**Data Management Plan**

***This data management plan takes into consideration i) primary MRI data collected on the University’s scanner; ii) secondary MRI data from the University’s repository, and iii) questionnaires. DON’T FORGET to account for other forms of data you might be using.***

Insert name of project:

1. **Background**

**1.1 Aim and study design**

Write a very concise description of 1) the aim of your project, and 2) the research methodology.

**1.2 Principal investigator and lead researcher**

The principal investigator is [insert name here] and the lead researcher is [insert name here].

*In the case of student projects, the principal investigator is usually the supervisor and the lead researcher is the student doing the project.*

1. **Data Summary**

**2.1 Types of data**

*If using primary data:*

This project will make use of primary medical resonance imaging (MRI) data, acquired in the DICOM format and subsequently transformed into other standard formats (such as NIfTI and GIFTI).

*If using secondary data:*

This project will make use of secondary medical resonance imaging (MRI) data, available in the DICOM or NIfTI formats and subsequently transformed into other standard formats, such as NIfTI (if originally DICOM) and GIFTI.

*If using questionnaires:*

Data will also be collected by means of questionnaires in paper/electronic format. *or*

Secondary data in the form of filled-in paper/electronic questionnaires will also be utilised.

**2.2 The origin of the data**

*If using primary data:*

MRI data will be acquired on the Siemens MAGNETOM Vida3 Tesla scanner belonging to the University of Malta (UM) and housed at Mater Dei Hospital.

*If using secondary data:*

MRI data will be acquired from the repository owned by the University of Malta (UM) and managed by its MRI platform, UMRI.

*If using questionnaires:*

Questionnaire text has been sourced from [provide details].

*If the questionnaires were given to you already filled in, specify their source (i.e. the repository or data set from which you obtained them) and the project they were originally produced for.*

**2.3 The expected size of the data set**

Provide an estimate of the number of data subjects and the size of the data set per subject; the latter may be calculated by employing the following rule of thumb:

* Structural MRI data: 30 MB
* Functional MRI data: 700 MB
* Diffusion MRI data: 180 MB

**2.4 Data processing**

*If using primary data:*

After collection, the MRI data will be transferred through the pseudo-anonymising pipeline established by UMRI, the University of Malta MRI platform, to the repository owned by UM and managed by UMRI. Only then will it be released to the researchers working on this project.

*If using secondary data:*

The MRI data received from the UM repository would already have been pseudonymised (this is done prior to storage).

*In either case, add (text might need to be customised):*

When it reaches the researchers, the data will first be quality-controlled with the software MRIQC. Any scans that are heavily compromised will be eliminated, while the rest will be preprocessed with standard software such as fMRIPrep. Preprocessing gets the data ready for the ensuing analysis, and also improves its quality by correcting for susceptibility distortion and motion.

*Briefly describe the next steps (including how questionnaire results, if any, will be incorporated into the analysis).*

1. **Data Security***(If using primary data)*

**3.1 Organisational measures with an emphasis on data minimisation**

*Data collection*

**No names or ID card numbers will be entered into the scanner software.** Instead, each research volunteer will be assigned a code (PatientID). Once all MRI data has been acquired, only the principal investigator will retain information that links the code to the name of the volunteer. This information will be kept secure, confidential and separate from the scans and questionnaires (delete ‘questionnaires’ if not relevant), and will be disposed of one year after the last day of data collection. It will not be shared with any third parties, except with the designated UMRI staff in the eventuality of an incidental finding (in which case the identity of the participant concerned would have to be established so that they could be notified).

While a copy of the MRI data on the scanner is transmitted to the UM repository shortly after collection, the original data (which would not yet have been pseudonymised) remains in place for about 4 weeks before being automatically deleted (the exact duration depends on how many people are scanned and the amount of data acquired). During this time, physicists, Magnetic Resonance radiographers and the system engineer have access to the data.

Consent forms will be retained by the principal investigator for a period of five years or for the duration of the project, whichever is the shorter. The unique code assigned to each volunteer will not be written on their consent form.

