



## **Psychosis MRI Shared Data Resource (Psy-ShareD) Partnership Data Transfer and Storage Protocol.**

### ***Relevant Documents/Templates***

1. *Psy-ShareD Memorandum of Understanding*
2. *Psy-ShareD Data Transfer Template*
3. *Information Commissioner's Office - Introduction to anonymisation*
4. *Information Sheet for New Partners*

### ***Data Sharing Agreement Documents***

Psy-ShareD Terms and Conditions Appendix (to be completed by external institution)

Psy-ShareD Contributor Letter

### **1. Introduction**

*1.1 Overview:* The Psy-ShareD Partnership will combine several pre-existing MRI datasets, with linked clinical and cognitive data, into one free-to-access, i.e., no cost database resource. The Partnership will bring together pre-existing structural MRI scans in patients with schizophrenia, first episode psychosis, clinical high risk for psychosis, familial high risk for psychosis, high levels of schizotypy, bipolar disorder, adults with an experience of childhood trauma, and age and sex matched healthy controls. This document outlines the procedures for data transfer from external sites to King's College London.

*1.2 Objectives:* The main objectives of the Psy-ShareD Partnership are as follows:

- i. Seek out and integrate within the Psy-ShareD database MRI datasets acquired in diverse psychosis and psychosis risk cohorts.
- ii. Build a sustainable 'free-to-access', well-documented and well-supported structural MRI data repository, utilising pre-existing MRI scans across psychosis illness/risk stages and in control participants, and include within the database linked clinical, cognitive, and demographic data.

- iii. Manage and promote the Psy-ShareD database to ensure wide global uptake by researchers across all career stages.
- iv. Develop and employ state-of-the-art methods to harmonise and standardise structural MRI, clinical and cognitive datasets.
- v. Once established, use Psy-ShareD to conduct and publish a proof-of-principle study addressing a current knowledge gap in the schizophrenia/psychosis MRI literature.

### *1.3 Data types (anonymised and de-identified prior to transfer)*

- i. T1 weighted (structural) MRI scan. The settings and acquisition parameters for T1 will depend on the make and field strength of each site's MRI scanner. Examples of typical T1 weighted acquisition sequence are SPGR, IR-SPGR MPRAGE, ADNI and ADNI-GO.
- ii. Demographic data will (ideally) comprise of age (at time of MRI scan), sex, ethnicity and handedness. Details on substance, tobacco, and alcohol use may also be included where available.
- iii. Clinical data will comprise of diagnosis, symptom assessment measures (i.e., PANSS or SAPS/SANS for psychosis populations and CAARMS/PSQ (or similar) for psychosis risk populations), the CTQ/ACE will be used to quantify childhood trauma. Other clinical data may comprise measures of depression and anxiety (e.g., HAM-D and HAM-A).
- iv. Cognitive measures will be quantified into domains of intellectual function, working memory/executive function, cognitive processing speed, and theory of mind (where data is available).
- v. Data anonymisation will be in accordance with the Commissioner's Code of Practice for Anonymisation ([ico.org.uk](http://ico.org.uk)) and appended.

### *1.4 Timeline*

It is anticipated that all data transfer will commence in August 2023 and continue for the duration of the set-up phase of the Psy-ShareD Partnership (as needed).

## **2. Data Sharing Schedule and Memorandum of Understanding**

Founding members of the Psy-ShareD Partnership are provided with Psy-ShareD Data Sharing Schedule (hereafter DSS) and Memorandum of Understanding (MoU) documents (both appended here). The DSS has been prepared by the Contracts Office at King's College London (KCL) and sets out the terms of data sharing between the Psy-ShareD Partners. Before data transfer can commence the DSS must be

completed by each site partner/investigator. Part 2 of the DSS requires details and information about the data being transferred to and received by KCL. The Table in Part 2 of the DSS must be fully completed and returned to KCL before data transfer can commence. The MoU sets out terms for data access by future Psy-ShareD users as well as terms of authorship. Please note that KCL may request future amendments to the DSS in relation to data access for Psy-ShareD users. Potential new Psy-ShareD partners are provided with a Psy-ShareD Schedule Terms and Conditions document for which Appendix A (data table) must be completed. A Psy-ShareD Contributor Letter that must be signed by a representative of the contributing institution and returned to KCL. A new data sharing agreement (DSA) can then be initiated.

