



RESPIRE Data Management Plan (DMP): Template (adapted from the University of Edinburgh)

Name:	Dr.Tabish Hazir
Project Title:	Community perception on effective approaches to promote acceptance and adherence to public health measures for COVID-19 prevention and control
Institute:	Maternal, Neonatal and Child Health Research Network (MNCHRN)
Start Date:	15 th June 2020
End Date:	30 th November 2020
DMP version number and date:	DMP_V01_20201130

Responsibilities & Resources (applicable across the sections below)

Who will be involved in the data management of this research?

Name	Role in the Data Management
Dr Hana Mahmood	Data deposit and oversight of implementation of the plan
Syed Yahya Sheraz	Data storage
Hira Kiani	Data collection and management
Rakhshanda Hameed	Data analysis and reporting
Sajid Khan	Data collection
Saman Mujeeb	Maintenance of data

1. Data Capture

What data will be generated or reused in this research?

Interviews are being done from a nationally representative sample of respondents across Pakistan through Computer Assisted Telephonic Interviewing (CATI) panel. A total of 2651 respondent will be interviewed, out of which 1890 interviews have been conducted so far. Numbers are automatically dialled from an already available panel of respondents who have already agreed to take part in public opinion based surveys. The system is using the principle of random digital dialling (RDD) to ensure random sampling. The data is being collected telephonically through tablets/computers with incorporated questionnaires with logical checks. Monitoring of the data collection process is being conducted through an online dashboard by the project manager. Data entered in the computer system by each enumerator is sent to the central server which can be accessed remotely and extracted by the data manager for analysis and reporting. The data will be analyzed in SPSS. This data (raw and anonymised) in the form of excel sheets and SPSS data files along with preliminary reports will be stored in password protected folders and computers with access to authorised members of

the research team only. These computers have also been encrypted using bitlocker and folders using the 7-zip.

How much data will be generated?

Up to 20 GB of data will be generated.

2. Data Management

How will the data be documented to ensure it can be understood?

For the questionnaire, data in the form of excel sheets, SPSS data files and preliminary reports, following will be recorded in a text file:

- name of the file
- file location
- file size
- file format
- software used or required to create or run the file (e.g. Microsoft Word for transcripts and translated documents)
- date of creation
- file creator
- file version
- access rights set for the file

Where will the data be stored and backed-up?

Storage:

All data (raw and anonymised) in the form of excel sheets and SPSS data files along with preliminary reports will be stored in password protected encrypted folders and computers with access to authorised members of the research team only. These computers have been encrypted using bitlocker and folders using the 7-zip.

Backup:

3 backup copies of the original data have been created and stored on different locations. Original copy has been stored on the password protected computer of the data manager located in the MNCHRN headquarters. One copy has been backed up fortnightly on the main server located at the central field office of MNCHRN and two copies on two separate drives with our cloud storage (My Cloud EX4100 with enterprise-class WD Gold Hard Drives- 2 drives of 10TB each). These hard drives provide up to 2.5M hours MTBF, vibration protection technology, RoHS compliant, have SATA 6 Gb/s interface, 262 MB/s data transfer rate, 256 MB cache, 7200 RPM performance class.

File naming:

Files have been named as per the standard naming convention of the organisation i.e. MNCHRN name_projectID_type of data file e.g. MCNHRN_RES007_excelsheet.txt

Organization Of folders:

Week-wise folders have been created for the data gathered during that week. These folders are password protected with access to only authorized members of research team. Files have been placed in their respective folders.

Data transfer:

All data collected or generated by the study (including personal data) will not be transferred outside the research site. Only anonymised, non-identifiable data will be shared at University of Edinburgh data repository (DataShare).

3. Integrity

How will you quality assure your data?

Following measures have been taken to assure the quality of the data:

- All versions of data collection tools will be saved for traceability.
- Data collection through tablets/computers with incorporated questionnaires with automatic logical skips and data coding checks to provide error free or complete data entry.
- The system automatically dials numbers of available panel of respondents who have already agreed to take part in public opinion based surveys using random digit dialing (RDD) principle to ensure random sampling. None of the enumerators have access to the phone numbers of the respondents.
- The manager can access the dashboard and observe the performance of each enumerator by noting the number of calls made per day, and the disposition status given by the enumerator against each call including the number of successful interviews, number of incomplete interviews, number of refusals, not responding numbers etc.
- Remote meetings are being conducted with research team through teleconferencing to give feedback on their performance and suggest ways to improve it.
- Data stored in three different locations to prevent the loss of the data. Original copy has been stored on the password protected and encrypted computer of the data manager placed at the corporate office of MNCHRN at the headquarters. This has been backed up on the main server located at the central field office of MNCHRN and also on cloud storage on fortnightly basis.

4. Confidentiality

How will you manage any ethical and Intellectual Property Rights issues?

De-identifying the data of all respondents using pseudonyms where required. Additionally, consents have been taken from the participants before the start of the interview indicating their voluntary participation and permission to share the data as per the participant information sheet.

5. Retention and Preservation

Which data do you plan to keep and for how long?

All data generated from this study will be kept for a minimum of 5 years after the study is finished. The tool used for data collection i.e. the finalized questionnaire along with its CATI version and code book will be archived.

How will the data be preserved?

The anonymised and cleaned data of interviews, questionnaire and meta data will be preserved in stable file formats i.e. note pad for text files and comma-separated value (.csv) file for spreadsheets. These will be uploaded in University of Edinburgh's online digital repository –DataShare.

The University of Edinburgh undertakes to maintain the digital outputs of this project for the long-term and will utilise University infrastructure (namely the Edinburgh DataShare repository) to ensure preservation and continued access. Edinburgh DataShare is an online digital repository of multi-disciplinary research datasets produced at the University of Edinburgh, hosted by the Data Library in Information Services. It acts as a trusted repository, ensuring that research data will be preserved.

6. Sharing and Publication

Which data will be shared and how?

The anonymised data generated from this study will be shared and archived in DataShare which is an open access data repository where data will be held indefinitely. The repository is optimised to be indexed by google scholar and google dataset search to promote data discovery.

Are any restrictions on data sharing required?

No, there are no restrictions on sharing of the anonymised data shared and archived in DataShare.