Royal Papworth Hospital
NHS Foundation Trust

VOICE-CF: Vocal Biomarker Analysis for Health Condition Detection in People with Cystic Fibrosis (CF)

PROTOCOL

DATE 19 March 2021

CHIEF INVESTIGATOR

Andres Floto (CI)
Professor of Respiratory Biology, University of Cambridge
Research Director and Honorary Consultant Physician, Cambridge Centre for Lung Infection
Royal Papworth Hospital NHS Foundation Trust
Papworth Road, Cambridge Biomedical Campus, Cambridge CB2 0AY
Tel: 01223 – 638000
Email: arf27@cam.ac.uk

Investigation Sites:	Royal Papworth Hospital, United Kingdom
Protocol Contributors:	Dr Charles Haworth, Royal Papworth NHS Foundation Hospital Judy Ryan, BSC, University of Cambridge Sonde Health 30 Newbury St, 3rd floor Boston, MA. 02116
Protocol Identification Number:	PO2706
Date of Protocol:	19 March 2021
Protocol Version Number:	3.0
Trial Sponsor:	Royal Papworth Hospital NHS Foundation Trust Papworth Road Cambridge Biomedical Campus Cambridge CB2 0AY

Contents

TRIAL OVERVIEW

TRIAL OVERVIEW

1. Trial Overview

Primary Research Question

Study Design

This is a two-year single-centre, observational, prospective cohort study. We aim to recruit initially 100 and then, following interim analysis, up to 300 adult males and females with CF, who are already participating in the Project Breathe home monitoring service improvement project at Royal Papworth Hospital. The study is designed to run as a non-disruptive study with no impact on routine clinical practice. Participants will be enrolled for 12 months.

Primary Outcome

1) To determine whether vocal biomarkers can contribute to existing predictive algorithms for pulmonary exacerbations
2) To determine whether vocal biomarkers can substitute for other data collection methods, in particular spirometry and pulse oximetry.

Rationale

The aim of this study is to explore the potential of vocal biomarkers to assist in or improve the detection of health conditions in people with CF. The emphasis is on identifying onset and progression of pulmonary exacerbations.

Using their smartphone, participants submit daily voice samples that will be analysed for markers of various health conditions using labelled training data (physiological, activity and self-reported data) that is also collected on a daily basis. The Sonde voice sample collection and other home monitoring recordings will happen sequentially. Participants will continue with voice sample collection if admitted for a pulmonary exacerbation if they wish to.

Sonde Health is a health technology company with a platform that provides voice-enabled health detection and monitoring capabilities in brain, respiratory and muscular impairment. Analysis of vocal biomarkers can provide early insight into one's health and wellness. Sonde's API platform enables most companies and developers with Android or iOS mobile applications to integrate health condition detection and monitoring capabilities powered by voice.

Royal Papworth Hospital has demonstrated that daily recordings of a range of physiological parameters (e.g. spirometry, pulse oximetry, activity, heart rate, etc.) in combination with self-reported survey responses (e.g. general wellness, cough, sleep quality) in people with CF holds promise to more accurately track fluctuations in health condition, in particular the start of pulmonary exacerbations (PEx). These data are collected remotely at the study participant's home using remote/ virtual online tools, as part of a larger effort to take a more virtual approach to clinics. Smartphone-based collection of voice samples for vocal biomarker analysis is a natural extension of this approach.

We propose a phased approach that starts by enrolling 100 participants, who would provide daily voice recording on the Sonde app, in addition to continuing their ongoing daily data recordings described above.

Following interim analysis at 3 months from starting the study, we will then decide whether to expand to include up to 300 participants.

2. Patient Recruitment Criteria

Study Population

- Initially 100 and then a further 200 adult male and female CF patients
- 12 - month enrollment period

Inclusion

- Diagnosis of Cystic Fibrosis based on genetic testing and /or sweat chloride levels.
- Age between >/= 18 to </= 65 years of age at time of consent
- Able to provide written informed consent
- Patients who are currently undertaking home monitoring / virtual clinics.
- Patients who have use of a smartphone

Exclusion

- Patients unable to provide written informed consent
- Patients who are currently not undertaking home monitoring / virtual clinics
- Lung transplant recipients.

