Plan Overview

A Data Management Plan created using DMPonline

Title: Neuronal circuits for perceptual inference

Creator: Christopher Petkov

Principal Investigator: Christopher Petkov

Data Manager: Fabien Balezeau

Project Administrator: Yukiko Kikuchi

Affiliation: Other

Funder: Medical Research Council (MRC)

Template: MRC Template

ORCID iD: 0000-0002-4932-7907

Project abstract:

The overarching goal of this research is to causally perturb and advance a predictive model of a fundamental neuronal circuit for perceptual inference in a primate model and to scale insights on this circuit from microscopic levels, only feasible in the animal model, to mesoscopic levels which will inform human and human patient studies.

ID: 91556

Start date: 01-01-2023

End date: 31-12-2026

Last modified: 20-01-2022

Grant number / URL: NU-010029

Copyright information:

Neuronal circuits for perceptual inference

---

0. Proposal name

0. Enter the proposal name

Neuronal circuits for perceptual inference

1. Description of Data.

1.1 Type of Study

Fundamental science research in nonhuman primate (macaque) model at Newcastle University.

1.2 Types of Data

All data are quantitative generated from neuronal recordings and neuroimaging research with behavioural measures: 1) Extracellular neurophysiological recordings form neurons and local field potentials using high density laminar arrays, with and without optogenetic stimulation of pyramidal projection neuron modulation. 2) Functional neuroimaging data. 3) Behavioural data during these measures in the form of eye tracking information (eye looking location, saccade measures, pupil dilation) and behavioural responses of the animals during the tasks.

1.3 Format and scale of the data

All data are electronic unencrypted files. Neuroimaging and behavioural data are in the form of data files that are Matlab/Python readable. All neurophysiological recording data are converted to Matlab/Python readable files using the Neurodata Without Borders (.nwb) standard. All neuroimaging raw data are converted to NIfTI format with header file information and .json standard descriptors for analysis using freely available neuroimaging software (SPM/FSL). We will also collect other metadata (animal colony behaviour and genome fingerprint).

2. Data collection / generation

2.1 Methodologies for data collection / generation

Neuroimaging data: will consist of high resolution anatomical scans (T1, T2), fMRI time series and header file information. The Brain Imaging Data Structure (BIDS) data format is used where we will share these data in the specified structure (.json files for header and scanner parameter information and .nii files following the NIfTI neuroimaging format). MRI event codes and experimental and stimulus event timing information will also be shared via Matlab or Python readable files, with specification of the analyses software and parameters used. ***Neurophysiological array and optogenetics data:*** will consist of TDT or Neuralynx converted data in the Neurodata Without Borders (.nwb) format. Event codes and experimental and stimulus timing information will be shared within the data structure along with optogenetic parameters. ***Behavioural data:*** we incorporate eye tracking (x, y location and pupil dilation) data alongside the task and even timing information in the MRI or neurophysiology data as metadata. ***Metadata:*** we will collect animal behavioural/phenotyping information via Matlab/Python and DeepLabCut processed videos in Matlab/Python readable files. Genome sequencing data will be mapped to the rhesus genome (Mmu Mmul_10/rheMac10). ***Code and analysis scripts:*** will be managed via GitHub, with documented code and instruction files on running the code.

2.2 Data quality and standards

Neuroimaging data: will follow community defined standards (Autio et al., NeuroImage 2020) and quality control measures (Milham et al., 2018, 2020, 2021). The Brain Imaging Data Structure (BIDS) will be used, which includes .json files for header and scanner parameter information. ***Neurophysiological array and optogenetics data:*** will be assessed for quality using internal checks on signal to noise and quality control measures (denoising) following quality control measures determined by datasets shared in the Neurodata Without Borders (.nwb) format and OpenNeuro communities. The optogenetics data will be assessed for quality in relation to the primate optogenetics resource (Tremblay et al., 2020). ***Behavioural data:*** will include all training data and steps and the animal's performance values and eye tracking data that were collected. ***Metadata:*** will be evaluated for behavioural phenotyping quality in relation to prior reports and validated measures wherever possible. Genome sequencing data quality will be assessed by genetics experts in the UK/US as described in relation to the mGAP data, the Macaque Genotype and Phenotype Resource. ***Code and analysis scripts:*** will follow coding practices and documentation https://google.github.io/styleguide/docguide/best_practices.html.

3. Data management, documentation and curation

3.1 Managing, storing and curating data

Data at Newcastle fileservers are secure, redundant and multi-site server located. Data are backed up daily onto secure servers in the host Institution. To aid collaboration an Office 365 Team space will be created. It is a collaborative space for storing, sharing and working on digital information. It will enable researchers from all sites to share and view files in a secure 1TB central repository though a web interface and reduce the need to replicate files across sites. Balezeau and Kikuchi will oversee curating the data and code throughout the research study and prepare for sharing.

