Project Acronym: BRAINSTORM

Project title: gloBal RetinAl Imaging coNSorTium fOR alzheiMer's disease

DATA MANAGEMENT PLAN

HISTORY OF CHANGES		
Version	**Publication date**	**Change**
1.0	05.01.2022	▪ Initial version

1. Introduction

This document contains the initial Data Management Plan (DMP) for BRAINSTORM. The DMP is prepared as part of the Work Package (WP) 4: Coordination, project management and dissemination and is labelled as 'Deliverable D.4.1'. The DMP outlines the data management approach that will be used throughout the project. It will be developed based on FAIR data sharing guidelines and data management rules of the partner institutions. The DMP aims to ensure that the data are collected, structured, curated and preserved according to the FAIR (Findability, Accessibility, Interoperability, and Reusability) principles, and are made available according to the "as open as possible, as closed as necessary" principle, in compliance with the Belgian and European legal standards, and notably the General Data Protection Regulation[1] (hereafter "GDPR"). The DMP implementation will be regularly evaluated by the lead investigators at each site and a log of modifications will be updated every time changes are made to the DMP. Lead investigators of each partner will oversee local data management.

Lead investigators at each site, overseen by the local human research ethics committee/s, will be responsible for the development of informed consent form and participant information sheet that meets Good Clinical Practice Guidelines and requirements of the country where consent is obtained. The lead investigators will coordinate with each other to determine whether and how the data collected and generated by the project are shared and become available for re-use. Data quality assurance and the availability of openly available data will also be managed by the lead investigators.

Lead investigators at each site will act as **data controllers**[2] of the data collected/generated in the frame of the tasks that fall under their leadership, determining the purposes and means of processing this data as well as safeguarding its appropriate and timely processing. If necessary, they will undertake any necessary actions to prepare the data collected/generated through the tasks they are leading for sharing either within the consortium or openly (e.g., use of proper naming conventions, application of suitable anonymisation techniques, the creation of appropriate metadata and documentation, etc.). Pursuant to Article 26 GDPR (the "Joint Controllers"), the partners have determined their respective responsibilities for compliance with the obligations under the GDPR in Annex D of the Consortium Agreement (CA).

[1] Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data and repealing Directive 95/46/EC (General Data Protection Regulation), OJ L 119, 4 May 2016 (see: https://eur-lex.europa.eu/eli/reg/2016/679/oj)

[2] According to the GDPR data controller means the natural or legal person, public authority, agency or other body which, alone or jointly with others, determines the purposes and means of the processing of personal data.

2. Data Summary

The collection of all data within the BRAINSTORM project happens on a lawful basis in the context of academic research. Academic research is carried out in the public interest, this means that it is conducted to contribute to an increase of knowledge and insight that will benefit (directly or indirectly) society.

The scope of the BRAINSTORM project is to evaluate retinal hyperspectral, photographic, optical coherence tomography (OCT), and OCT-angiography biomarkers of Alzheimer's disease (AD) in persons with, or at risk of, AD. Retinal imaging findings will be compared with CSF, neuroimaging, blood, and cognitive biomarkers of AD and to establish whether common age-related eye diseases have hyperspectral imaging signals that are distinct from those of AD (WP1). The project will also identify retinal imaging biomarkers of cerebrovascular disease (WP2), and determine pathological correlates of retinal imaging biomarkers of AD in preclinical models and post-mortem human tissues (WP3).

This research will re-use, collect/generate and analyse the following type of data:

Human data:

1. Observational data: patient details (clinical measurements, medical/ electronic health records)
    a. Regular personal data:
        - Name and address (human).
        - Contact details (tel. number, e-mail address, ...) (human).
        - Year of birth and age (human).
        - Personal identification number assigned to data subjects participating in the study such as EAD number (human).
    b. Sensitive personal (health) data
        - Description of characteristics of physical features of the body, the retina and systemic microvasculature in particular, further explained below.
        - Medical history, medication use and medical test information (such as blood samples results from scans and biopsies).
        - Family history, medical history, medication use, personal history – nicotine and medical test information, neuropsychological performance, cognitive reserve score.
        - Ocular data: visual acuity measurement, biomicroscopy and funduscopy, fundus photographs, hyperspectral retinal images, Dynamic Vessel Analyser (DVA), Optical Coherence Tomography (OCT) retinal images and OCT angiography (OCT-A) retinal images.
2. Experimental data: digital images (DICOM) including brain images (MRI and PET scans), retinal images (hyperspectral, OCT, OCT-A, colour photographs; microscopy images of post-mortem retina), video files, oximetric data, quantified imaging device outputs, protein gels and ELISAs.

