# Plan Overview

*A Data Management Plan created using DMPonline*

**Title:** BrainTrack

**Creator:**Eliana Maria Vasquez Osorio

**Principal Investigator:** Mauricio Alvarez, Angela Davey

**Affiliation:** University of Manchester

**Funder:** Engineering and Physical Sciences Research Council (EPSRC)

**Template:** EPSRC Data Management Plan Customised By: University of Manchester

**ID:** 176587

**Start date:** 01-03-2026

**Last modified:** 28-01-2026

# BrainTrack

## Manchester Data Management Outline

### 1. Will this project be reviewed by any of the following bodies (please select all that apply)?

- Ethics
- Funder

### 2. Is The University of Manchester collaborating with other institutions on this project?

- Yes - Part of a collaboration and owning or handling data

Collaborators:

- The Christie NHS Foundation Trust
- Universidad de Antioquia (UdeA)
- Comfamiliar Risaralda *(via collaboration with UdeA)*
- Hospital Pablo Tobón Uribe *(via collaboration with UdeA)*

### 3. What data will you use in this project (please select all that apply)?

- Re-use existing data (please list below)

Data will include:

- Clinical records of treatment including demographics, treatment characteristics, and treatment outcomes.
- Radiotherapy planning data including medical images, contours and dose distributions.
- Follow-up magnetic resonance imaging (MRI) for children with cancer, and children with perinatal asphyxia.
- Longitudinal MRI for healthy children

### 4. Where will the data be stored and backed-up during the project lifetime?

- Other storage system (please list below)
- University of Manchester Research Data Storage Service (Isilon)

Isilon resources will be used to store open-access or requested-access data of healthy children.

Local storage servers at the Advanced Radiotherapy group within the Radiotherapy Related Research team for UK-Computer Aided Theragnostics (ukCAT) data (from The Christie). These servers are

regularly backed up on twin machines located in different locations within The Christie NHS Foundation Trust and University of Manchester.

Data from Colombian collaborators will be stored and handled by UdeA.

Derived data from all data sets will be stored on a shared folder using One Drive.

**5. If you will be using Research Data Storage, how much storage will you require?**

- 1 - 8 TB

**6. Are you going to be receiving data from, or sharing data with an external third party?**

- Yes

Raw data (images) will not be shared between collaborators. However, we will share derived data (including automatic segmentations of brain sub-structures) between University of Manchester (UoM) and UdeA.

**7. How long do you intend to keep your data for after the end of your project (in years)?**

- 5 - 10 years

***Guidance for questions 8 to 13***

**Highly restricted information defined in the Information security classification, ownership and secure information handling SOP is information that requires enhanced security as unauthorised disclosure could cause significant harm to individuals or to the University and its ambitions in respect of its purpose, vision and values. This could be: information that is subject to export controls; valuable intellectual property; security sensitive material or research in key industrial fields at particular risk of being targeted by foreign states. See more examples of highly restricted information .**

**If you are using 'Very Sensitive' information as defined by the Information Security Classification, Ownerships and Secure Information Handling SOP, please consult the Information Governance Office for guidance.**

**Personal information, also known as personal data, relates to identifiable living individuals. Personal data is classed as special category personal data if it includes any of the following types of information about an identifiable living individual: racial or ethnic origin; political opinions; religious or similar philosophical beliefs; trade union membership; genetic data; biometric data; health data; sexual life; sexual orientation.**

**Please note that in line with data protection law (the UK General Data Protection Regulation and Data Protection Act 2018), personal information should only be stored in an identifiable form for as long as is necessary for the project; it should be pseudonymised (partially de-identified) and/or anonymised (completely de—identified) as soon as practically possible. You must obtain the appropriate ethical approval in order to use**

**identifiable personal data.**

**8. What type of information will you be processing (please select all that apply)?**

- Anonymised personal data

**9. How do you plan to store, protect and ensure confidentiality of any highly restricted data or personal data (please select all that apply)?**

- Not applicable

**10. If you are storing personal information (including contact details) will you need to keep it beyond the end of the project?**

- Not applicable

**11. Will the participants' information (personal and/or sensitive) be shared with or accessed by anyone outside of the University of Manchester?**

- No

**12. If you will be sharing personal information outside of the University of Manchester will the individual or organisation you are sharing with be outside the EEA?**

- No

**13. Are you planning to use the personal information for future purposes such as research?**

- No

**14. Will this project use innovative technologies to collect or process data?**

- No

For clarification, artificial intelligence (AI) and machine learning will be used on anonymised data.

**15. Who will act as the data custodian for this study, and so be responsible for the information involved?**

Dr Eliana Vasquez Osorio

**16. Please provide the date on which this plan was last reviewed (dd/mm/yyyy).**

2025-07-03

## Data Collection

**What data will you collect or create?**

Cancer cohort (data access approval in place): De-identified data (imaging, hospital records and treatment data) for all patients treated at The Christie NHS Foundation Trust, UK, are available via the UK Computer Aided Theragnostics (ukCAT) research database. We have all imaging and patient data for 36 patients with mixed-diagnosis brain tumours, aged 3 - 25 years at the time of treatment. We have approval to extend this cohort during this project as part of Brainatomy 2 grant.

Perinatal asphyxia cohort (data access approval in place): Anonymised data from 34 children, including imaging, parent's demographics, Bayley assessments, and other relevant clinical data. This dataset includes 12 children of which were treated with therapeutic hypothermia at birth (born and treated at Hospital Pablo Tobón Uribe) and the rest not treated (born at Comfamiliar Risaralda).

