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

Name:	Professor Dr Ee Ming Khoo
Project Title:	Pulmonary Rehabilitation for Patients with Chronic Respiratory Diseases in Malaysia: A Feasibility Study
Institute:	University of Malaya, Kuala Lumpur, Malaysia
Start Date:	1st May 2020
End Date:	30th June 2021
DMP version number and date:	DMP Version 1 and 06.03.2021

Responsibilities & Resources (applicable across the sections below)

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

Person in charge:
Principal investigator, principal investigator at the site and research manager of the project are listed below:

1. Principal Investigator: Professor Dr Ee Ming Khoo
2. Principal Investigator at the site: Dr Soo Chin Chan, Associate Professor Dr Julia Patrick Engkasan
3. Research Manager: Jayakayatri Jeevajothi Nathan

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, password protected and kept in a locked cabinet)
3. Solid-state drive (encrypted, password protected and kept in a locked cabinet)
4. DataStore (to back up active research data)
5. Microsoft Word
6. Microsoft Excel
7. NVivo Qualitative Data Analysis Version 12.0 Software
8. SPSS Software version 25.0
9. Locked cabinet for any hard copies (research protocol, topic guide, participant information sheets, sociodemographic forms, informed consents, educational booklet home exercise diary and NVivo coding framework)

1. Data Capture

What data will be generated or reused in this research?

Data that were generated are listed as below:

1. Phase 1 (Baseline):
   (a) Patient - sociodemographic and personal profiles such as age, gender, ethnicity, diagnosis, smoking status, educational status, occupation, medication usage, recruitment response rate, attendance throughout the programme and retention rate of participants with reasons (if provided).
   (b) Pre-quantitative assessment - COPD Assessment Test (CAT), 6 Minutes Walking Test (6MWT).
   (c) Post-quantitative assessment - COPD Assessment Test (CAT), 6 Minutes Walking Test (6MWT).
2. Phase 2:
   (a) Quantitative (Health care professional) - age, gender, profession, and length of experience in specialty
   (b) Qualitative transcript of patients, family members and health care professionals to assess factors that influence the implementation of the home-based pulmonary rehabilitation programme.

How much data will be generated?

50GB – 250GB

2. Data Management

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

1. Phase 1 (Baseline):
   Source documents will be documented in the Microsoft Excel and SPSS. A codebook containing information about each of the variable in the dataset were created to ensure that the data is understood and interpreted properly.
2. Phase 2:
   Each interview was assigned with a participant code and transcribed verbatim in Microsoft Word. The de-identified interview transcripts were exported into NVivo Qualitative Data Analysis Version 12.0 Software for data analysis.

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, educational booklet, home exercise diary and NVivo coding framework) were stored in locked cabinet.

3. Integrity

How will you quality assure your data?

Data entered were or will be checked by researcher for completeness and accuracy.

The interviews were recorded by using two audio recorders and the non-verbal cues were noted down by researcher. Interviews were transcribed verbatim and the transcripts were checked by two researchers for accuracy.

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 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?

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.

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.

This research was funded by the National Institute for Health Research (NIHR) Global Health Research Unit on Respiratory Health (RESPIRE) using UK aid from the UK Government to support global health research.

The views expressed in this publication are those of the author(s) and not necessarily those of the NIHR or the UK Department of Health and Social Care.