3. Derived and compiled data: research documentation (text, spreadsheets, protocols, notes and diaries), manuscripts, statistical output, algorithms and scripts, software.
4. Biological samples: post-mortem retinal and brain tissue from humans and rodents.

Experimental data:

- Experimental data: blood pressure, blood flow measurements, histology, whole-mount images, and microscopy images, protein gels, DNA gels, gene expression analysed by qPCR, ELISA results, cognitive testing of mice, levels of metabolites in pericytes, DVA measurements (quantified imaging device outputs); brain images (MRI and PET scans), retinal images (hyperspectral, OCT, OCT-A, colour photographs; microscopy images of post-mortem retina), video files (hyperspectral), quantified imaging device outputs.
- Derived and compiled data: research documentation (text, spreadsheets, protocols, notes and diaries), manuscripts, statistical output, algorithms and scripts, software.

The processing of data can be regarded as primary (prospective trial data) and secondary (retrospective data from the electronic health records of participants) and is derived from participants (healthy volunteers and patients) of the University Hospitals of Leuven, Umeå University Hospital and Centre for Eye Research Australia and its collaborators. The subjects will be informed of the processing of their data in case of secondary processing as part of the informed consent process for study participation.

Patient data will be digitally recorded in the form of questionnaires, clinical examination and image files, as well as body fluid (blood, CSF) and tissue (post-mortem brain and retinal tissue) samples. Each partner will utilise their institute's secure data storage system with automated onsite back-up and mirroring.

Partner 1: In Melbourne, participants will be assigned a study identification number for entry onto an Electronic Case Report Form (eCRF) and into a secure REDCap research database, maintained in compliance with CERA standard operating procedures.

Partner 2: In KU Leuven (P2), numeric data will be stored in REDCap, the logged Electronic Case Report Form (eCRF) that is approved and supported by the UZ Leuven Clinical Trial Centre and Ethics Committee.

Partner 3: In Umeå, participants will be assigned a study identification number. Data will be recorded on Electronic Case Report Forms (eCRF) at Umeå University Hospital in accordance with standard operating procedures and ethical approval.

Written reports of the project progress will be stored for internal purposes. Methodological and clinical manuscripts will be drafted to disseminate the results from the proposed project. A detailed description of human data collected in the study can be found in the tables below:

Partner 1 - The Centre for Eye Research Australia

Mode of acquisition	Type of data	Format	File size	Field of view	Image dimension	Pixel size	Bit-depth
Metabolic Hyperspectral Retinal Camera (Optina)	Hyperspectral retinal imaging	.h5	430MB per image (8 images per patient)	30-degree fov	1099 x 1099	8.2µm x 8.2µm	16
CERA's custom hyperspectral camera	Hyperspectral retinal imaging	.h5	550MB per image (2 images per patient)	38-degree fov	1500 x 1500	7.8µm x 7.8µm	12
Plex Elite 9000 (Zeiss)	OCT-angiography	.TIF	1MB per image (4 images per patient)	56-degree fov	1023 x 1023	17.67µm x 17.67µm	8
Canon CR6-45NM Optometry Non-Mydriatic Retinal Fundus Camera	RGB retinal imaging	.TIF	50MB per image (4 images per patient)	45-degree fov	5183 x 3455	2.71µm x 4.08µm	24
Spectralis (Heidelberg Engineering)	Spectral-domain optical coherence tomography	.vol	200MB per image (2 images per patient)	55-degree fov	1535 x 1535	11.53µm x 11.53µm	8
Visit (general ophthalmological)	Ophthalmological evaluation	eCRF	Some kB	N/A	N/A	N/A	N/A
Mini-Mental State examination, Clinical Dementia Rating Staging Instrument, Clock Drawing, Stroop colour word test,	Neuropsychological evaluation	Medical file to eCRF	Some kB	N/A	N/A	N/A	N/A

Controlled Oral Word Association Test, Category fluency, Rey Complex Figure Test, Digit Span, Digit Symbol Substitution Test, Boston Naming Test, California Verbal Learning Test-II, Weschler Test of Adult Reading							
3D MPRAGE, 3D T2-weighted FLAIR image, 5 echo Gradient Recalled Echo (GRE) images for susceptibility-weighted imaging and quantitative susceptibility mapping	Magnetic Resonance imaging	DICOM	Multiple MB to GB	N/A	N/A	N/A	N/A
Philips Allegro and Gemini TF64, Siemens mCT	PET scan	DICOM	Multiple MB to GB	N/A	N/A	N/A	N/A

