



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

Name:	Dr. Tabish Hazir
Project Title:	Exploration of Pneumonia Related Policy Formation and Implementation in Pakistan
Institute:	Maternal, Neonatal and Child Health Research Network (MNCHRN)
Start Date:	01 May 2018
End Date:	30 September 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 Project
Dr Hana Mahmood	Data deposit and oversight of implementation of the plan
Hira Kiani	Data collection and management
Sajid Khan	Data collection
Omair Sheikh	Data collection
Saman Mujeeb	Maintenance of data
Rakhshanda Hameed	Data analysis
Syed Yahya Sheraz	Data Storage

1. Data Capture

What data will be generated or reused in this research?

This project has/is being accomplished through two phases and the first phase is in turn divided into two sub-phases (A & B).

Phase 1 –A. Content Analysis:

Following types of document (in paper and/or electronic form) were acquired from government websites, government archives, and through requests to Ministry of Health officials (and those from related Ministries), development partners, and researchers involved in pneumonia related research in the past.

- Government strategic plans (5-10 year)
- Health ministry action plan



- National or provincial pneumonia related strategies/guidelines
- General program guidelines with pneumonia as a component
- Any training guidelines on pneumonia
- Any facility guidelines
- Any community intervention guidelines
- Any monitoring guidelines

Both federal and provincial level policy and program documents were included. Although majority of the documents have been obtained through open public domains, however, for those documents which were not open access, permissions will be obtained from the givers to deposit the documents in the repository.

Phase 1 B. In depth Interviews:

In depth interviews were conducted using a semi structured pretested interview guide translated into Urdu with key stakeholders who are or were involved in the formation or implementation of the policies or programs related to pneumonia in children under five years of age. These were conducted either face to face or telephonically, based upon respondent convenience, by a member of the research team and were audio recorded. 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.

20 interviews with key stakeholders have been conducted.

Phase 2 - Net Map Activity:

Net map activities were conducted in the form of group discussions with key stakeholders who are involved in the formation of the policies related to pneumonia in children under five years of age. Each activity had around 8 to 10 participants. These were audio and video recorded. Field notes were also taken during the activity by a note taker.

4 of the discussions were conducted face to face before COVID-19 but due to the ongoing COVID-19 crisis, face to face activities were challenging and therefore an option of online/ telephonic meetings was introduced for remaining two sites.

Recordings were transcribed and translated for analysis.

Data including audio and video recordings, transcriptions, translations, and field notes have been stored in password protected and encrypted computers with access to only members of the research team. These computers have been encrypted using Bitlocker and the folders are stored in encrypted containers using the 7-zip.

Four group activities (Federal, Punjab, Khyber Pakhtunkhwa, and Sindh) and eight telephonic activities (Azad Jammu Kashmir and Gilgit Baltistan) were conducted with 45 participants.

How much data will be generated?

30 - 50 GB of data will be generated

2. Data Management

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

Phase 1. A. Content Analysis

For the content analysis of the documents, a matrix has been developed from the data extracted from the analysis of the acquired documents and is stored as a CSV file.

For each policy document, following details have been noted:

- a. Name of Policy
- b. Year of release
- c. Released by (e.g., which ministry)
- d. Component/section of the document with information related to pneumonia in children under five years
- e. Any mention of pneumonia

Phase 1. B. In-depth Interviews

For audio recording, video recording, transcript and translation of each interview 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

Phase 2- Net Mapping

For audio recording, video recording, transcript and translation of each net mapping activity 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., notepad 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:

Phase 1

A. Content Analysis

All the policy documents and the matrix developed from the data extracted from the analysis of the acquired documents are stored in password protected and encrypted computers with access to authorized members of the research team only.

B. In-depth Interviews

Data including interview guide, audio and video recordings, transcriptions, translations, code books and field notes are stored in password protected and encrypted computers with access to authorized members of the research team only.

Phase 2- Net Mapping

Data including interview guide, audio and video recordings, transcriptions, translations, photographs, and field notes are stored in password protected and encrypted computers with access to authorized members of the research team only.

Backup:

3 backup copies of the original data from all the phases have been created and stored on different locations. Original copy has been stored on the password protected computer of the data manager placed in the MNCHRN 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 a 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., organization name_projectID_subjectID_type of data file e.g., MCNHRN_RES004_I.1.1_transcript.docx

Organization of folders:

Folders according to the geographical locations of the sites of data collected have been created. Files have been stored in their sub folders created according to the file type (audio, photos, transcripts, translations).

Data transfer:

Data was transferred to the computer of the data manager on the same day after the interview / net map activity. The original is deleted after transferring data to Data Manager's computer. The audio file was trimmed by the Data Manager to delete any identifying information before sharing for transcription.

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 net map activities were recorded in two recording devices i.e., two separate audio sticks to prevent the loss of data
- Back checking of transcripts and translations.
- Data stored on 3 locations to prevent the loss of the data. Original copy has been stored on the password protected computer of the data manager placed in the MNCHRN headquarters. This has been backed up on the main server located at the central field office of MNCHRN and on cloud storage on a 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 translations. The audio files have also been trimmed to remove the identifiable information. Additionally, consents have been taken from the participants before the start of the interview and net map activity.

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 tools used for data collection i.e., interview guide and the policy documents will also be preserved.

How will the data be preserved?

The data from all phases of the study will be cleaned and de-identified before preservation. This anonymised data including the matrix of content analysis, 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 –DataVault.

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 DataVault repository) to ensure preservation and continued access. DataVault is restricted access long term storage. Data is held encrypted and access to the data is restricted to nominated individuals. Datasets can receive a DOI via the University Pure system and the metadata is publicly available, promoting data discovery.

6. Sharing and Publication

Which data will be shared and how?

Despite data being anonymised, there is a possibility that respondents might still be identifiable. Therefore, the anonymised data generated from this project except the audio recordings will be shared and archived through Edinburgh DataVault which is a restricted access archive controlled by the data owners and appropriate for sensitive information.

Are any restrictions on data sharing required?

Despite being anonymised, there are concerns of potential misuse if data will be shared openly, therefore data will be stored on DataVault and will be shared through a data sharing agreement. Thus, we will make the anonymised data and associated documentation available to users only under a data-sharing agreement that provides for

(1) a commitment to using the anonymised data only for research purposes and not to identify any individual participant.



- (2) a commitment to securing the anonymised data using appropriate computer technology; and
- (3) a commitment to destroying or returning the anonymised data after analyses are completed

