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

Name:	Dr. Mohammod Jobayer Chisti
Project Title:	Feasibility and acceptability of bubble continuous positive airway pressure for treatment of Bangladeshi children with severe pneumonia
Institute:	The University of Edinburgh, UK & International Centre for Diarrhoeal Disease Research, Bangladesh (icddr,b)
Start Date:	01 September 2018
End Date:	30 June 2021
DMP version number and date:	Version 1.0; 08 December 2020

Responsibilities & Resources (applicable across the sections below)

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

Task	Responsible Personnel	Contributors
Development of case report forms (to align data collection with requirements of the study protocol)	Site PIs	Site Coordinator Medical Officer Data Management Team Scientists from the University of Edinburgh (UoE)
Development of data management system	Data Manager	Site Coordinator Data Management Team Other Site Staff
Study implementation: • Screening of potential participants • Enrolment • Planning and completion of follow-ups • Collection and entry of data • Monitoring of visit completion and QC	Site PIs	Site Coordinator Site Data Collectors Site Data Management Team
Review and quality control of data collection and entry	Site PIs	External Clinical Monitor Site Coordinator Data Manager
Transmission of data to UoE	Data Manager	Site Data Management Team

Design, development, and production of study reports: • Data descriptive • Monthly study monitoring reports • Query reports	Data Manager	Site Coordinator Site Data Management Team
Resolution of data issues	Site PIs	Site Coordinator Site Data Management Team Site Data Collectors

1. Data Capture

What data will be generated or reused in this research?

Both Quantitative and qualitative data will be collected.

Quantitative Data:

Study personnel will collect data on case report forms (CRF) initially which will have a study number. Patients who will be screened as eligible for the study criteria will receive bCPAP. Baseline data will be collected when the informed consent is given by the participant's (i.e. the child's) legal guardian during enrolment. The baseline dataset includes present complaints, clinical findings, oxygen therapy, socio-demographic and diet information.

The participants will be followed-up till discharge and information regarding clinical appearance, oxygen flow, systemic examination, and complication will be collected and documented in the follow-up form.

At the time of discharge the conclusion data of the participant will be collected that will include duration of oxygen therapy, treatment procedures, nurses and physicians' follow-up, outcome etc.

Paper-based forms containing the participant's information will be transcribed later into electronic format by using personal computer or TAB into REDCap. All screening patients' information will be preserved in the screening log. The information of patients who will not be eligible for enrollment according to the inclusion and exclusion criteria will also be preserved. This will help to know the prevalence as well as basic characteristics such as distribution of age, sex, exclusion reasons etc. of the patients the study period.

We are experienced in capturing data into REDCap which is safe and secured. The online database of REDCap system will help to send the quantitative data to a central server from two sites. Central data management personnel will download and save the files into different formats (SPSS, Stata, SAS, R, CSV/Microsoft Excel etc.) and devices as required.

Computers and all other devices (laptop, tab – we will encrypt our devices) for the purpose of this study's data management are secured at icddr,b by its own system providing with

individuals' secured login (users with passwords) which is monitored by a group of strong IT professionals. All licensed software used here. In icddr,b there is no problem with electrification as it has own power distribution capacity (UPS, IPS, high production capacity generator) beside the regular distribution.

Qualitative Data:

RAs and Md. Fakhar Uddin will collect qualitative data through Focus Group Discussion (FGD) and In-depth Interviews (IDI) using flexible semi-structured questions. Physicians, nurses, associated staff and patient's parents/legal guardian will participate in FGDs and IDIs. Participant's information will be kept confidential, none other than the investigator of this study, and only the Ethical Committee of the icddr,b will have an access to that information. They will use an audio recorder to capture the conversation/discussion during FGD/IDI, copy the audio recorded data file to transfer from the recorder to their official computers, and then will transcribe verbatim to transform the audio-recorded data into textual data in Microsoft word file and the textual data files will be saved in their computer with a specific file name. We will generate additional anonymised transcripts where participants name will be replaced with a pseudonym in order to maintain the confidentiality- that of keeping participants' identities secret. For quality control, in order to maintain the originality of the discussion/conversation in the written transcription, they will transcribe the audio-recorded data immediately after conducting the IDI and FGD as well as will listen to the audio recording again to re-confirm the accuracy of the written textual data. Then, they will share the audio-recorded data file and a typed textual data file with qualitative study lead (Md. Fakhar Uddin) and PI of the study to check the accuracy of data and perform data analysis.