Partner 2 - Ku Leuven

Type of data	Format	Volume	How created
RGB retinal imaging	.TIF	12,613 MB/image (2 images / patient)	VISUCAM (Zeiss)
Hyperspectral retinal imaging	.TIF	2,2 MB/image (5 images / patient)	XIMEA (Imec)
Hyperspectral retinal imaging	.h5	550MB per image (2 images per patient)	CERA's custom hyperspectral camera
OCT(A) imaging	.E2E	220 MB/image	OCT(A) Spectralis (Heidelberg Engineering)
Oximetry retinal imaging	.ret	5 MB/image (1 images / patient)	OXYMAP
Dynamic Vessel Analysis Video	.dv	1-1,5 GB/video (1 video / patient)	DVA (IM-RC 3.0;Imedos)
Dynamic Vessel Analysis graphs	.rva	1,5-10 MB/ image (1 image / patient)	DVA (IM-RC 3.0; Imedos)
mRNA	Blood sample (EDTA, serum, heparine)	some kB	Blood sample
Ophthalmological evaluation	medical file to eCRF	Some kB	Visit (general ophthalmological)
Neuropsychological evaluation	medical file to eCRF	Some kB	MOCA, RAVLT, CST, Stroop, COWAT
Magnetic Resonance imaging	DICOM	Some kB	3T Achieva dstream Philips MRI scanner

Device	Dimension	Image Type	Bit-depth	Pixel Size
VISUCAM	2048 x 1958	RBG retinal imaging	24	96 dpi, 96 dpi

XIMEA	2048 x 1088	hyperspectral retinal imaging	8	96 dpi, 96 dpi
CERA's custom hyper-spectral camera	1500 x 1500	hyperspectral retinal imaging	12	7.8µm x 7.8µm
OXYMAP	NA	oximetry retinal imaging	NA	NA
Spectralis	NA	OCT(A) imaging	NA	NA
DVA	NA	Dual vessel analysis	NA	NA
OCT(A) (Zeiss Cirrus)	NA	OCT(A) imaging	NA	NA

Interview Opportunit-ies	**Setting**	**subjects discussed**
Visit (general ophthal-mological)	In UZ Leuven	General ophthalmological evaluation
Neuropsychological evaluation	In UZ Leuven	Montreal Cognitive Assessment (MOCA), Auditory Verbal Learning Test (AVLT), Stroop Colour Word Test, COWAT/semantic fluency

Partner 3 - Umeå

Type of data	**Format**	**Volume**	**How created**
Hyperspectral retinal imaging	.h5	550MB per image (2 images per patient)	CERA's custom hyperspectral camera
OCT(A) imaging	.E2E	220 MB/image	OCT(A) Cirrus (Zeiss)
Ophthalmological eval-uation	medical file to CRF	Some kB	Visit
Magnetic Resonance imaging	DICOM	Some kB	3T General Electric MRI scan-ner

Device	**Dimension**	**Image Type**	**Bit-depth**	**Pixel Size**

CERA's custom hyper-spectral camera	1500 x 1500	hyperspectral retinal imaging	12	7.8µm x 7.8µm
OCT(A) (Zeiss Cirrus)	NA	OCT(A) imaging	NA	NA

2. FAIR data

Through the life cycle of the BRINSTORM data, the FAIR principles will be followed as far as possible, while ensuring compliance with national and European ethic-legal framework. The FAIR data approach is described by the acronym: ***F****indable,* ***A****ccessible,* ***I****nteroperable and* ***R****eusable data*.

2. 1. Making data findable, including provisions for metadata

Data will be stored in REDCap, the logged Electronic Case Report Form (eCRF) that is approved and supported by each institution. Information included on the eCRF include:

- Enrolment and consent of the participant
- Medical and Ocular history
- Vision Preliminary Tests
- Devices used in retinal imaging and image quality score

Furthermore, the data structure of REDCap foresees metadata of device acquisition being for each device type separately:

- Whether or not the device was used
- The date of image acquisition
- The image quality score
- The operator of the device
- Whether or not any problems or malfunction occurred and if yes what kind

All columns carry a specific and logic code which specifies the parameter that is described and the possible answers in REDCap are limited to prefilled options which limits error in the final database. All imaging acquisition is done by means of a Standards of Procedure (SOP) that is made available to all study personnel and is updated regularly.