*Data storage - UMRI repository*

*i) Access*

The data collected for this project will be transferred to the UM MRI repository. Data in the repository is made available on a strictly “as needed basis”, and only to perform duties and/or research authorised by UM. This extends to any personal metadata – like sex, weight and height – that is sometimes needed in research and is therefore not removed during pseudonymisation. Any such personal details are not automatically released with the scans, but are solely forwarded to researchers if they are essential for their project.

Specifically, the copy of the data in the repository will be accessible to: 1) the research team working on this study\*; 2) UM employees administering the system; and 3) after a 2-year embargo (which may be shortened at the discretion of the principal investigator), possibly (i) teaching personnel at UM for teaching activities (and students attending their courses); (ii) other researchers and students at UM involved in approved projects (examiners assessing student projects might also need access, usually for purposes of verification); and (iii) authorised external entities for research purposes - in which case any sharing of data would be subject to a data sharing agreement incorporating the safeguards that apply to the data stored in the University repository;

Research volunteers participating in a study will be able to obtain access to their own data upon a formal request to UMRI. Such requests will be handled by the principal investigator of the study who can link the volunteers’ ID (patientID) to their name.

**All applications for doing research using the collected data will have to be approved by a recognised ethics committee, as well as by the information and data protection commissioner.**

\*The raw data, as well as derivatives that are not in aggregate form, can be shared within UM provided that this is done exclusively i) by the principal investigator ii) with staff or students involved in the study, and iii) using UM-managed infrastructure or, if not possible, encrypted storage devices.

*If data will be shared with external entities, add:*

Sharing with [name of external entity] will take place subject to the terms and conditions specified in the attached **data sharing agreement**.

*In case the principal investigator is affiliated with an external institution, replace the above paragraph with:*

\*The raw data, as well as derivatives that are not in aggregate form, can be shared within [insert name of principal investigator’s entity of affiliation] provided that this is done exclusively i) by the principal investigator ii) with staff or students involved in the study, and iii) using encrypted storage devices. Such sharing is covered by the attached **data sharing agreement**.

*If data will be shared with a third external entity, add:*

Sharing with [name of external entity] will take place subject to the terms and conditions specified in the attached **data sharing agreement**.

*In all cases, add:*

ONLY individuals who have completed and submitted UMRI’s researcher compliance form will be given access to the data.

*You should be very clear about who will have access to your copy of the data. For instance, if you plan to save this copy on a computer cluster at the University, the data would also be accessible to the system admins.*

*If your research study includes questionnaires, make sure you specify*

1. *Whether they will be online (e.g. as a Google form) or in paper format*
2. *How you will safeguard data privacy e.g. by marking each questionnaire with the participant code, rather than writing down names/ID card numbers; by putting in the F/REC study code part of the participant code yourself (in the case of online questionnaires), so that research volunteers only need to supply the participant number – refer to the document ‘Information sheet for UMRI scanning service applicants’ for more details about the way the participant code is constructed.*
3. *Where the filled-in questionnaires will be stored (e.g. in the UM repository) and for how long.*

*ii) Retention period*

UM will store the data for a period of 10 years from the date of last use. After 10 years, it will be deleted if no longer deemed necessary for scientific research purposes, although data that has been approved for use in certain projects may be kept longer.

*Data storage - individual researchers*

Any raw data, as well as derivatives that are not in aggregate form, will be erased from computers/personal storage systems once the project has been completed – unless the data is needed for a follow-up study, in which case a new application will be submitted to UMRI, and the data will not be further processed or shared until authorization from UMRI and ethical clearance from F/REC [the relevant (Faculty) Research Committee] have been obtained.

**3.2 Technical measures**

*UMRI repository*

UMRI has put a lot of effort into ensuring that the data transmission pipeline and storage servers are secure and protected against unauthorised access, and that the repository only holds properly pseudonymised data. An additional layer of protection has been implemented by adopting the principle of duty segregation in relation to UM staff having access to sensitive components of the pipeline.

Data collected on the MRI scanner is first transferred from the hospital’s internal network to UM’s through an encrypted VPN tunnel. Once at UM, it is automatically pseudonymised by a compute node known as the Anonymiser prior to being stored in the repository.

