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

Name:	Dr GM Monsur Habib
Project Title:	**Enhancing access to pulmonary rehabilitation (PR) through implementation research in Bangladesh**
Institute:	Bangladesh Primary Care Respiratory Society
Start Date:	June 2018
End Date:	August 2021
DMP version number and date:	Version 1; 03 December 2020

Responsibilities & Resources (applicable across the sections below)

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

The research team will have direct involvement in the data management and they are:

1. Dr GM Monsur Habib
2. Professor Hilary Pinnock (Main Supervisor; Edinburgh)
3. Dr Roberto Rabinovich (Second Supervisor; Edinburgh)
4. Professor Samir K Saha (Site Supervisor; Bangladesh)
5. Dr Aftab Uddin (Site Supervisor; Bangladesh)

My project will have five phases:
Phase 1: Initial stakeholder engagement
Phase 2: Systematic Review
Phase 3: Developing PR protocol for low resource setting
Phase 4: Conducting feasibility study of PR
Phase 5: Second stakeholder engagement

As a part of my responsibility in this PhD programme, I will collect, store, manage data and all the data will be stored safely in the DataStore of the network of the Edinburgh University.

By this time I will complete a course on data management from Coursera or other recognised online courses.

We will only keep unidentifiable data. Access to collated unidentifiable participant data will be restricted to individuals from the research team treating the participants, representatives of the sponsor(s) and representatives of regulatory authorities.

2. Data Management

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

Each phase will be based on protocols being developed at the start of each phase. I will index each phase in a codebook which will contain all the details on fieldwork (including notes on deviation from protocol and reasons behind this) and discussion on why individual decisions are made. This documentation will help using the data in the future can be done effectively and accurately.

Phase 1: Anonymous statements of the stakeholder with the development of theme http://www.jogh.org/documents/issue202002/jogh-10-020384.htm.

Phase 2: The published paper (Systematic review: https://www.nature.com/articles/s41533-020-00210-y) will have all the data and tables as included and supplementary file.

Phase 3: Table of the components of PR as well as the figures building on the information obtained from systematic review, evaluation of standard guidelines, training on PR at various centres, feedback from the stakeholders will be persevered with clear description and usable form.

Phase 4: Both the quantitative (age, weight, height, lung volume, Incremental Shuttle Walking Test (ISWT), Endurance Shuttle Walking Test (ESWT), modified Medical Research Council (mMRC), COPD Assessment Test (CAT), Hospital Anxiety and Depression Score (HADS) etc.) and qualitative data (transcription of interview with analysis e.g. thematic framework) will be tabulated in a worksheet with the de-identification of personal information. A summary will be drawn e.g. mean age±SD, so that it can be easily interpreted for future use.

De-identification: We will give a code against each participant e.g. Mrs Monowar will identified by the number B0012, indicating the participant was in group B and her serial was 12. None of the participant and provider will know the name of the participant in this group. As the lead researcher only I will keep and be able to access the code or ID and an original copy of the list where I can identify the participant.

Phase 5: Second stakeholder engagement: This stakeholder engagement data will also be recorded and transcript in Bengali. The transcript will be translated in back-to-back translation method. The theme will be generated as per the rule of qualitative analysis (I have completed MPH course of qualitative analysis under the University of Edinburgh) and data will be stored anonymously.

Qualitative interviews will be recorded on an encrypted digital recorder, transferred securely to a Bengali-speaking transcriber who is one of the research team members employed by the Khulna pulmonary rehabilitation service. As an ongoing study, we have conducted interviews of four participants only and these have been transcribed. English translation is also done with the quality assurance by back-translation. The participants are de-identified by using a unique ID.

Where will the data be stored and backed-up?

I will keep my data/files on the University of Edinburgh DataStore, either through my university username account access or through the DataSync facility. I will be guided by the RESPIRE Data Science and Methodology Coordinator and UoE Research Data Service.

The data will be stored on the University of Edinburgh's filestore ensures high-quality storage with guaranteed back-up (automatic data replicate to an off-site disaster facility). If I am outside the university network (doing data collection in Bangladesh), I will have access to this facility through VPN network.

The information we will be collecting in paper copies will be stored under lock and key in Bangladesh Primary Care Respiratory Society, Bangladesh. Despite various available methods of data storing, in our practice setting (LMICS) we will adapt external storage media, such as tapes or discs along with the hard copies.

3. Integrity

How will you quality assure your data?

Phase 1: Data of Stakeholder engagement, recorded by two independed reviewers and disputes were resolved by the supervisory panel. The paper is peer-reviewed and published, please find the following link: http://www.jogh.org/documents/issue202002/jogh-10-020384.htm.

Phase 2: Systematic Review: From the very beginning we have been careful about the quality of data. Every step was cross checked by supervisory team. For quality assurance, literature search was endorsed by senior librarian of the University of Edinburgh, data selection, extraction, analysis, and final reporting were made by two independent reviewers and final decisions we made through consensus or discussion with the supervisory team. All of these procedures are published in the protocol (https://www.nature.com/articles/s41533-019-0122-1) and in the review with both the included and supplementary file (https://www.nature.com/articles/s41533-020-00210-y)

Phase 3: Table of the components of PR: Building on the findings of stakeholders meeting, systematic review, and extensive evaluation of existing standard guidelines, we developed the deliverable components of PR for LMICs, which is rechecked by supervisory team.