Recruitment

- The planned recruitment target rate is:

  - Year 1: Initially *100 and then a further 200 patients after interim analysis.*
  - The recruitment rate has been calculated based on previous similar studies and recognized recruitment rates for this institution.

Section 3. Sample Size and Data analysis

Sample Size

Based on previous studies using vocal biomarkers to assess respiratory disorders (Stasak et al, 2020, submitted to IEEE ICHI on COVID-19 identification from voice; additional unpublished internal data gathered by Sonde Health on 3,000+ patients with asthma, COPD, and CHF), we expect to be able to correlate voice signals with other physiological signals with 20-30 participants. We would like to recruit between 100 and 300 participants to examine whether signal changes in the Sonde app provide useful additional information of impending acute pulmonary exacerbations.

Statistical Analysis

- All data collection i.e. voice, physiological and self-reported data will be jointly analyzed using a variety of statistical and machine learning techniques. The Sonde App data will be linked to other study data using the existing participant study ID, which participants will enter in the Sonde App. No identifiable information will be shared with Sonde.

4. Consent Process and Visit Schedule

Visit

- Patients will be approached for consent in one of the following ways (1) whilst attending their normal CF clinic outpatient appointment with hardcopy consent (2) a telephone call & discussion with documented verbal consent (3) email introduction with follow-up telephone call and email consent or (4) whilst an in-patient on the ward with hardcopy consent.

Daily Samples

- Each patient will collect daily short voice samples on the Sonde app using their personal smartphone device (iOS or Android) along with their other normal daily home monitoring data
- Each daily voice sample will take approximately one minute to complete.
- You will submit 2 identical voice samples daily during the study.
- 1)There will be two identical voice sample recordings
  - a 6 second held vowel “ahhh” which will be done twice.

Informed Consent Procedure

- Patients will receive written and verbal information about the trial
- Patients interested in participating in the study will be provided with a study-specific patient information sheet and given time to consider the study. A member of the research team will contact the patient to discuss the study further and possible participation in the study.
- Written or verbal informed consent (which can be provided in the form of email or telephone) will be obtained by a member of the trial team after a suitable time has elapsed during which the patient has had ample time to read the information sheet, consider the trial and ask any questions. The Investigator must explain to each patient the nature of the study, its purpose, the procedures involved, the expected duration, the potential risks and benefits involved and any discomfort it may entail (*Ref. International Conference of Harmonisation of Good Clinical Practice (ICH/GCP)5 4.8.7*
- The ultimate responsibility for obtaining written or verbal or electronic informed consent lies with the Investigator, but this responsibility may be delegated to a suitably trained, and experienced person.
- Prior to the patient’s participation in the study, the written informed consent form must be signed and personally dated by the patient and by the clinician who conducted the informed consent discussion (*Ref. ICH/GCP 4.8.8*). Each box at the end of each statement on the consent form must be signed by the patient.
- Each subject must be informed that participation in the study is voluntary and that he/she may withdraw from the study at any time and that withdrawal of consent will not affect his/her subsequent medical treatment (*Ref. ICH/GCP 4.8.10*).
- A copy of the informed consent document will be given to the patient (*Ref. ICH/GCP 4.8.11*). One copy will be filed in the patient’s hospital case notes and a copy filed in the Trial Management File by the Clinical Research Nurses.
- The Clinical Research Nurse will ensure that the patient does not have any trial specific procedures prior to giving informed consent.
- If a patient withdraws their consent a line should be drawn diagonally through the consent form and labelled ‘consent withdrawn’ and signed and dated by a member of staff.

- The informed consent is also presented within the Sonde App before study participants can start their daily study activities (one-time after starting the study). Participants indicate agreement by checking an electronic box. The app also allows participants to review the consent information at any time via the app menus.

5. Follow-up Visits and Travel Expenses

Follow up Visits

There are no formal follow up visits. All patient data will be monitored by the research team to ensure data collection is occurring on a daily basis. If there are any problems with the data, patients will be contacted either via email or telephone to offer help and guidance to the patient if needed.

Patients will also have access to support from the research team via telephone / email if they are having problems with data collection.

Travel Expenses

There will be no compensation or inconvenience payments given for participating in the project.

6. Data Collection

Data Collection Form Completion

- The consent form and other project documents will be stored in the research office in a locked cabinet.
- All participant data will be linked-anonymized and stored on a secure password protected computer.
- All consenting participants will be registered on the study's registration log which will be filed in the study's site file. The registration log will include the participants name, date of birth and study ID.

