you reuse existing data?

- Yes, please specify

Voor de ontwikkeling van AI voor automatische diagnose WP4 zal gebruik gemaakt worden van bestaande data van de Rotterdam Studie en de Alzheimer's Disease Neuroimaging Initiative (ADNI). Het gaat hier om MRI data, demografische en diagnostische gegevens. Voor de Rotterdam Studie is door alle deelnemers informed consent getekend en de studie is goedgekeurd door de METC van het Erasmus MC. Voor de ADNI studie zijn de gegevens publiek beschikbaar gesteld (https://adni.loni.usc.edu/).
Parallel hieraan zal nieuwe data gegenereerd worden.

2.3 Describe who will have access to which data during your study.

Type of data	Who has access
Direct identifying personal data	Research team with care relationship to patient, Datamanager
Key table linking study specific IDs to Patient IDs	PI (with care relationship to patient), Datamanager
Pseudonymized data	Research team, Datamanager

2.4 Describe how you will take care of good data quality.

#	Question	Yes	No	N/A
1.	Do you use a certified Data Capture Tool or Electronic Lab Notebook?	x		
2.	Have you built in skips and validation checks?	x		
3.	Do you perform repeated measurements?		x	
4.	Are your devices calibrated?	x		
5.	Are your data (partially) checked by others (4 eyes principle)?		x	
6.	Are your data fully up to date?	x		
7.	Do you lock your raw data (frozen dataset)	x		
8.	Do you keep a logging (audit trail) of all changes?	x		
9.	Do you have a policy for handling missing data?	x		
10.	Do you have a policy for handling outliers?	x		

2.5 Specify data management costs and how you plan to cover these costs.

#	Type of costs	Division ("overhead")	Funder	Other (specify)
1.	Time of datamanager	x		
2.	Design of eCRF		x	
3.	Data Capture Tool license fee		x	
4.	Questionnaire license fee	x		
5.	Storage		x	
6.	Archiving		x	

2.6 State how ownership of the data and intellectual property rights (IPR) to the data will be managed, and which agreements will be or are made.

UMCU is owner of the source data and the pseudonimized data is owned by UMCU, Erasmus MC, Amsterdam UMC and Quarijn.
We volgen hier de procedure voor pseudonymisatie van research data zoals vastgelegd in SOPs door ICT Divisie Beeld conform GDPR voor het Research Imaging Archive. Hierbij worden voor de MRI scans dmv CTP (https://www.health-ri.nl/preparation-phase-xnat)

herleidbare gegevens verwijderd. Vanuit het gepseudomyseerde RIA worden de research data naar de XNAT verstuurd.
De MRI data wordt gepseudonymiseerd en via beveiligde verbindingen (DICOM protocol, CTP) verzonden. Er is een verwerkersovereenkomst tussen UMCU en Health-RI voor opslag van de data in XNAT.

3. Personal data (Data Protection Impact Assessment (DPIA) light)

Will you be using personal data (direct or indirect identifying) from the Electronic Patient Dossier (EPD), DNA, body material, images or any other form of personal data?

- Yes, go to next question

I will process personal data for a complicated multicenter study. I therefore need to fill out a full DPIA and store the DPIA report of mitigating measures in the same folder as this DMP.

3.1 Describe which personal data you are collecting and why you need them.

Which personal data?	Why?
indication for disease	to train AI for image classification
MRI data	to investigate diagnostic performance
Gender, age	study population

3.2 What legal right do you have to process personal data?

- Study-specific informed consent

3.3 Describe how you manage your data to comply to the rights of study participants.

Right of Access	Research data are coded, but can be linked back to personal data, so we can generate a personal record at the moment the person requires that. This needs to be done by an authorized person.
Right of Rectification	The authorized person will give the code for which data have to be rectified.
Right of Objection	We use informed consents.
Right to be Forgotten	In the informed consent we state that the study participant can stop taking part in the research. Removal of collected data from the research database cannot be granted because this would result in a research bias.