Cohorts of healthy children (data access approval in place/in progress): We will use open datasets of imaging and demographics for healthy children to model normative growth. These include: i) Lifespan Baby Connectome Project, ii) OpenNeuro, iii) the Adolescent Brain Cognitive Development study (ABCD, release 5.1, https://abcdstudy.org/). Datasets i) and iii) are handled by NIMH Data Archive (NDA), while dataset ii) is completely public.

**How will the data be collected or created?**

Cancer cohort: Already been collected from anonymised records as part of the UK Computer Aided Theragnostics (ukCAT) research database. Ethics approval has been obtained for the retrospective analysis from the UK Computer Aided Theragnostics Research Database Management Committee (research ethics committee reference no. 21/NW/0347). Under this agreement, a database management committee handles data access applications by UoM researchers. Data is already collected, but if extending during the lifetime of the project, new data will be collected from

Perinatal asphyxia cohort:  This dataset will be accessed via our key collaborators Dr Hernan Garcia Arias, via the Universidad de Antioquia. Note that this study is still recruiting, so the numbers may increase. All ethical permissions and data-sharing agreements are in place and are handled by our Colombian collaborators.

Cohorts of healthy children: The team has gained access in the past to NDA-managed datasets, including ABCD, and will renew the relevant agreements for this grant.

## Documentation and Metadata

### What documentation and metadata will accompany the data?

Documentation and metadata will be used to help secondary users understand and re-use the data, and will include information generated from the patient processing for analysis, definitions of variables, quality control mechanisms, and methodologies.

For medical images, metadata is part of the DICOM format. Data exported as spreadsheets will have their data dictionary stored in an additional tab following the guidance of the Open Science Framework (https://help.osf.io/article/217-how-to-make-a-data-dictionary) .

Delivered data will be stored in folders clearly indicating if it is intermediate data or final results within each project folder. Every project folder will have a "readme.txt" file following the guidance described in the Cornell's Research Data Management Service Group (https://data.research.cornell.edu/data-management/sharing/readme/). As a minimum, the readme files will include a description of the analysis run, the person responsible for the analysis and a time period (following W3C/ISO 8601 date standard YYYY-MM-DDThh:mm:ss). To keep track of all inputs, identifiers, the version of the source code, and parameters/configuration files used will be recorded (automatically generated from scripts). To ensure all outcomes are recorded, results, artifacts, and logs generated from the experiment will be kept, as well as metrics, charts, or visualizations used to interpret the outputs. These is in accordance to standard practices to document data-science experiments

## Ethics and Legal Compliance

### How will you manage any ethical issues?

Ethical issues in terms of use of data will be in line with ethically approved research database (ukCAT).

Ethical issues in terms of release of research data will be dealt with via stipulations included in research collaboration agreements between the university and the data providers (if any agreements are relevant) and accordingly with University of Manchesters data sharing processes.

### How will you manage copyright and Intellectual Property Rights (IPR) issues?

Intellectual property rights will be dealt with and discussed during the formation of data sharing research collaboration agreements (if appropriate). Contracts will be put in place via The University of Manchester contracts office, and will include details of publication policy for the data.

## Storage and Backup

### How will the data be stored and backed up during the research?

UoM Isilon resources will be used to store open-access or requested-access data of healthy children. Local storage servers at the Advanced Radiotherapy group within the Radiotherapy Related Research team for UK-Computer Aided Theragnostics (ukCAT) data (from The Christie). These servers are regularly backed up on twin machines located in different locations within The Christie NHS Foundation Trust and University of Manchester.

Data from Colombian collaborators will be stored and handled by UdeA.

Derived data from all data-sets will be stored on a shared folder using One Drive.

### How will you manage access and security?

Data on the UoM servers will be accessible only to members of the project team, with access only available through a University of Manchester log-in to those who have access granted. Derived data from all data-sets will be stored on a shared folder using One Drive, which is only accessible to members of the project team. Access management will be controlled by Dr Eliana Vasquez Osorio.

## Selection and Preservation

### Which data are of long-term value and should be retained, shared, and/or preserved?

Developed models of brain development will be shared for future use for benefit to researchers, and downstream to children with disease or developmental conditions.

### What is the long-term preservation plan for the dataset?

Relevant research outputs will be retained in our research servers and the University of Manchester Research Data Management Service (RDMS) in alignment with University and EPSRC guidelines on retention of research outputs (minimum of 10 years). All data stored in our research servers and at UoM RDMS will be replicated.

## Data Sharing

### How will you share the data?

Data generated during this project will be made discoverable by publications and presentations. Published data will be made available via UoM eScholar repository (which is the publishing face of the RDMS). Published outputs will be assigned a DOI (Digitial Object Identifier) which can be used to reference the data in publications.

Appropriate metadata will be published with the research data to enable other researchers to identify whether the data could be suitable for their own research.

Public release of data will occur in a timely manner, with the derived sharable results available no later than the acceptance for publication of the main findings from the final dataset. For volumetric data

(e.g., sub-structure models - WP4) file formats that conserve the geometrical information will be used (nifti or mhd files). For growth models, the complete model will be shared. We will also make scripts for model-building open access  (via GitHub), derived data (e.g., vectors from correspondence in WP2) will also be made available.These will be included in the supplementary materials alongside the published articles (for publishers that allow this), and/or in the University of Manchester escholar repository.

**Are any restrictions on data sharing required?**

Anonymised patient data access is governed by the ethical approval and data sharing agreements in place with the various institutes. These data are not generated by the project and as such will not be available for sharing.

## Responsibilities and Resources

**Who will be responsible for data management?**

Dr Eliana Vasquez Osorio

**What resources will you require to deliver your plan?**

University storage facilities (costed).
Departmental storage systems (in-kind) - already paid for by large grants.