Interim Analysis

An interim analysis will be conducted at the 1-year point after study start to further inform on how many patients will be needed to complete the study. We will file a protocol amendment with the ethics committee should this number differ significantly from that in the current protocol.

The current estimate of approximately 300 participants rests on assumptions detailed below. These assumptions will be refined at the interim analysis point with the data acquired up to that point, so that the powering can be assessed with greater confidence.

To determine the ability for the vocal biomarkers to assist in the prediction/ early detection of pulmonary exacerbations in people with cystic fibrosis, it is helpful to consider this performance in terms of sensitivity and specificity of the vocal biomarker data with respect to observed pulmonary exacerbations during the study. Other studies using these biomarkers in respiratory conditions have indicated that approx. 60% sensitivity and 60% specificity could be used as an approximate and somewhat conservative performance level. Sensitivity and specificity calculated from the collected data will be used to calculate odds ratio (2.25 at the given sensitivity and specificity) with a 95% confidence interval and p-value according to customary statistical methods. At the stated assumption around performance, about 60 pulmonary exacerbations would need to be observed and examined to have a p-value at 0.05.

From clinical experience in the CF clinic, it can be estimated that following roughly 100 patients with CF over a 12-month period will result in approximately 60 pulmonary exacerbations. Vocal biomarker models are gender dependent, so to assess this level of significance in both genders would require a total of 200 patient participants. Accounting for an as yet unknown degree of missing data and drop-outs from the study would further increase the sample size estimate to somewhere between 200-300.

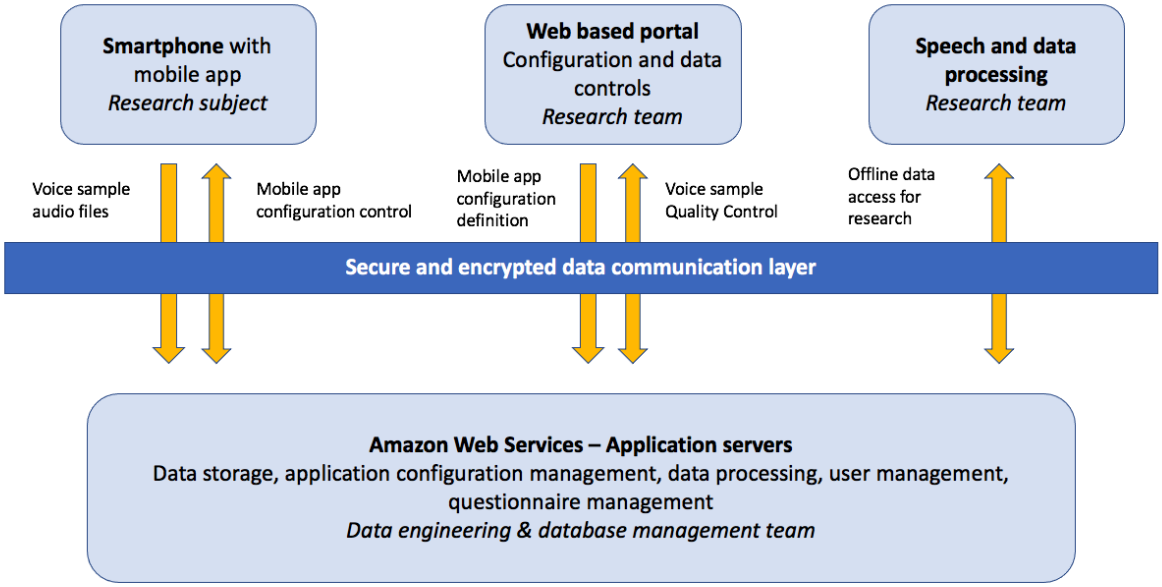
At the 12-month interim analysis points all figures used in these calculations can be refined from the observations made up to that point.