How much data will be generated?

Quantitative data:

- Patients screening will take place as much as require till 20 enrolments.
- Approximately 0-50 GB

Qualitative data:

- In total 4-6 FGDs and 16-20 IDIs
- 0-50 GB

2. Data Management

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

Quantitative Data:

Paper CRFs will be Scrutinized and try to find out the gaps and inconsistencies with logical flow and if found anything unusual, we will consult/return to the data collection team who

will help for the correction. This process will help us to have robust data into the REDCap. In REDCap we will also use the quality assurance tool to resolve all the possible inconsistencies (may require to consult with GPs if necessary as they will be responsible for data collection in site) and this will help for the correction.

Qualitative data:

In the study field/site, RAs will organize qualitative data following the key steps include labelling all materials (recorded data, consent forms, field notes, and impression notes), take photos of consent forms and field notes (convert hard materials to soft materials), save recorded data and photos of consent forms, field notes in computer, external hard drive, and the UoE network (Datastore) facilitated through a drop like service called DataSync (to allow file storage) , to prepare a data collection update table (background/demographic information, interview code, original names, pseudonyms, age, sex of each respondent). After coming back to the Dhaka office of icddr,b, we will organize and manage data following the key process such as transcription of recorded data, expanding field notes and impression notes, back up of all data (in computer, email, hard copies), preserve written transcriptions, writing pads and recordings, maintain an update table with the list of all transcriptions. All hard copies (consent forms, field notes, and photo) will be transferred into soft copies and saved in computers, an external hard drive icddr,b repository and the UoE network (Datastore) facilitated through a drop like service called DataSync (to allow file storage). The soft copies (transcripts, scan copy of consent form & field note, impression note, photos) will be saved in an assigned password protected computer. All hard copies and external hard drive will be stored in a durable file cabinet under lock and key of icddr,b hospital building which is a well secured office place with 24 hours supervision of central security team. The keys will be kept by qualitative study personnel (Md. Fakhar Uddin) and PI of the study.

Where will the data be stored and backed-up?

All the paper based data will be preserved in a secured room using file cabinet under lock & key. The electronic data will be saved in a secured server and also in password protected and encrypted personal computer, external hard drive, for each site and in the UoE network (Datastore) facilitated through a drop like service called DataSync (to allow file storage).

Study personnel will bring the paper based data to central data management team from sites. The data will be stored and saved in the Bank Building of icddr,b using a file cabinet under lock & key. This is a safe and secured place to keep all the recorded data for five or more years.

As we are supposed to capture all quantitative data into REDCap, it will be stored centrally the electronic data. Besides these will export and store in our computer hard disk, Portable Hard drive, flash drives, etc in weekly basis.

Qualitative data:

Md. Fakhar Uddin (qualitative study lead) and PI of the study will store the audio recorded and textual qualitative data (Microsoft word files) in their official computers. In addition, data will be backed-up in computers, external hard drive, and the UoE network (Datastore) facilitated through a drop like service called DataSync (to allow file storage) and kept it in a secured locker under lock and key of icddr,b hospital building which is a well secured office place with 24 hours supervision of central security team. The keys will be kept by qualitative study lead (Md. Fakhar Uddin) and PI of the study. The soft copies such as audio files, photos, emails etc. will be saved in an assigned password protected computer.

3. Integrity

How will you quality assure your data?

