# Plan Overview

A Data Management Plan created using DMPonline

Title: Healthworker Attitudes towards Tobacco Use and Cessation Among People with Severe Mental Illness

Creator:RUSHAM RANA

Affiliation: University of York

Template: York DMP Template for Postgraduate Researcher Projects

ID: 183023

Start date: 04-08-2025

End date: 30-11-2027

Last modified: 29-09-2025

Copyright information:

The above plan creator(s) have agreed that others may use as much of the text of this plan as they would like in their own plans, and customise it as necessary. You do not need to credit the creator(s) as the source of the language used, but using any of the plan's text does not imply that the creator(s) endorse, or have any relationship to, your project or proposal

# Healthworker Attitudes towards Tobacco Use and Cessation Among People with Severe Mental Illness

## 1. Defining your data

### 1a. What data will you create or use?

This project will generate both primary and secondary data across three components: a scoping review, an online survey, and qualitative interviews.

For the scoping review, bibliographic records and extracted information (such as author names, publication year, study setting, methodology, and key findings) will be collected from databases including PubMed, Medline, Embase, and PsycINFO. These records will be managed in Covidence. Extracted data will remain under 100 megabytes and will not include full texts, which will only be accessed through institutional subscriptions.

Primary data will be collected through a web-based survey of health workers in mental health settings in Pakistan using Google Forms. Survey data will include both quantitative and short qualitative responses, exported into CSV or SPSS files. The total volume is expected to remain below 50 megabytes.

Qualitative data will be generated through semi-structured interviews with health professionals. Interviews will be recorded using encrypted devices, transcribed, and anonymised before analysis. Transcripts and coding files will be stored in Word or NVivo formats. The combined volume of audio and transcripts is expected to remain under 2.5 gigabytes. Raw recordings will be deleted once transcription and verification are complete.

Additional materials will include survey instruments, interview guides, consent forms, ethics approvals, and memos. These will be stored in Word or PDF format and remain under 100 megabytes. Non-sensitive materials (such as anonymised questionnaires or coding frameworks) may be shared under a Creative Commons licence. Sensitive data, such as interview transcripts, will not be made openly available. Only anonymised excerpts or thematic summaries may be included in the thesis and publications.

### 1b. Who owns the data you will create or use?

As a registered PhD student at the University of York, I, Miss Rusham Zahra Rana will retain intellectual property rights over the data generated throughout the course of this project, in accordance with the University's Regulation 12 on student intellectual property. This includes all primary data collected through surveys and qualitative interviews, as well as any supporting materials such as interview guides, consent forms, and analytic frameworks developed during the project. Ownership of this data will remain with me unless otherwise specified in any future collaboration or funding agreements. At present, there are no external sponsors or collaborators with claims on intellectual property.

For the scoping review component, existing data in the form of published literature will be reused under fair research use and citation practices. The bibliographic records and extracted content from these sources will remain the intellectual property of the original publishers and authors. No full text articles will be stored or shared. Extracted data used in the review will be limited to information already publicly available and will be appropriately referenced. Use of these sources will comply with licence agreements available through the University’s institutional subscriptions.

Access to sensitive datasets such as interview transcripts will be managed carefully in accordance with ethical approvals and participant consent. Where possible, anonymised data and research tools will be shared openly, while ensuring compliance with data protection and copyright obligations. Should any new agreements arise during the course of the research, for example with collaborating institutions or transcription services, these will be reviewed for intellectual property implications and recorded accordingly.

### 1c. Are there any ethical or legal considerations relating to the data you will create or use?

This project involves collecting data from human participants and will therefore handle personal and sensitive information. Ethical approval will be obtained from the Health Sciences Research Governance Committee at the University of York and, where required, from local review boards in Pakistan.

Survey data will be collected anonymously, with no identifying information stored. Interview data will be anonymised during transcription, with names, workplaces, and any identifying details removed. Audio files will be deleted once transcription and verification are complete.

Participation will be voluntary, and informed consent will be sought. Participants will be told how their data will be used, stored, and shared. They will have the right to withdraw their data within two weeks of participation.

Anonymised survey data may be made available for reuse in line with participant consent. Qualitative data will not be shared in raw form, but thematic summaries and anonymised illustrative quotations may be included in publications. All data will be stored securely on University systems, classified under the Information Classification Policy, and disposed of when no longer required.

## 2. Working with your data

2a. How much data storage will you require during your project's lifetime? (tick the storage requirement you estimate for your project)

- 250GB - 1TB
- < 250GB

2b. Where will your data be stored for the duration of the project? (tick the storage option(s) you have chosen to use)

- University of York Google Drive
- University of York filestore (my personal filestore - the userfs drive - on the centrally managed network)

I will save personal data on the University of York filestore and the anonymised data on the University Google Drive.

2c. How will your data be backed up?

All sensitive data will be stored on a secure, managed University of York filestore.

No data will be stored solely on local devices such as laptops or external drives. Temporary working files created during travel or fieldwork will be transferred to secure University storage as soon as possible and deleted from portable devices immediately after confirmation of successful transfer. If internet access is limited during fieldwork, encrypted files will be saved temporarily on password-protected devices and synced with University systems at the earliest opportunity.

