A red sign with white text

Description automatically generated

**Psy**chosis MRI **Share**d **D**ata Resource (Psy-ShareD)

Information sheet for new partners

A key line of research in schizophrenia and psychosis populations has been to investigate neuroanatomical markers, relative to non-schizophrenia control groups. In recent decades, numerous Magnetic Resonance Imaging (MRI) studies investigating brain structure in schizophrenia and psychosis populations have been funded in the UK and internationally, progressing our understanding of the neuroanatomical basis of schizophrenia. However, many of these studies are underpowered, due to small sample sizes resulting from the costs of MRI, difficulties recruiting patients with psychotic illnesses, and a failure to consider differences across illness stages. Few efforts have been made to combine pre-existing datasets; and while notable recent multi- centre initiatives have begun, these datasets are generally restricted to MRI scans in cohorts at later stages, have patient sample sizes less than 300, and importantly, are not representative of the diversity of patients who experience psychosis, despite indications that patients from different ethnicities may show distinct patterns of neural markers.

The UK Medical Research Council funded Psy-ShareD database will host pre-existing structural MRI data collected at different sites across Europe, North America, South America, Asia and Australia (currently N ~3,000; projected N ~5,000) that will be suitable for region of interest, voxel-based morphometry, cortical thickness and surface area analytical approaches. We will also seek out and integrate further datasets within the Psy-ShareD database. All MRI T1 data included in Psy-ShareD will be linked to standardised clinical, demographic, IQ/cognitive data in patient, risk, and control populations. MRI datasets will be harmonised using software that can remove unwanted variation induced by scanner differences, while preserving biological variability between individuals using an empirical Bayes framework (i.e., [ComBat](https://github.com/Jfortin1/ComBatHarmonization)). Datasets in patient and risk cohorts have linked PANSS or SAPS/SANS ratings, which will allow classification of symptoms. Harmonised cognitive data for specific cognitive domains, such as attention, working memory, executive function, and social cognition, will also be available. Further information is available via the project website ([psyshared.com](http://psyshared.com/)).

The Psy-ShareD database will be hosted using open access platforms and data from all sites will be curated into BIDS format (Brain Imaging Data Structure, https://bids.neuroimaging.io) an international standard for organising neuroimaging data. Data dictionaries will be developed for clinical data ensuring the information is readily transferable between datasets. Curated, anonymised, and de-identified data will be allocated a DOI number ensuring a persistent and durable link to the files.

It is planned that the Psy-ShareD Partnership will:

1. Seek out and integrate within the Psy-ShareD database additional MRI datasets acquired in diverse psychosis and psychosis risk cohorts.

2. 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.

3. Manage and promote the Psy-ShareD database to ensure wide global uptake by researchers across all career stages.

4. Develop and employ state-of-the-art methods to harmonise and standardise structural MRI, clinical and cognitive datasets.

Benefits for Data Contributors: Firstly, *Data Conributors (Partners)* will be asked to approval all *Data User* requests to access their data (i.e. review and approval Data Access Request forms that seek to use their dataset(s)). No access to *Data Users* will be granted without site PI approval. To ensure that all PIs and sites contributing data to Psy-ShareD benefit and have their contribution recognised, we stipulate that all outputs (papers/articles) that make use of their data will need to include them as named co-authors. This is a condition which we impose on all researchers who make data access requests. We encourage, where possible, strong collaborative relationships between *Data Users* and Data Contributors through the data analysis and manuscript drafting phase. The *Psy-ShareD Partnership* will ensure that *Data Contributor/ Partners* contributing data to that paper have received proper authorship recognition, and those PIs will also be provided with the draft paper for their checking and approval at that stage as well, prior to any submission of the work.

In addition, *the Psy-ShareD Partnership* will be listed on the author line of all Psy-ShareD publications. As such, all publications and citations resulting from Psy-ShareD data will be linked to all data contributors via the *Psy-ShareD* author listing DOI. Full details related to authorship and publications are included in the Psy-ShareD [Memorandum of Understanding](http://psyshared.com/).

Data Transfer: Full details are provided in the attached Data Transfer Protocol document. Briefly, data are anonymized, and no personal data is transferred. Data are used only for the Project in the Principal Investigator’s laboratory. Upon receipt, the data will be given a protocol code by the project team which will ensure that there is no chain of identification, and the Data Provider and Data Recipient can no longer link the data back to an individual. Prior to the transfer of the data, the Data Provider must ensure it will remove precise features of individual neuroanatomy from the brain scans (the project team can facilitate this process). Upon receipt of the data, the Data Recipient will create harmonized and normalized data which will not constitute Confidential Information and will be made available to the research community, via an approval process whereby the Psy-ShareD team will review data access requests on a case-by-case basis.