Quantitative Data:

The individual supervisor will monitor the quality of each person's work following the SOPs and will ensure the quality. These will be done randomly and regular basis. These should be essentially discussed in weekly meeting among the study personnel which will definitely help to improve and maintain the high quality of the study. We will assign unique Screening IDs for all the screened patients. Among the screened patients who will be given consent will consider enrolled patients. The enrolled children aged between 2 month and 24 months will then be identified with their different Study IDs which will also be unique. In all case report forms the Study IDs will be inscribed.

Before data analysis, data cleaning and management will be carried out. In a process of scrutinization by both GPs and Data Management Team will generate a list of inconsistencies and problems occurring in data collection process which will be shared with whole data collection team for improvement. After the scrutinization of CRFs data management team will capture the data into REDCap and with a systematic process Data Management Team will preserve the paper CRFs of Participants in a secured locked File Cabinet for reuse.

Qualitative data:

Prior to qualitative data collection, PI of the study and Md. Fakhar Uddin (qualitative study lead) will provide one-week intensive training to the newly recruited RAs to make them knowledgeable about the research topic, study design, and methods of qualitative data collection, techniques (neutrality, insider perspectives) and procedures of data collection. In addition to the RAs, including the lead anthropologist (Md. Fakhar Uddin), will conduct the IDIs and FGDs. Immediately after completion of an IDI/FGD, interviewer will discuss the collected data amongst team members and then will share the audio-recorded data file and its written transcription with the study lead (Md. Fakhar Uddin) and PI of the study to check data quality and identify data gaps in order to collect rich data from the next IDI/FGD. We will do triangulation of findings comes from different data collection methods, participants and interviewers. The team leader (Anthropologist) will oversee the

whole data collection process and maintain communication daily with RAs to ensure the quality of data. We will conduct IDIs/FGDs until obtaining data saturation.

4. Confidentiality

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

During the course of the research all the collected information will be kept confidential and there are strict laws which safeguard the patients' privacy at every stage. We do hereby affirm that privacy, anonymity and confidentiality of data/information identifying the patient will be strictly maintained. All medical information and the description of treatment of the patient will be kept confidential, under lock and key.

We will share the anonymised data with the co-investigators at the University of Edinburgh using SPSS software or as required.

All relevant documents regarding the participants' research data including reports, baseline data, clinical information, follow-up data, discharged data etc. will be kept in a secure storage area with limited access. We will not share any clinical information except study personnel 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 purpose of the study. The disclosure of any said confidential information to other parties will not be done without the written agreement from the sponsor or its designee.

5. Retention and Preservation

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

Our scheme is to obtain the high quality data by the well trained study personnel and to keep the data (raw, hard copies, form) in a well secured place for a long-term without any violation of the terms and policy.

All study documentation will be kept for a minimum of five years from the protocol defined end of study point. When the minimum retention period has elapsed, study documentation will not be destroyed without permission from the sponsor.

How will the data be preserved?

We will keep all the consent forms, recorded data, field notes, hospital notes, impression notes, track records, audio files and photos in the form of hard copies in a secure storage area with limited access and the specific files in electronic transcribed form using appropriate software. Materials which contain patient identifying data will be kept separate from any anonymised material and then will be sent for data processing. All the paper-based anonymised documents will then be scanned as soon as possible to make sure as shareable. All data as well as digitized documents will be preserved in the password protected and encrypted computer, , external hard drive, and the UoE network (Datastore) facilitated through a drop like service called DataSync (to allow file storage) etc. From our repository the final electronic database will be later shared or given access to The University of Edinburgh. The University of Edinburgh undertakes to maintain the digital outputs of this project for the long-term and will utilise University infrastructure to endure preservation and continued access.

All Investigators and study site staff involved with this study must comply with the requirements of the appropriate data protection legislation including 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. Access to collated participant data will be restricted to individuals from the research team treating the participants, representatives of the sponsor(s) and representatives of regulatory authorities.

Computers used to collate the data will have limited access measures via user names and passwords, and encryption where appropriate.

