Bridge2AI eConsent - MIT - English

Please complete the form below.

Thank you!

UNIVERSITY of SOUTH FLORIDA
USF RESEARCH & INNOVATION

Study ID: STUDY004890_MOD000035 Date Effective: 4/10/2025

Informed Consent to Participate in Research Involving Minimal Risk

Information to Consider Before Taking Part in this Research Study
Title: Bridge2AI Voice Data Acquisition
Study # 004890

Overview: You are being asked to take part in a research study. The information in this document should help you decide if you would like to participate. The sections in this Overview provide the basic information about the study. More detailed information is provided in the remainder of the document.

Study Staff: This study is led by Yael Bensoussan, MD, who is a laryngologist at University of South Florida. This person is called the Principal Investigator (PI). In this study, there are other site-specific PIs. Other approved study staff may act on behalf of the PIs.

Study Details: This study is conducted at University of South Florida (USF), Weill Cornell Medicine (WCM), Massachusetts Institute of Technology (MIT), University of Toronto (UofT), Vanderbilt University Medical Center (VUMC), Mount Sinai Hospital (MSH), Hospitals for Sick Children (HSC), Massachusetts Eye and Ear (MEEI), Emory University, and is supported/sponsored by The National Institutes of Health (NIH).

The human voice is unique to each individual and contains audio features that have been linked to diseases such as Parkinson's, dementia, mood disorders, and certain cancers. Given the advances in artificial intelligence, there is increased interest in using voice in the diagnosis of disease. Although early results are promising, many limitations remain including the limited size of databases available, and their questionable quality and diversity. The main purpose of this study is to build a large open science database of human voices, speech and respiratory sounds, other health data, and associated metadata. This database will respect applicable best ethical practices and ensure representation of a diverse population to fuel artificial intelligence research related to voice. Some of the information collected will further be shared with the general public in full open access, in a de-identified or aggregated form.

Voice, speech, and respiratory sound data will be preserved as audio recordings. Data collected will include voice tasks such as vowel sounds, free speech, oral text reading, snoring, coughing and breathing sounds. Demographic data such as age, sex, gender, race, ethnicity, and language will be collected through an application. Different types of imaging may be collected from your medical record, such as chest x-rays, brain CT scans, and brain MRIs. No imaging will be performed in the context of this study and only imaging data produced from prior or ongoing clinical evaluations will be reviewed and collected from your chart. You may also be asked to answer surveys and/or validated questionnaires related to your health status.

You will be offered to enroll with the option to contribute with single time point data (data collected in a single session) or longitudinal data (data collected over multiple sessions). If you are enrolled as a control for this study, data collection will be for a single time point.

- For single time point data collection, voice data collection will be performed in one of two ways.
    - Data collection may occur in clinic during your regular office visit, where you will be offered to enroll in the study and escorted to a study room before or after your regular appointment. In most cases, the consent and data collection will be performed on the same day with a research assistant. In some cases, enrollment and consent may occur remotely through an application, and/or may occur at an earlier point in time than data collection. We generally expect a session to last 45 - 60 mins, but this may vary due to individual

circumstances. If you wish or request to come back or have a separate appointment for recording data we can accommodate this request, but it is not necessary.
    - Data collection may occur remotely, through access to an application.
- For longitudinal data collection, data will be collected in 2 forms:
    - 1: Longitudinal data in clinic during clinic visits:
        - This data collection will happen in clinic during your REGULAR follow-up visits. No additional visits outside of your regular follow-up visits are required. The maximum follow-up time will not exceed the study period of 4 years.
    - 2: Longitudinal data "at home":
        - In between clinic visits, you may be asked to perform voice data collection at home through an application that will be downloaded to your personal smart device with the help of our research assistant in the clinic. Depending on your condition, this data collection will be at an interval of 1-6 months and for a maximum of 4 years total duration. You can decide to leave the study at any point during that time.

Participants: You are being asked to take part in this study either because you have been diagnosed with a voice disorder/ neurological disorder/ respiratory disorder/ mood and anxiety disorder that is within the scope of interest of the study, or because your voice can serve as a control for the database. This means that your voice can be compared to the voices of those with the conditions within the scope of interest so we can better understand how to identify and distinguish them.

Voluntary Participation: Your participation is voluntary. You do not have to participate and may stop your participation at any time. There will be no penalties or loss of benefits or opportunities if you do not participate or decide to stop once you start.

Benefits, Compensation, and Risk: You will not directly benefit from participation in this study. However, we expect that the results of the study will contribute to improving disease detection, clinical care, and public health activities.

There is no financial cost to participate in this study. You will receive gift cards as compensation for your participation. Each participant will be given a $40 gift card for sessions lasting under 90 minutes, and an $80 gift card for sessions lasting over 90 minutes, for a maximum of three sessions and $120.