2.2. Making data openly accessible

Prior to publication, data sharing is restricted to members of the BRAINSTORM consortium. All research outputs sampled through the BRAINSTORM project will be made openly accessible, whenever possible, via existing platforms that support FAIR data sharing (www.fairsharing.org). Personal data will only be published after de-identification and identifiers will not be published. If, despite all efforts, it is not possible to protect the identities of subjects even after removing all identifiers, personal data will not be made public. Participants in the study are informed via consent forms about the policies regarding data sharing. The ethical aspects are covered in the context of the ethics review at each institution. The storage and transfer of data on human subjects to the repositories used by the consortium will only be considered in case of informed consents, ethics approval, compliance with GDPR and, when applicable, approval by local data protection authorities. Any controls or limitations on access to or usage of human data is described below in the ethics section of this DMP (Section 7).

2.3. Making data interoperable

The work packages are comprised of separate research lines, each focusing on another research pillar and question. There is no need for pooling of the data afterwards, hence different ontologies or vocabularies will arise. That being said, international approved terminology is used when describing parameters. These nouns are validated by the respective international societies (European Society of Ophthalmology, European Academy of Neurology). Furthermore, the monthly meetings will streamline terminology and specific attention will be attributed to the mappings towards the most common ontology.

2.4. Increase data re-use (through clarifying licences)

As explained in Section 2.2 all research outputs will be made openly accessible, whenever possible, via existing platforms that support FAIR data sharing (www.fairsharing.org). The Consortium Agreement will further clarify access rights for implementation, use and exploitation of data.

3. Allocation of resources

Each partner institution will be responsible for and pay for the costs relating to the back-up and secure storage of the data generated and received by that organisation and each partner will be solely responsible for digital file storage and management of that data.

Fee for publication in open access journals is either foreseen in the Partner's budget or carried by the respective Research Team based on external funds.

Each partner is responsible for the implementation of this Data Management Plan.

4. File naming

Separate folders will be made for each participant and visit. Participants will be sorted by their study identification number and the naming strategy for most files will be: SubjectID_Eye_ImageType_ImageDate YYYYMMDD.

5. Processing of data

Personal data will be processed[3] in the context of the research, as per Section 1. The research will comply with the rules pertaining to the processing of personal data provided in the GDPR. All partners will register the research activities that involve personal data according to the rules and procedures at their own institution. In addition, in application of the European Regulation EU 2016/679 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data (GDPR), for the purpose of the BRAINSTORM, the partners jointly determine the purposes and means of the processing of the data and are joint controllers within the meaning of Article 26 GDPR (the "Joint Controllers"). Pursuant to Article 26 GDPR, the partners have determined their respective responsibilities for compliance with the obligations under the GDPR in Annex D of the Consortium Agreement (CA).

[3] For the purpose of this DMP and of the entire research project, processing "means any operation or set of operations which is performed on personal data or sets of personal data, whether or not by automated means, such as collection, recording, organisation, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction" (Art. 4.2 of the GDPR).

The obtained experimental data will be stored and processed for purposed of better interpretation and the processing methods planned to be used will be:

1. Parametrical, cross-sectional statistics that can be executed after the first visit.
    - The strength of association between ophthalmic measures and the relevant ground truth (PET or CSF amyloid beta levels for instance) will be evaluated.
    - A Receiver Operating Characteristic (ROC), Pearson's correlation and Shapiro-Wilk test will be used to analyze the data.
2. Machine learning, cross-sectionally that can be executed after the first visit.
    - The purpose of this is to derive hyperspectral imaging estimates of the parameters of interest – such as brain amyloid beta burden.
    - Methods such as Savitzky-Golay spectral smoothing procedure to compute and smooth the images, and a machine learning method - Dimension Reduction by Orthogonal Projection for Discrimination (DROP-D) will be used to discern between Aβ PET− controls from PET+ cases on the basis of spectral data. Deep learning networks will also be used.
3. Parametric, longitudinal statistics that will model the follow-up repeated measurements.
    - The purpose of this is to identify temporal changes in biomarker values.
    - Linear mixed models and Pearson's correlation will be used.
4. Recurrent deep learning to derive information hidden within the time-sequential imaging
    - The purpose of this is to enhance data analysis of the above mentioned if conditions are met, and to get information on each image and the period in between.

The precise methods of analysis to be used in the project are likely to evolve during the course of the study in accordance with preliminary findings. The investigators undertake to utilise the optimal analytical and statistical methods for each research question.

