



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

Name:	Dr. Tabish Hazir
Project Title:	Understanding the reasons for delay in seeking care for pneumonia among caregivers of children under five in Pakistan
Institute:	Maternal, Neonatal and Child Health Research Network (MNCHRN)
Start Date:	15 December 2018
End Date:	30 March 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 YahyaSheraz	Data storage
Hira Kiani	Data collection and management
Rakhshanda Hameed	Data analysis and reporting
Nyla Muzaffar	Data analysis
Sajid Khan	Data collection
SamanMujeeb	Maintenance of data

1. Data Capture

What data will be generated or reused in this research?

In depth interviews:

In depth interviews (IDIs) with mothers of children under the age of five residing in urban and rural communities across Punjab, Sindh, Khyber Pakhtun Khawa (KPK), Azad Jammu Kashmir (AJK) and the federal capital were conducted either face to face or telephonically, based upon the respondent convenience, by members of the research team using a semi structured pre tested interview guide. These were audio recorded with the permission of the respondent. Field notes were also recorded during the interview by a note taker. Recordings were transcribed and translated. Thereafter, they were coded by two members of the research team.

32 in depth interviews have been conducted in all.



Focus group discussions

Focus group discussions (FGDs) were conducted with fathers and grandmothers of children under five years residing in urban and rural communities of federal capital, Punjab, KPK, Sindh, and AJK. Each face to face activity had around 8 to 10 participants. These were audio recorded. Field notes were also taken by a note taker. Data including audio recordings, transcriptions, translations, field notes and pictures have been stored in password protected and encrypted computer with access to authorized members of the research team only. All the data is de-identified by replacing actual names with pseudonyms or using initials only. 20 focus group discussions including -10 with grandmothers and -10 with fathers have been conducted.

Computers holding this data have been encrypted using Bitlocker and the folders are stored in encrypted containers using the 7-zip.

How much data will be generated?

Up to 30 GB of data will be generated.

2. Data Management

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

For audio recording, transcript and translation of each in depth interview (IDI) and focus group discussion (FGD), 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:

Data including audio recordings, transcripts, translations, photographs and field notes are stored in password protected and encrypted computers with access to authorized members of the research team only. These computers have been encrypted using Bitlocker and the folders are stored in encrypted containers 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 and encrypted computer of the data manager placed at the corporate office of MNCHR at the headquarters. One copy has been backed up on the main server located

at the central field office of MNCHRN and two copies on cloud storage(My Cloud EX4100 secured by volume encryption with enterprise-class WD Gold Hard Drives- 2 drives of 10TB each) on fortnightly basis. 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 organization i.e. MNCHRNname_projectID_subjectID_type of data file e.g. MCNHRN_RES002_K.1.1_transcript.docx

Organization Of folders:

Folders according to the geographical locations of the sites of data collected have been created. Separate sub folders have been created for each interview and focus group discussion with each including the following

- Audio recording
- Field notes of the note taker
- Photos
- Transcript
- Translated document

Data transfer:

Data was transferred to the computer of the data manager on the same day after the interview or focus group discussion. The original audio file in the audio stick was deleted after copying it in the computer of data manager. The audio file was trimmed by the Data Manager to delete any identifying information before sharing for transcription. A copy is saved in the server as well. A record file has been maintained to track the transfer of the data among research team members.

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 were saved for traceability.
- Interviewer was accompanied by the supervisor to monitor the quality of the data being collected
- Interviews and focus group discussions were recorded in two recording devices i.e. two separate audio sticks to prevent the loss of data
- Back checking and proof reading of transcripts and translations to ensure nothing is missed out
- 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?

All transcripts are anonymous and identifying details are removed or given pseudonyms in the transcript and translated documents. The audio files have been trimmed in order to remove those parts that contain any identifiable information. Additionally, consents have been taken from the participants before the start

of the interview and focus group discussion 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 anonymised data generated from this study will be kept for a minimum of 5 years after the study is finished. The tools used for data collection i.e. interview guide and focus group discussion guide will also be archived.

How will the data be preserved?

The anonymised and cleaned data including the translations, interview and focus group discussion guides, code books 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 data will be shared in separate folders e.g. a folder of IDIs will have 33 separate folders representing each of the 32 in depth interviews and 1 for the interview guide and metadata. Within each of these folders will be all data relevant to that interview i.e. the audio file, the transcript and the translation.

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 project except the audio recordings will be shared and archived in Edinburgh 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.