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

Name:	Dr. Osman Mohammad Yusuf
Project Title:	Hypertonic saline nasal irrigation and gargling with hypertonic saline for suspected or confirmed COVID-19: pragmatic web-based Bayesian adaptive randomised controlled trial
Institute:	The Asthma & Allergy Institute Pakistan
Start Date:	01 September 2020
End Date:	31st June 2021
DMP version number and date:	Ver 1.0 3 May 2021

Responsibilities & Resources (applicable across the sections below)

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

Data will be maintained, after vetting by the Principal Investigator and by the Data Management Team, under Dr. Aimal Rextin for entering, maintenance, and security.

1. Data Capture

What data will be generated or reused in this research?

A website will be developed to give participants guidance and information. It will also contain all forms and surveys needed for this study. Data that will be collected include daily diary entries, participant information like name, age, gender, address, contact information, and their replies to questionnaires they will be asked to fill out. This data will be collected through an online web-based database. This server will be located at AAIP office in Islamabad, Pakistan and will be secured with a Web Vulnerability scanner and intrusion detection system. All communication from the participants to the server will be transmitted on HTTPS secure protocol. Advice from a web security expert will be taken and further measures can also be added if it is suggested. In addition, as per advice from the ELVIS COVID-19 Study, we plan to use the Research Electronic Data Capture (REDCap) platform for storing the data collected. We will be in communication with the concerned Data Manager of the ELVIS Study in the Edinburgh Clinical Trials Unit (ECTU) to arrange the logistics of this arrangement.

The advertisement for participation in the study, based on the ELVIS COVID-19 advertising, will be published on local media and different social media platforms in the form of posters video advertisements, social media (Facebook, WhatsApp etc.) and emails to healthcare providers.

Participants will visit the associated webpage given in the advertisement and enrol themselves to participate in this project. The eligibility criteria for each participant to be part of this study are given below:

Inclusion Criteria:

1. Adult (≥18 years)
2. Resident of major cities of Pakistan (for laboratory access)
3. Must be self-isolating within 48 hours of facing symptoms
    a) new continuous cough and/or shortness of breath or difficulty breathing, OR at least two of these symptoms: fever, rigor with chills, muscle pain, headache, sore throat, or recent loss of taste or smell and/or
    b) those with virologically confirmed SARS-CoV-2 infection and clinical symptoms indicative of COVID-19 (as detailed in (a) above).

Exclusion Criteria:

1. Onset of illness (symptoms) more than 48 hours
2. Inability to consent
3. Pregnancy
4. Immunosuppression
5. Inability to perform HSNIG
6. Those taking part in another interventional medical trial

7. Those with suspected/confirmed COVID-19 in whom hospital admission is recommended
8. Those who do not have access to email / internet
9. Those living in a household with another person currently participating in this study

If the participant meets the above criteria, he will be asked to fill out the following ONLINE forms:

1. An Eligibility and Consent Form to show their agreement for using their data in this research project (after reading the Patient Information Sheet)
2. Baseline information at recruitment of the study. This form will contain personal identifier information, including:
    - Name
    - Date of Birth
    - Sex
    - Address
    - Phone number
    - Email
    - GP / Physician Address
3. Questionnaire about the current symptoms being faced by the participant.

After this, the patient will be automatically randomised into either the Interventional group (HSNIG group) or the Control Group (either simple ablution with plain water or no intervention).

The following data will be collected from each participant:

- A Daily Study Diary to be filled online, every day for 14 days by each participant (irrespective of study group).
- A Daily Symptom Report. This will be separate for the Study group and the Control group.
- An End of Illness Questionnaire to be filled online by each participant. This will be separate for the Study group and the Control group. In case a participant improves before 14 days of treatment, he or she will be asked to fill out the "End of Illness" questionnaire when he improves, and once again at the end of a 14 day period from start of his treatment.

All data from the participants including their personal information, questionnaires and daily diaries will be collected through a web portal by designing appropriate forms for each. The diary data will be stored in .txt format, while questionnaires data will be stored in .csv files. The data generated in this project will be entirely co-owned by AAIP and the University of Edinburgh. However, it will be available to be used for research purposes.

Details:

A web application will be used to collect the data, it will also randomly form two equal sized groups: the intervention and the control. The control group will follow the general medication for COVID-19 as suggested by their physician and intervention group will perform HSING a maximum of 12 times a day for 14 days or till the end of illness. Participants will be asked to write daily diary which includes a modified Wisconsin Upper Respiratory Symptom Survey (WURSS 24)– Daily Symptom Report (modified according to the symptoms of COVID-19 as defined by the Government of Pakistan) to analyse length and severity of symptoms. Each participant will be asked to complete an end of illness diary and fill a questionnaire. Those who complete the end of illness diary before day 14 will be asked to answer a questionnaire on day 14.