6. Data security

Partner 1: In Melbourne, participants will be assigned a study identification number for entry onto an Electronic Case Report Form (eCRF) and into a secure REDCap research database, maintained in compliance with CERA standard operating procedures. Data for this study will be collected using REDCap, a secure web application for building and managing online surveys and databases. The REDCap platform consists of a MariaDB relational database and a web server. These servers are in a physically secure location on premise at the Centre for Eye Research Australia and managed by the CERA Information Technology team. Frequent backups are made daily and stored in accordance with CERA's retention schedule which incorporates both cloud and offsite storage. The servers provide a stable, secure, well-maintained, and high-capacity data storage environment.

Clinical and imaging data taken on the day of the visit will be automatically saved to the local hard drive of the imaging equipment and to the secure server of the Centre for Eye Research Australia. Other data such as clinical, imaging and fluid biomarker data received from the recruiting site for each participant will be stored on CERA's secured network as part of the data collected during the study. Data on the server is mirrored in two locations, making data loss extremely improbable. Data will be stored in a password protected file. No identifying data will

be shared with any external party. Analysis will be performed on deidentified images. The images with key identifiers will be saved on the server in a master file. Records will be kept for 15 years after the completion of the study, in accordance with the requirements of the Therapeutic Goods Administration and Health Privacy Principals.

Partner 2: In KU Leuven we will utilise our institute's secure data storage system (University Hospitals Leuven servers) with automated onsite back-up and mirroring. All data will be examined and kept in a coded format, with each participant assigned a unique anonymous identification. Biological samples will be stored in locked facilities and will be accessible only to the relevant researchers. Acquisition, processing and storage of patient data related to the clinical studies will be supervised by Prof. I. Stalmans. Technological developments and source codes will be documented under the supervision of Prof. I. Stalmans.

Numeric data is saved in online servers (following GDPR legislation) in an online electronic case report form (eCRF), RedCap. Image data (VISUCAM, OXYMAP, XIMEA, DVA, OCT(A) will be stored on the internal UZ Leuven Database and an external coded (password only known by PI and delegate) hard disc.

Data will be preserved for a period of at least five years from the end of the research project, except for clinical trial data. Datasets collected in the context of clinical research, will be archived for a minimum of: 10 years in Sweden (Archives Act SFS 1990:782), 15 years in Australia (Australian Code for the Responsible Conduct of Research) and 25 years in Belgium (European Regulation 536/2014).

With the exception of the data related to the clinical trials, which will be solely stored on the servers of the University Hospitals Leuven, as mandatory by the UZ Leuven ethical committee, all data will be stored and duplicated on the secure FTP servers hosted by KU Leuven and ESAT. Personal data will only be published after de-identification and identifiers will not be published. If, despite all efforts, it is not possible to protect the identities of subjects even after removing all identifiers, personal data will not be made public.

Partner 3: In Umeå we will use the institute's secure data storage system with automated onsite back-up and mirroring. All extracted data will be examined and kept in a coded format, with each participant assigned a unique study ID. The code key will be stored in a locked space. Acquisition, processing and storage of patient data related to the Brainstorm project will be supervised by Ass.Prof G.Johannesson.

Clinical data that is acquired from participants in the clinical study and is filled in electronic Case Report Forms (CRF), REDCap. Imaging data taken on the day of the visit will be automatically saved to the local hard drive of the imaging equipment and to the secure server of the Umeå University Hospital as well as on an external coded hard drive (password only known by PI and delegate).

Data that will be shared with CERA for analysis will be pseudonymised and will be sent through secure methods such as Cryptshare in order to ensure the protection of personal data.

7. Ethical aspects

For personal and sensitive data, we will abide by local laws, the General Data Protection Regulation 2016/679 and US FDA Good Clinical Practice guidelines. The Data Management Plan will be approved by institutional research governance committees and ethics review

authorities. As stipulated in the informed consent (IC), no patient information will be disclosed to third parties, nor will it be used for any purpose other than described in the IC. All data will be analysed and stored in a coded fashion, with a unique anonymous identifier for every subject that is stored separately from the research data and only accessible by key researchers.

Study protocols and methods will be approved by the ethics committees governing the jurisdiction of each consortium partner and all studies will be conducted in accordance with the Declaration of Helsinki. Informed consent will be obtained from all participants following a thorough explanation of test procedures. Research on human biomaterials will be compliant with ICH-GCP E6 Guideline principles, the EU Directive 2004/23/EC and related national legislations. All human biological material will be registered in the Biobank of the Leuven University Hospital and Centre for Eye Research Australia. Following study completion extracted and/or derived materials will be stored for the legislated duration applicable in each study site jurisdiction. Biobank samples will be anonymised.

8. Other issues

No other issues are foreseen within scope of this project.