3.4 Describe the tools and procedures that you use to ensure that only authorized persons have access to personal data.

1. We use the secured Research Folder Structure that ensures that only authorized personnel has access to personal data, including the key table that links personal data to the pseudoID.
2. We make use of a certified Electronic Data Capture (EDC) tool (Castor). To send surveys, email address will be used in the EDC, but this is encrypted for the users in such a way that users can send emails to subjects without seeing the actual email address. No personal data other than email address will be used in the EDC.
3. Patient digital imaging data for study purposes will be stored at the Research Imaging Archive (RIA) facility of the imaging

division of UMC Utrecht. For safe processing of images, RIA will be used (uses pseudonymization in order to guarantee safe processing). Only authorized personnel can access the (pseudonymized) imaging in the RIA container via personal login. The linkage table for the pseudonymized images will also be stored at the RIA. The container can only be accessed by users with the proper rights. Hospitals may transfer digital data into the RIA through secure connections. The RIA shields patient identifiable information through pseudonymized identifiers (i.e., study number) and only allows access to authorized researchers.

3.5 Describe how you ensure secure transport of personal data and what contracts are in place for doing that.

We have a signed consortium agreement with the UMCU, Amsterdam UMC, Erasmus MC and Quarijn that ensures the safe transport of data. In addition, we are setting up a more detailed contract between Quarijn (second location for obtaining new data) and UMCU (primary location of acquiring data) to ensure secure transport of data. (in progress)

4. Data Storage and Backup

4.1 Describe where you will store your data and documentation during the research.

The digital files will be stored in the secured Research Folder Structure of the UMC Utrecht.

4.2 Describe your backup strategy or the automated backup strategy of your storage locations.

1. During data collection, automatic backups will be made in the Electronic Data Capture Tool Castor. Upon completion of data collection, all data are exported and saved in the Research Folder Structure where they are automatically backed up by the UMC Utrecht backup system.
2. Data in Research Imaging Archive (RIA) is stored in two data centers in the UMC Utrecht that are synchronized hourly. These centers are present at different locations within the UMC Utrecht.

5. Metadata and Documentation

5.1 Describe the metadata that you will collect and which standards you use.

1. For the data collected in Castor, a codebook of my research database is available in Castor.
2. During the project, we will make the data available via HealthRI that includes the description of metadata.

5.2 Describe your version control and file naming standards.

File naming: subjects included will have a VU- or VQ-number that is increased by one after every new inclusion (V = volunteer), U= UMCU, Q is Quarijn.
Github version control is used for scripts.

6. Data Analysis

6 Describe how you will make the data analysis procedure insightful for peers.

1. I will make an overview of datasets and analysis scripts, such that it is fully clear how the statistical analysis is performed. Peers will be able to repeat the analysis based on my overview.

7. Data Preservation and Archiving

7.1 Describe which data and documents are needed to reproduce your findings.

1. The data package will contain: the DICOM data, the study protocol describing the methods and materials, the script to process the data, the scripts leading to tables and figures in the publication, a codebook with explanations on the variable names, and a 'read_me.txt' file with an overview of files included and their content and use.

7.2 Describe for how long the data and documents needed for reproducibility will be available.

1. In view of the regulation for Clinical Trials, I need to store all data for at least 15 years with the goal to be able to go back to patient level.

7.3 Describe which archive or repository (include the link!) you will use for long-term archiving of your data and whether the repository is certified.

1. After finishing the project, the data package will be stored at the UMC Utrecht Research Folder Structure and is under the responsibility of the Principal Investigator of the research group. Data will be published on the UMC Utrecht repository DataverseNL.

7.4 Give the Persistent Identifier (PID) that you will use as a permanent link to your published dataset.

I will be using a DOI-code and will update this plan as soon as I have the code

8. Data Sharing Statement

8.1 Describe what reuse of your research data you intend or foresee, and what audience will be interested in your data.

1. The raw data can be of interest for other researchers or for spin off projects.

8.2 Are there any reasons to make part of the data NOT publicly available or to restrict access to the data once made publicly available?

- Yes (please specify)

1. In the data repository, the pseudomised data will be available under a CC0 license, with terms of use that require the user to cite the data. To minimize barriers for reuse, the data is available without restrictions to access, immediately after the publication of the last paper.

8.3 Describe which metadata will be available with the data and what methods or software tools are needed to reuse the data.

all pseudomised data will be made available including the scripts that can view or process the data.

8.4 Describe when and for how long the (meta)data will be available for reuse

- (Meta)data will be available as soon as article is published

8.5 Describe where you will make your data findable and available to others.

We will use the policies of HealthRI to make the data findable, while the link will be mentioned in the paper that will be published prior to releasing the data.