### **3. External Site Data Preparation**

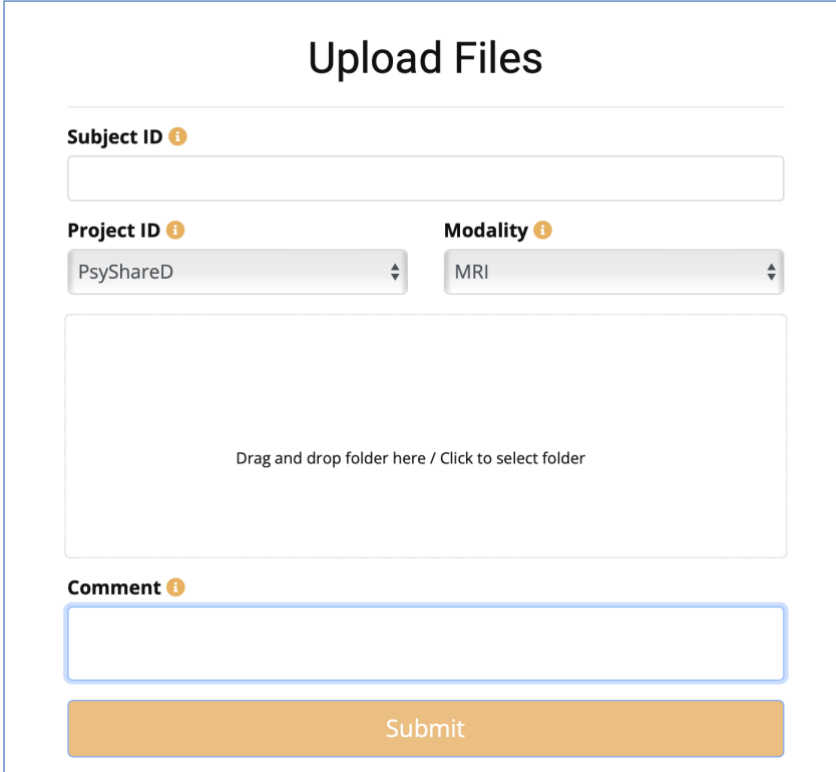
Before data transfer (from an external site) to KCL, the following procedures must be undertaken.

- i. The DSS/DSA must be completed (Part 2 Data Details for DSS or Appendix A data table) and the document signed by the local site PI/partner. Once the DSS/DSA has been completed this must be returned to the Psy-ShareD team at KCL. It is crucial that details of local ethics approvals are provided in section 3.2 of Part 2 of the DSS or section 4 of Appendix A for the DSA (new partners).
- ii. The transferring site must provide precise numbers and detail of the data being transferred in both the DSS/DSA and in the Psy-ShareD Data Transfer Template. These details should be sent to the Psy-ShareD team at KCL before data transfer begins.
- iii. Prior to data transfer, the Psy-ShareD team at KCL will provide the transferring site PI with computer scripts and instructions for data de-identification and full anonymisation according to the Psy-ShareD protocol. An additional script will also be provided that produces specific site, protocol, and participant codes for each MRI scan/file. Scripts will work in MATLAB (<https://www.mathworks.com/products/matlab.html>) and Python (<https://www.python.org>). Detailed information about scripts is provided in Appendix II and by the Psy-ShareD team liaison.
- iv. The Psy-ShareD team at KCL will also provide instructions and support for the implementation of the coding system to anonymously links MRI data with demographic, clinical, and cognitive data.

The Psy-ShareD KCL team are happy to help if any clarification is needed for the above process.

#### 4. Data Transfer

All data will be transferred via Bitbox (<https://www.bitbox-imaging.com>). Bitbox was developed at KCL for the purpose of imaging and clinical data transfer within multi-site studies. BitBox is used to transfer medical imaging data from external and independent sites into the KCL imaging network, securely and in accordance with GDPR regulations. All external sites must be registered in the system prior to transferring via Bitbox. Following registration, the Bitbox team will assign each site a unique Center ID. The sender associated with the specific center must use that ID to create his own account. Once the account has been created, the sender can send data using the GUI shown in Figure 1 below. The Psy-ShareD and Bitbox team at KCL will provide support for this process. Any comments and details about the data transferred can be inserted in the 'comment' tab. Please note that the maximum size for any single data transfer via Bitbox is 2GB of data organised in the folder. This should be sufficient for reasonably large MRI T1 samples. However, there may be occasions when MRI T1 data will need to be arranged into multiple 2GB folders, if sample sizes are very large.



The screenshot shows a web interface titled "Upload Files". It features a form with the following elements:

- A text input field for "Subject ID" with an information icon.
- Two dropdown menus for "Project ID" (set to "PsyShareD") and "Modality" (set to "MRI"), both with information icons.
- A large rectangular area with a dashed border and the text "Drag and drop folder here / Click to select folder".
- A text input field for "Comment" with an information icon.
- An orange "Submit" button at the bottom.

Figure 1: BitBox GUI interface for data transfer. Prepared folders containing anonymised de-identified data can be dragged and dropped into the box above.

## 6. Data Storage

All data (MRI, demographic, clinical, cognitive), will be stored on secure NaN servers at KCL Centre of Neuroimaging Science. This is a secure network that can only be accessed via a username and password. All data will be stored in the PSY\_SHARED project folder and only the KCL Psy-ShareD team will have permissions to access this folder. If you require any further detail regarding data storage at KCL please get in touch with the [Psy-ShareD KCL team](#). Figure 2 below shows a flow chart of the data transfer process.

Figure 2: Flow chart showing data transfer approval, preparation, transfer, and storage stages.

