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

Name:	Siti Nurkamilla Ramdzan
Project Title:	Culturally tailored school-based intervention for asthma in Malaysia (CuT-AsthMa)
Institute:	University of Malaya, Kuala Lumpur Malaysia
Start Date:	1st April 2018
End Date:	31st March 2021
DMP version number and date:	DMP Version 1 and 01.12.2020

Responsibilities & Resources (applicable across the sections below)

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

Person in charge:

1. Coordinating Investigator: Dr Siti Nurkamilla binti Ramdzan
2. Principle Investigator for RESPIRE Malaysia: Professor Dr Ee Ming Khoo
3. Research Manager: Dr Jayakayatri Jeevajothi Nathan
4. Research Assistant: Nursyuhada binti Syukri

Resources:

Hardware and software used in the management of data are listed below:

1. Laptop (encrypted and secured with username and password and kept in a locked cabinet)
2. Thumb drive (encrypted and secured with username and password and kept in a locked cabinet)
3. External hard disk (encrypted and secured with username and password and kept in a locked cabinet)
4. DataStore (to back up active research data)
5. Microsoft Word
6. Microsoft Excel
7. SPSS SPSS Software version 25.0
8. NVivo Qualitative Data Analysis Version 12.0 Software
9. Locked filing cabinet for hard copies (questionnaire and other relevant forms)

1. Data Capture

What data will be generated or reused in this research?

There are 3 phases in this project. Data that were or will be generated are listed below:

1. Phase 1: Systematic review of school-based intervention for asthma among primary school children. Eligible studies were obtained from electronic databases and data collected are available online.
2. Phase 2: Qualitative study of stakeholders to inform development of school-based intervention for asthma in Malaysia.
3. Phase 3: Feasibility study of developed intervention. Qualitative and quantitative data will be collected in this study.

Data generated for phase 2 and 3 will comprise of:

1. Quantitative data:
    i. Socio-demographic data (gender, age group, professional role) will be directly obtained from participants/parents of participants.
    ii. Implementation outcomes data: Participants will be de-identified and questionnaire for outcome data will not contain any identifiable information. Questionnaire data on self-management practices, physical activity and anonymous feedback forms will be obtained directly from children/parents. Data on adherence to intervention and drop out will be obtained from the attendance/participation list. Data on the number of children with asthma attacks and adherence to new practices will be obtained from school records. The researchers will collect data on inhaler technique score by assessing the technique of participants. Data on asthma action plan and review date for asthma from healthcare providers via the participants.
    iii. Health outcomes data: Questionnaire data on asthma control and other asthma-related outcomes including asthma diary will be obtained directly from participant/parents of participants.
2. Qualitative data:
    i. This qualitative data will be collected via audio recording of focus groups/individual interviews. Participant interviews will be recorded with digital audio recorders and subsequently downloaded into the encrypted laptop and secured with username and password. Interviews will be transcribed into word documents and de-identified at the time of transcription. All interview recordings and transcripts will be assigned participant codes and not associated with participant's information. Audio recordings will be deleted at the end of the research project.
    ii. Identifying data such as participant names and codes will be kept in paper format. This hard copy will be kept and stored in a locked cabinet for five years after which all hard copies will be shredded and disposed in secure bins.

How much data will be generated?

50-250GB

2. Data Management

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

1. Phase 1: data will not be shared since the data was obtained from published articles and data are available online.
2. Phase 2 and 3:
   i. Quantitative data
   All data from the completed questionnaires were entered into the Excel spreadsheet. The numbering of the variables will follow the numbering in the questionnaire. A codebook containing information e.g. the information on the study design, sampling methodology and other details of the data will be archived along with the data to assist understanding.
   ii. Qualitative data
   Each interview was given a unique identification code. The interviews were transcribed verbatim in a Microsoft Word file. The interview transcripts were exported into NVivo for data analysis. All data generated from an interview were tagged with the same identification code.

To ensure that all members of the project and a future user would understand the research data and outputs, all supporting documents (s) will be kept with the data. This will preserve the data to support the research findings and allow others to replicate your results.

Where will the data be stored and backed-up?

The electronic data were stored in two RESPIRE laptops, a thumb drive, a solid-state drive and will be backed-up in DataStore. Hard copies (research protocol, topic guide, participant information sheets, sociodemographic forms, informed consents, questionnaires and NVivo coding framework) were stored in locked cabinet.

3. Integrity

How will you quality assure your data?

The completed questionnaires were checked by research assistants for completeness of the data before it was entered into Excel spreadsheet. Data entered were double checked by a member of the team for possible errors. If any missing information were found, participants would be contacted for clarification.

All the transcripts were transcribed assisted by field notes containing information e.g. non-verbal cues not captured using audio recorders. The transcripts were de-identified and checked by a member in the researcher team for accuracy. Some transcripts will be translated to English and the translated transcripts will be checked by a member in the research team.

4. Confidentiality

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

1. The study will be conducted in accordance with the principles of the International Conference on Harmonisation Tripartite Guideline for Good Clinical Practice (ICH GCP). All required approvals have been obtained.
2. The investigator is responsible for the overall conduct of the study at the site and compliance with the protocol and any protocol amendments.
3. All evaluation forms, reports, and other records will be identified in a manner designed to maintain participant confidentiality. Clinical information will not be released without the written permission of the participant. The Investigator and study site staff involved with this study may not disclose or use for any purpose other than performance of the study, any data, record, or other unpublished, confidential information disclosed to those individuals for the study. Prior written agreement from the sponsor or its designee must be obtained for the disclosure of any said confidential information to other parties.
4. All Investigators and study site staff involved with this study must comply with the requirements of the appropriate data protection legislation including the General Data Protection Regulation and Data Protection Act 2018 with regard to the collection, storage, processing and disclosure of personal information and will uphold the Act's core principles.

5. Retention and Preservation

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

1. Phase 1: not applicable
2. Phase 2 and 3:
   i. Identifiable Data
   All sociodemographic forms and informed consents of the participants will be kept for five years. Audio recordings will be permanently deleted at the end of the research project.
   ii. Non-identifiable Data
   De-identified electronic data will be preserved for long term.

How will the data be preserved?

The hard copies of these identifiable data will be kept in a locked cabinet and will be shredded and disposed in secure bins after five years. The non-identifiable data will be preserved in two RESPIRE laptops, a thumb drive, a solid-state drive and on DataVault.

6. Sharing and Publication

Which data will be shared and how?

The final de-identified analysed data, conference abstract presentations, journal articles and dissemination event will be shared in DataShare.

Are any restrictions on data sharing required?

Yes, prior to release for sharing, there remains the possibility of deductive disclosure of subjects with unusual characteristics. Thus, for data archived in Edinburgh DataVault, we will make the data available to users only under a data-sharing agreement that provides for:

1. A commitment to using the data only for research purposes and not to identify any individual participant;
2. A commitment to destroying or returning the data after analyses are completed.