How much data will be generated?

There will be about 405 participants, each will write daily diary for 14 days and fill a daily questionnaire. They will also write either an end of illness diary or a questionnaire on day 14. All data will be in .txt or .csv format and will be stored in an online database. We estimate that we will need less than 5 GB of memory storage.

2. Data Management

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

Some metadata such as questionnaires, technical notes, etc. will be necessary. This will be published along the research data. We will generate the following:

1. A **technical report** will be prepared for the secondary users to understand the collection procedure and processing of the generated data with a bibliographical citation for users to cite in future publications.
2. A **readme.txt** file to explain how the daily diary and end of illness diary is maintained. It will also include the details of parameters considered in daily diary and the differences between control and intervention group.

All documentation and metadata will be generated in a text file (.txt).

Where will the data be stored and backed-up?

During the project, all data will be stored on .txt or .csv files. The .txt files will be used for daily dairies while the .csv file will be used will be used for questionnaire data. These files will be stored on the local hard drive of a computer at AAIP. These files will also be backed up on a shockproof external local hard disk dedicated to the project and will be password protected to ensure its security. The password of these files will be shared with a very limited number of concerned personnel working for the project. Furthermore, a backup will also be maintained on DataStore, a data facility on a UoE server. The data will be sent securely to DataStore through DataSync.

Once the project is completed all sensitive information including the age and gender of the participants will be removed and stored on DataShare to ensure maximum possible availability.

3. Integrity

How will you quality assure your data?

Two types of data will be collected during this project.

The first is daily diary written by the participants for 14 days of illness on the online web portal. The project team working under the supervision of the PI, Dr. Osman will ensure that each participant must update his/her health condition on daily diary. In case, any participant forgets to update the daily diary, the project team will contact him/her through email/text message to give them a reminder. The authorized members of project team will routinely assess the integrity of data entered by the participants and guide them in case of any mistake in health-related entries.

The second type of data is in the form of questionnaire that the participant will fill on each of the 14 days on the online web portal. The questionnaire will be close ended and proper form validation will be applied on each question to maintain the integrity and correctness of the entered information

4. Confidentiality

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

Local Ethical Clearance has been obtained after getting approval from the sponsor.
All human participants in this study will be informed about the data being collected, the purpose of the study and that it might be shared with other researchers after anonymization. Their data will only be recorded and preserved if they give their informed consent. We will give a unique identifier to each human participant.

We will collect several personal information as listed in Section 1. This data will be stored in password protected files to ensure their privacy. The reasons why we collect this information is listed below:

- Name and Address: For correspondence and if needed for getting some of the participants tested. Address will also be needed to assess the locality of infection, and nearest diagnostic laboratory
- Date of Birth: This will be used to calculate their age which would later be used to check if their clinical response to HSNIG is age related
- Sex: Their gender will be recorded to see if males and females respond differently to HSNIG. It is a very important parameter since males are affected by CoV2 virus more than females.
- Phone number: For communicating with the participants through calls, and SMS text messages, if needed.
- Email: Will be used for sending instructions etc.

GP / Physician Address: May be used if further information is needed to get further information about the clinical symptoms, treatment etc. (only if the participant agrees and permits). Once all analysis is completed and project finishes, all personal data listed above including age and gender will be removed from the data. At this stage, the participants will be identified only through an anonymized identifier. Hence, the data will not contain any personally identifiable information, however, we will know that when two data entries concern the same person. As the datasets will not contain any personal identifiable information, these will be deposited into Edinburgh DataShare.

5. Retention and Preservation

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

All data mentioned above will be stored indefinitely after anonymization, i.e., we will remove any personally identifiable data (Name, address, email, phone, physician/GP, age and gender etc.). We would like to preserve anonymized HSING data for many years so that other research projects may also benefit from it.

How will the data be preserved?

Since the shared data will not contain any personal information (including age, gender) it will be stored at DataShare to make it widely available for the scientific community.

6. Sharing and Publication

Which data will be shared and how?

Only anonymized, daily diary and questionnaire data will be shared for research purposes to the authorized researchers. The anonymized patient's data will be shared by the Data management team of Dr. Aimal Rextin. Since, the data will not contain any personal information of the participants, so the data and metadata will be made available through DataShare.

Are any restrictions on data sharing required?

There will be no restrictions on the anonymized data and the metadata that will be available on DataShare.

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.