The main data security measures in place are summarised below:

* **Pseudonymisation Server:** Technical access granted to a few named UM employees, regulated physical access, encryption of data storage partitions, restricted network exchanges, internal firewall.
* **Cold Storage Repository:** Technical access limited to a small number of authorised staff, regulated physical access, encryption of data at rest, firewall protection.
* **Research Repository:** This repository holds copies of data released for research and has similar safeguards to the Cold Storage Repository. It is envisaged, however, that researchers/teaching personnel will eventually be given access to it via user accounts that restrict said access to the data needed. The research repository does not contain encrypted hashes of ID card numbers.

*Individual researchers*

Data will be processed by UMRI personnel to get it ready for release prior to being made available to researchers. This administrative processing will include the removal of personal attributes not necessary for the research. In the case of this project, the data will be scrubbed of any such attributes except for the data subject’s age/sex/height/weight/scan date [choose as appropriate], since this/these is/are needed [specify why].

Raw data (and any derivatives not in aggregate form) will be:

1. stored on an encrypted data storage device;

 2) processed, analysed and viewed on computers which are password-protected, have appropriate antivirus software installed *(if applicable)*, and are never connected to an unsecured network.

**3. Data Security** *(If using secondary data)*

**3.1 Organisational measures with an emphasis on data minimisation**

The raw data, as well as derivatives that are not in aggregate form, can be shared within UM provided that this is done exclusively i) by the principal investigator ii) with staff or students involved in the study, and iii) using UM-managed infrastructure or, if not possible, encrypted storage devices.

*If data will be shared with external entities, add:*

Sharing with [name of external entity] will take place subject to the terms and conditions specified in the attached **data sharing agreement**.

*In case the principal investigator is affiliated with an external institution, replace the above paragraph with:*

\*The raw data, as well as derivatives that are not in aggregate form, can be shared within [insert name of principal investigator’s entity of affiliation] provided that this is done exclusively i) by the principal investigator ii) with staff or students involved in the study, and iii) using encrypted storage devices. Such sharing is covered by the attached **data sharing agreement**.

*If data will be shared with a third external entity, add:*

Sharing with [name of external entity] will take place subject to the terms and conditions specified in the attached **data sharing agreement**.

*In all cases, add:*

ONLY individuals who have completed and submitted UMRI’s researcher compliance form will be given access to the data.

*You should be very clear about who will have access to your copy of the data. For instance, if you plan to save this copy on a computer cluster at the University, the data would also be accessible to the system admins.*

*If the secondary data includes questionnaires, make sure you specify where they will be stored and for how long.*

Any data received from the UM repository for this project, as well as derivatives that are not in aggregate form, will be erased from computers/personal storage systems once the project has been completed – unless the data is needed for a follow-up study, in which case a new application will be submitted to UMRI, and the data will not be further processed or shared until authorization from UMRI and ethical clearance from F/REC [the relevant (Faculty) Research Committee] have been obtained.

**3.2 Technical measures**

Data from the repository will not contain the names and/or national ID card numbers of the data subjects, and will have been scrubbed of all personal attributes except for the data subject’s age/sex/height/weight/scan date [choose as appropriate], since this/these is/are needed [specify why].

Raw data (and any derivatives not in aggregate form) will be:

1. stored on an encrypted data storage device;
2. processed, analysed and viewed on computers which are password-protected, have appropriate antivirus software installed *(if applicable)*, and are never connected to an unsecured network.
3. **Ethical Aspects**

This project shall be subject to UM’s Research Code of Practice.

*If using primary data:*

Data collection will not start until the study has received ethical clearance from F/REC.

1. **Dissemination of Data and Findings**

**5.1 Norms**

Dissemination of results will conform to the norms of the discipline.

**5.2 Academic dissemination**

The primary mode of dissemination will be through published articles in reputable scientific journals. No personally identifiable data will be divulged in such articles.

*If using primary data:*

A comprehensive description of possible uses of the data is provided in 3.1.

**5.3 Public dissemination**

There may be cases where aggregate results or pseudonymised images are included in outreach material intended for science communication to the lay public. **Any findings will always be presented in a manner which makes it impossible to identify individual volunteers.**

1. **Other Issues**

This document will be reviewed on an as-needed basis and every version will be dated and stored for transparency and record-keeping. For the avoidance of doubt, ‘as-needed basis’ will be understood to mean ‘as needed to ensure the strictest safeguards for data privacy’. UMRI will be notified of any substantial changes.