Data Access / Sharing

- The Sonde health app will collect linked anonymized data. All data will be linked-anonymized and uploaded to Sonde's cloud-based server systems, which meet HIPAA and GDPR requirements. Sonde uses Amazon Web Services (AWS) for all data collection and app deployment. Documentation is available upon request. Collected data is transferred through secure (HTTPS) Protocol and is stored in encrypted form in AWS.
- The home monitoring equipment linked-anonymized data is already being uploaded to a secure NHS approved web-based site to allow researchers within the University of Cambridge to analyse it via Microsoft Azure which implements the transmission integrity and confidentiality control by ensuring that cryptography is implemented through a hybrid model. The following is a high-level list of the symmetric and asymmetric keys used for encrypting and protecting confidentiality of data.
    - Use AES for symmetric encryption/decryption
    - Use 128-bit or better symmetric keys
    - Use RSA for asymmetric encryption/decryption and signatures
    - Use 2048-bit or better RSA keys
    - Use SHA-256 or better (SHA-384, SHA-512) for hashing and message- authentication codes.

- Sharing of data between Royal Papworth Hospital and Sonde is as noted below:
    - Sonde will share with Royal Papworth, vocal biomarkers including voice feature level information that enables health detection research, with some limitations to protect Sonde IP and metadata collected on the Sonde app.
    - Royal Papworth will share with Sonde daily home monitoring data, participant demographics and basic medical history / comorbidities.
    - Data sharing between Sonde and Royal Papworth is planned during each 3-month increment and upon study completion.

Data Flow Chart



Storage of Documents

- There will be minimal storage of documents. The only documents that will be stored in paper form will be the consent forms and any letters of approvals or other paper related documents which will be kept in a locked filing cabinet in a locked room.

Retention of Documents

- All study documentation will be stored for 15 years after the last patient has completed their last study visit.
- Documentation will be archived in a fire safe, secure Sponsor approved archive facility.

Monitoring and Audit

- All study-related records will be made available upon request of the monitor, auditor, Sponsor, R&D, REC, MHRA or other regulatory authority.
- The project data will be monitored by the Research Officer quarterly. The first five patients along with the site and sponsor files will be monitored. Monitoring will be increased if any issues are found.

Section 7. Adverse and Serious Adverse Events

The definition of an adverse event is: 'Any untoward medical occurrence in a patient which does not necessarily have a causal relationship with this treatment'. This includes 'any unfavourable and unintended sign (including an abnormal laboratory finding), symptom or disease temporally associated with the study drug'. This may include, for example, a cold or an accident.

The definition of a serious adverse event is one that fulfils at least one of the following criteria:

- Is fatal- results in death
- Is life threatening

- Requires inpatient hospitalization or prolongation of existing hospitalization
- Results in persistent or significant disability/incapacity

OR

- Is a congenital anomaly/birth defect

The definition of a suspected unexpected serious adverse reaction (SUSAR) is a serious adverse event that is thought to be possibly or definitely related to the study drug.

Recording and Reporting

- As there is no therapeutic intervention involved in this study, we do not anticipate any risks
- We do not expect to have any AEs or SAEs
- All research staff in contact with patients are responsible for noting adverse events that are reported by the patient and making them known to the Principal Investigator.
- At each visit or study assessment, adverse events that have occurred since the previous visit should be elicited from the patient. The event should be detailed in the patient's notes, as source document verification, including the start date (if known) and the end date.
- The action taken regarding the study procedure should be documented
- Document any treatment/medication given for the event, including the dates the treatment/medication was commenced and the date it was stopped/changed, if applicable (an example of an AE form is in section 10 of the site file section). Documenting of adverse events is the responsibility of the Principal Investigator, Co-Investigator and Clinical Research Nurse/Assistant
- Events, which are ongoing at the final study visit, should be followed up as clinically indicated.

- All serious adverse events (SAEs) and suspected unexpected serious adverse reactions (SUSARS) will be documented as above using the appropriate reporting documentation in section 10 of the site file documents.
- SUSARS must be reported to the Sponsor within 48 hours of the Chief / Principle Investigator/Clinical Research Nurse being aware.
- All Unexpected SAEs and SUSARs should be reported to the hospital R&D department and may be require analysis through the hospital incident reporting system.
- SUSAR reports to be submitted to the main REC (*timelines to be specified)*
- Expected SAEs need to be reported to the Main REC in the annual REC report; the REC who performed the locality assessment does not need to be informed.
- Store all completed SAE forms with the patient's study documentation.

8. Amendments

Substantial amendments will be submitted to the Research Ethics Committee for review and approval.