The responsibility for managing data backups and ensuring regular transfer to secure storage lies with me as the primary researcher. I will follow University of York guidance on data security and good research data management practices throughout the life of the project.

2d. How will you manage access to your data and its security?

Some of the data collected during this project will be classified as confidential, particularly audio recordings and transcripts of interviews with health workers, as well as consent forms and contact information. These data will be handled in line with the University of York's Information Security Policy and Information Classification guidance.

All confidential data will be stored on secure, access-controlled University of York systems and the Health Sciences departmental filestore, which are protected by institutional authentication and encryption. Access to identifiable or sensitive data will be restricted to myself and my academic supervisors. If any temporary storage is required during fieldwork, files will be saved in encrypted formats on password-protected devices and transferred to University storage at the earliest opportunity. Once transferred, all local copies will be deleted permanently.

No data will be stored on unencrypted USB drives, personal cloud storage, or local hard drives without protection. If required, guidance from IT Services will be sought to ensure appropriate encryption and secure handling of devices. Collaboration with supervisors will be managed using University-approved platforms, and any shared files will be accessible only through accounts with two-factor authentication.

Public or non-sensitive data such as anonymised survey instruments or codebooks may be shared more widely during the project. In these cases, care will be taken to ensure that files do not contain identifying information or metadata that could

compromise confidentiality. All data will be regularly reviewed to ensure compliance with University security standards throughout the project's duration.

2e. How will you organise your data?

All research data will be organised in a clear folder structure based on project components (scoping review, survey data, interviews, ethics documents, analysis outputs). Subfolders will separate raw, processed, and final outputs. File names will use underscores, descriptive elements, and dates in the format DDMMYYYY, with version numbers where relevant (e.g., transcript_participant07_15012026_v01.docx).

Versioning will be supported through file naming conventions and supplemented by the automatic version history available in University-approved platforms such as Google Drive or OneDrive, reducing reliance on manual tracking. README files will be prepared for core datasets intended for archiving or sharing, explaining file naming logic, coding schemes, and abbreviations. The system will be designed to remain simple and accessible for supervisory oversight.

2f. How will your data be documented and described?

Documentation will focus on datasets intended for archiving or sharing, including README files and data dictionaries. Coding frameworks and software versions will be noted for transparency. Additional notes will be created where needed to support interpretation, rather than systematically for every intermediate output.

## 3. Archiving your data

3a. Which data will be retained long-term?

At the end of the project, anonymised survey data and supporting documentation (such as variable definitions and a data dictionary) may be archived in ReShare. Consent will be sought from participants for this purpose, and access may be safeguarded rather than open if any risk of indirect identification is identified.

Raw qualitative interview data (audio and transcripts) will be deleted after transcription, verification, and analysis. Only anonymised thematic summaries or extracts included in the thesis or publications will be retained. No raw qualitative files will be archived or shared.

Administrative records (consent forms, ethics approvals) will be stored securely for ten years in accordance with University policy, accessible only to the supervisory team and relevant governance bodies.

3b. Where will you deposit the data? (tick the data repository/archive you have chosen or are required to use)

- Other repository/archival system (please provide details below)
- A subject or discipline-specific data repository (please provide details below)

If any survey data is deemed sufficiently anonymised and suitable for sharing, I intend to deposit it through the UK Data Service's ReShare repository. This platform allows for appropriate metadata publication, flexible access controls (e.g. open or safeguarded access), and is aligned with good data governance practices.

All decisions regarding deposit will be reviewed with my supervisors, and where required, deposits will be made under the supervision or support of a University staff member to ensure compliance with institutional data protection requirements.

## 4. Sharing your data

4a. Will the data you archive be openly shared?

- No (please explain why the data cannot be openly shared below)

Survey data may be archived in ReShare, but sharing will only occur if participants consent and supervisory review confirms that the risk of reidentification is low. If any risk remains, the data will be deposited under safeguarded access rather than open access.
Qualitative data will not be shared in raw form. Only anonymised thematic summaries or selected extracts included in published outputs may be retained, and these will not be deposited as standalone datasets.

4b. What access restrictions will be imposed on the data you archive?

Access to archived data will be based on sensitivity and participant consent. Anonymised survey data may be made available through ReShare, but where there is any risk of indirect identification, it will be subject to safeguarded access.
Qualitative interview data will not be archived in raw or transcript form due to the risk of deductive disclosure. If any extracts or thematic data are retained beyond inclusion in publications, these will be available only through managed-access arrangements requiring justification of use and agreement to confidentiality safeguards.
Supporting documents such as survey instruments and codebooks may be shared openly where they carry no risk of identification.

## 5. Implementing your plan

5a. Who is responsible for implementing your plan?

The principle investigator, Rusham Zahra Rana.

5b. How will your plan be kept up-to-date?

As and when necessary, but at a minimum before every Thesis Advisory Panel meeting (every six months).

5c. What training or further information will you need to successfully manage your research data throughout the project

Training on how to write README files and how to store metadata will be needed. Training on how to manage data has already been undertaken via an online course offered by the University of York.