Phase 4: Feasibility study: Both the quantitative (age, weight, height, lung volume, ISWT, ESWT, mMRC, CAT, HADS etc.) and qualitative data (transcription of interview with analysis e.g. thematic framework) will be tabulated in a worksheet with the de-identification of personal information. A summary will be drawn e.g. mean age±SD, so that it can be easily interpreted in its future use. Despite continuous supervision of ensuring the quality of data by the supervisory team it will be finally rechecked and endorsed by principal supervisor.

Phase 5: Second stakeholder engagement: This stakeholder engagement data will also be recorded and transcript in Bengali. The transcript will be translated in back-to-back translation method. The theme will be generated as per the rule of qualitative analysis (I have completed MPH course of qualitative analysis under the University of Edinburgh) and data will be stored anonymously. Quality of data will be ensured by the authorisation of the supervisory team.

As I am a medical practitioner and patients will be recruited from my practice, I may not get the information from my patients with their actual perception, that's why we decided to appoint one trained independent research assistant who will conduct all the interviews from the patients, providers, and other stakeholders.

4. Confidentiality

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

I have received ethical approval from the Institutional Review Board BRAC James P Grant School of Public Health BRAC University, Bangladesh and sponsorship from the Academic and Clinical Central Office for Research & Development (ACCORD) for the University of Edinburgh's Research for all my PhD project. Due to COVID-19 pandemic amendments were made for phase 4, where I revised the method of data collection from one-on-one meeting to online meeting, and both IRB Bangladesh and ACCORD approved this.

In phase 4, I am currently continuing feasibility study as per the amended protocol. The intellectual property (IP) generated from this study will be shared between the University of Edinburgh and Bangladesh Primary Care Respiratory Society (BPCRS).

The participants will be referred from my practice in Khulna, Bangladesh, and the assessments will be performed within this practice setting. Their personal health information (PHI) therefore will be stored and recorded for clinical purposes in my usual practice record keeping system. However, the data collected for the purpose of feasibility study will be stored as per protocol described above.

The patient recruitment will start from the receiving the referral patients for PR. One research assistant will explain to the patient according to the information provided in the patient information sheet (PIS) and allow them to think about it to take decision. Once the patient decides to take part in the study, a consent form will be provided to them and the trained researcher who has completed a certificate course on consent taking will take signed consent from the patient.

All evaluation forms, reports, and other records must 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 the 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.

5. Retention and Preservation

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

We planned for the product of the final de-identified quantitative data to be available for use by the research and policy community for phase 2 and 4. For qualitative data in phase 1 and 4, we planned to share the overarching themes of the final data in the text version. All audio-files will be destroyed once data analysis is complete.

Phase 1 & 5: I will retain and not share further the qualitative (opinions given by the stakeholder and their position in the power/impact quadrant) and quantitative (how many stakeholders attended each meeting labelling their profession) data which are partly published (http://www.jogh.org/documents/issue202002/jogh-10-020384.htm)

Phase 2: Quantitative data (number of hits in searching, selected articles, total number of participants including intervention and control group, age, weight, height, dyspnoea, CAT, HADS and other variables, outcome measurement data) will be retained part of which are published: https://www.nature.com/articles/s41533-020-00210-y
It will be retained locally at BPCRS (both the electronic and paper version), part of it in published journal, and Edinburgh DataShare open access repository (number of hits in searching, selected articles, total number of participants including intervention and control group, age, weight, height, dyspnoea, CAT, HADS and other variables, outcome measurement data), as all of them will not be available in the published journal and BPCRS storage may be damaged in adverse circumstances.

Phase 3: It will include tables of components and models of PR from SR, training course on PR, and stakeholders view; part of it are already published in open access journal. Others will be retained in BPCRS and Edinburgh DataShare repository.

Phase 4: Feasibility study: As the patients will be referred from my practice, as per my practice trend I will keep the PHI for 7-10 years and sometimes more if required. I am not aware of local rule about how long to retain the data. However, we will not share the transcription and coding of the qualitative study as it might include non-shareable elements by law.

How will the data be preserved?

Data will mainly be preserved either locally at BPCRS or through Edinburgh DataShare (see above).

6. Sharing and Publication

Which data will be shared and how?

This study is part of a PhD programme of work and will be reported in a publicly available thesis. I have submitted the findings of this study to conferences for presentations and a high impact peer-reviewed journal for people who have access to the relevant sites. Other methods of dissemination will include innovative dissemination channels of RESPIRE (websites and Twitter) to raise awareness of our publications.

We will share de-identified quantitative and qualitative data in Edinburgh's DataShare.

Systematic Review: As few data captured in the systematic review are already published in the open access peer-reviewed journal and presented at various seminars and conferences, the only data that are not published will go into Edinburgh Datashare.

Feasibility: In the feasibility study the blank templates; coding guides; interview guides; blank consent form is open and need not to restrict, however, they are the intellectual property of the research. That's why it should be retained in the DataShare.

Are any restrictions on data sharing required?

The University of Edinburgh is the data controller. Any data breaches will be reported to the University of Edinburgh Data Protection Officer who will report to the relevant authority according to the appropriate timelines if required.

Yes, some of data are developed as intellectual property and kept closed until formal write up for publication or thesis writing. These data/items will be restricted for the time being.

Other reasons for the restriction of data is that the population of the study are from one site and one clinician will refer the patients from one community, which could lead to potential identification of the participants.

If the data is shared openly participants might be identified due the points mentioned above; in that case few sensitive elements may be retained in the closed repositories (to be determined and agreed).

We will share the data in DataShare with an embargo period until publication or thesis writing has been completed.