3.2 Metadata standards and data documentation

Filenames will clearly and uniquely identify each dataset. The neurobiological and behavioural data will all be stored with a description of the experimental protocol and any events that might have influenced the data.

3.3 Data preservation strategy and standards

Raw data will be retained 10 years after the project ends on secure local servers at the host institution. Analysed data and metadata will be made available after publication on publically accessible repositories including Open Science Framework, OpenNeuro and PRIME-DE.

4. Data security and confidentiality of potentially disclosive personal information

4.1 Formal information/data security standards

This is not a human study which requires protection of the identity of the individuals. The animals are identified with a unique code that is consistent across the sharing sites for identification of the data in the animal. See Newcastle University Data Security Policy (link below) which is in line with recognised legal information/data security standards.

4.2 Main risks to data security

The main risk to data security are unauthorised cyber attack and loss. Data theft is protected against by the use of secure institutional servers and by using animal codes in the data based on the use of alphanumeric identifiers. Loss is protected against by the use of secure institutional servers with internal backup.

5. Data sharing and access

5.1 Suitability for sharing

The data are suitable for sharing alongside publication with third parties interested in carrying out further analysis. Sufficient metadata and the use of widely available software and file formats
allow this. All data will be shared as part of published reports or during the research as part of collaboration-seeking model (Milham et al., 2020, 2021) via these venues. ***Neuroimaging data:*** will be shared via the PRIMatE Data (https://fcon_1000.projects.nitrc.org/indi/indiPRIME.html) and Resource Exchanges (https://prime-re.github.io/) using the Brain Imaging Data Structure (BIDS) data format. ***Neurophysiological array and optogenetics data:*** will be shared on OpenNeuro (https://openneuro.org/) using the Neurodata Without Borders (.nwb) format. ***Behavioural data:*** will be shared as metadata alongside the neuroimaging or neurophysiological data. ***Metadata:*** The whole genome sequencing data will be shared with mGAP, the Macaque Genotype and Phenotype Resource, as well as the National Center for Biotechnology Information (NCBI) Sequence Read Archive (https://www.ncbi.nlm.nih.gov/sra/docs/). ***Code and analysis scripts:*** will be shared via GitHub (https://github.com/) and to ensure long-term access through data.ncl (Newcastle Data Repository) with supporting metadata.

5.2 Discovery by potential users of the research data

Potential new users can discover the repository and contact details for research team in the publications that result from this work or from searching for these type of data in the shared repositories listed above. The datasets will be assigned a persistent identifier in the form of a DOI (digital object identifier). The DOI will ensure the data is findable in its own right but will also be used in research outputs, including publication, in the form of data access statements to link the data and outputs. In addition, Newcastle's data.ncl (powered by figshare) will create a collection that is assigned a DOI. Within this collection will be Newcastle data but also data records that link to data archived elsewhere. This will allow a potential reuser to see all published datasets from the project in one place.

5.3 Governance of access

The data will be made freely available in public repositories under appropriate licenses (e.g. CC-BY or MIT).

5.4 The study team's exclusive use of the data

The study team will have exclusive use of the data until publication after which they will be made openly available. However, we are also part of an open data community that is developing collaboration seeking models (PRIME-DE) which means that data sharing can occur with third parties with some restriction until publication when the data are openly and freely shared.

5.5 Restrictions or delays to sharing, with planned actions to limit such restrictions

There will be none.

5.6 Regulation of responsibilities of users

Data will be licensed and de-identified before sharing with users. Users will be required to acknowledge the provenance of the data and MRC funding.

6. Responsibilities

6. Responsibilities

Study-wide data management is the responsibility of the PI. Data and metadata creation and curation for sharing is the responsibility of the identified staff, overseen by Balezeau & Kikuchi. Data security and quality assurance of data is the responsibility of the IT team.

7. Relevant policies

7. Relevant institutional, departmental or study policies on data sharing and data security

Data Management Policy and Procedures
https://www.ncl.ac.uk/media/wwwnclacuk/research/files/ResearchDataManagementPolicy.pdf
Data Security Policy
https://services.ncl.ac.uk/itservice/policies/InformationSecurityPolicy-v2_1.pdf
Data Sharing Policy
See data management policy
Institutional Information Policy
https://www.ncl.ac.uk/media/wwwnclacuk/research/files/openaccess-to-research-outputs.pdf

8. Author and contact details

8. Author of this Data Management Plan (Name) and, if different to that of the Principal Investigator, their telephone & email contact details

Christopher Petkov (chris.petkov@ncl.ac.uk)
Fabien Balezeau (fabien.balezeau@ncl.ac.uk)
Yukiko Kikuchi (yukiko.kikuchi@ncl.ac.uk)
Colline Poirier (colline.poirier@ncl.ac.uk)