Published results will not contain any personal data that could allow identification of individual participants.

6. Sharing and Publication

Which data will be shared and how?

The final anonymised research data will be released as public with the adherence of data policies of icddr,b and The University of Edinburg (UoE). Considering the future use of the data and prevailing of potential research findings, this data will be available after three years of protocol completion date as recorded by the IRB, publisher, or extraordinary circumstances. Research Administration (RA) of icddr,b will make the data available after receiving the final data from study Principal Investigator (PI). RA will ensure the future use and availability of the data upon request of potential researcher. UoE will also take necessary steps for the availability of data accordingly to their policy. There may be some data/items that we can share openly in Edinburgh DataShare, and there may be some data/items that we cannot share openly, and therefore we will consider depositing the data in Edinburgh DataVaults. For the latter, we will ask each potential user to read the icddr,b data access policy (https://www.icddrb.org/images/stories/AboutUs/icddrb_data-access-policy_nov-07.pdf) and the icddr,b data licensing application and agreement (https://www.icddrb.org/images/stories/AboutUs/icddrb_dlaa_20140518_1.doc. The application will be referred to the icddr,b senior manager of research administration.

The final data in both the repositories of icddr,b and UoE will be made publicly available according to the data access policies. The interested party shall contact repository administrator who will organize a meeting of the Data Repository Committee (DRC) and the respective PI or her/his delegated co-investigator. In the event approval is accorded, DRC will establish a link between the requesting parties with the PI/ designee for provision of needed data and their metadata. The PI/designee shall inform DRC which data have been shared with the requesting party for records.

Further use of bubble CPAP data will require that any abstract, manuscript, publication, dissemination, presentation or other public disclosure based on this datasets properly acknowledges icddr,b, UoE, PI, the original donor and the core donors. The researchers publishing secondary analyses based on this dataset must give each icddr,b and UoE Researchers who made an intellectually significant contribution to the dataset an option to be listed as an author, provided it meets international standards for authorship in peer-reviewed journals and the policies of the publishers; and submit the manuscript to Repository Administrator for review and clearance.

After the end of the study the results will be reported nationally and internationally by attending national and international scientific conferences and by arranging dissemination programmes in icddr,b. Additionally we will publish the findings in international peer reviewed journal(s).

Each Party (icddr,b and University of Edinburgh) shall own the Results/Data generated by it under its Allocated Work and shall be responsible for securing ownership of such Results/Data from its employees, students, officers, representatives and consultants and contractors.

However, all the investigators will have access to the data and will be the authors of manuscripts written from this study. For the main manuscript, first author will be the icddr,b PI of this study. One investigator from the University of Edinburgh, UK may be the senior author and Dr Shams El Arifeen may be the co-senior author from Bangladesh of the manuscript written from this study. If there will be more than one manuscript from this study, first authorship, senior authorship and other authors will be agreed before preparing the draft. The RESPIRE Publication Strategy contains an authorship policy which recommends following the generally accepted rules for academic publication from the ICMJE (www.icmje.org) (The Vancouver Convention). The study team will adhere to this policy.

Are any restrictions on data sharing required?

Any data or derived data, including any primary research data, associated metadata, or any additional relevant research data necessary to understand, assess, and replicate any reported Results in totality ("Research Data"), then such Research Data must be deposited into and published on an open source platform, and be accessible and available to all. As this a contractual obligation icddr,b and UoE will ensure availability of final anonymised data publicly for future use and potential secondary research, and the recipient must adhere with our data sharing agreement.

We will not share any clinical as well as non-clinical information that may disclose participants will not be done without the written agreement from the sponsor or its designee.

All the paper based study documentation will be kept for a minimum of five years from the protocol defined end of study point in a secured room under lock and key and the electronic data in a password-protected personal computer, external hard drive, and the UoE network (Datastore) facilitated through a drop like service called DataSync (to allow file storage) When the minimum retention period has elapsed, study documentation will not be destroyed without permission from the sponsor.