This research is considered minimal risk. The risks of participation are the same as the risks you face in daily life. Nonetheless, certain residual risks do remain. There remains a low risk that either the raw unprocessed study data held at the original research institution, or the coded study data shared with other researchers, could be subject to a data breach, or could otherwise be associated to your identity. It is also possible that the de-identified or aggregated data shared with the general public could be re-identified in the future.

Confidentiality: Your raw and unprocessed study data, which is linked to your identity, will be held at the original research institution that collects the data. This data will only be made available to the Principal Investigator, and to select other members of the original study team, except in certain exceptional circumstances described in the "Privacy and Confidentiality" section below.

The research data that is to be held in a central, secure database, and subsequently shared with external researchers, will be double-coded (direct identifiers will be replaced with a unique alphanumeric code). This coded research data will be stored for an indefinite duration on cloud servers or another equivalent technological medium. It will be shared with researchers from all categories of organizations for commercial or non-commercial research uses. It can be shared domestically or internationally. Only the principal investigators and select individuals at the original research institution will have access to the linkage logs that can be used to associate your identity to the coded study data.

Data that poses a negligible risk of causing individual re-identification will be shared in full open access with the general public.

Why are you being asked to take part?

You are being asked to take part in this study because you have been diagnosed with a voice disorder/ neurological disorder/ respiratory disorder/ mood and anxiety disorder with known effects on voice and speech, or because your voice can serve as a control for the database. The PIs and study team are trying to see whether voice has biomarkers for particular disease categories that could be screened and differentiated through AI and machine-learning algorithms by building a large, ethically-sourced, and diverse database of related voice recordings. Researchers from around the world will be able to access the data that is held in this database to perform new research using your data. Some data that poses little or no privacy risks, such as data that has been aggregated with that of other people or stringently de-identified, will be shared with the general public.

Study Procedures

To determine your eligibility for participation in this study, authorized clinicians will review your medical record to screen for inclusion and exclusion criteria. Once you are deemed eligible, we will collect the following information from your health record: demographic data such as age, sex, gender, race, ethnicity, and language; clinical data related to your health status such as disease type, severity, and symptoms (except for the control group); and imaging data produced from prior or ongoing clinical evaluations such as chest x-ray for the respiratory disease group and brain MRI and CT scans for the neurology group. Please note that no imaging will be performed in the context of this study. Only imaging data produced from prior or ongoing clinical evaluations will be reviewed and collected from your chart. For some neurological conditions, clinical data that refers to genetic tests may be included. For example, the CAG repeat length may be included to assess the confirmation of Huntington's disease through genetic testing. No genetic samples will be taken for this study.

If eligible, you will be offered to enroll in the study. The process of consent and collection of voice, speech, and respiratory sound data will usually take from 45 to 60 minutes, depending on individual circumstances and the vocal tasks linked to your disease category.

You will be offered to enroll with the option to contribute single time point data (data collected at a single session only) or longitudinal data (data collected several times over a period of time).

Single time point data collection will be performed in the clinic during the participants' regular office visit, or in some cases remotely through an application. Longitudinal data may be collected in 2 ways: during your regular clinic visits and/or "at home" through a smart device application downloaded to your personal device in the clinic. Depending on your group (disease cohort or control), your data may be

collected at intervals of 1 to 6 months and for a maximum of 4 years total duration. You can decide to leave the study at any point during that time. If you wish to schedule a separate appointment for collecting data (outside your regular clinic visit), we can accommodate this request, but it is not necessary.

Total Number of Participants

About *5 000* individuals will take part in this study at USF. A total of *30 000* individuals will participate in the study at all sites.

Alternatives / Voluntary Participation / Withdrawal

You do not have to participate in this research study.

You should only take part in this study if you want to volunteer. You should not feel that there is any pressure to take part in the study. You are free to participate in this research or withdraw at any time. There will be no penalty or loss of benefits you are entitled to receive if you stop taking part in this study. If participants are students or employees, the decision to participate or not to participate will not affect your student status (course grade) or job status.

This project's main purpose is to build an open science database enabling researchers to access the data for future research purposes. The research data held in that database, including voice recordings, demographic data, clinical and imaging data, questionnaire data, and associated metadata will be stripped of direct identifiers and coded and be uploaded to a secure database. If you withdraw from the study after they complete the voice data collection, your voice recording, and other research data cannot be removed from the database. If you withdraw during or before the voice data collection, your data will not be included in the database. For longitudinal data collection, you can decide to withdraw from future instances of longitudinal data collection at any point. In that case, only the data collected prior to your withdrawal will be kept in the database. The minimal requirement for participation is 1 data collection time point.

Benefits

There are no direct benefits to participating in the study. However, we expect that the results of our study will contribute to improving disease detection through voice, with positive outcomes for clinical care, public health, and health research.

Risks or Discomfort

Participation in this study involves the collection of: your voice, speech, and respiratory sound recordings; demographic, clinical, and imaging data from your existing medical record; and survey/questionnaire responses. Some survey questions may make you feel uncomfortable. Know that the information that you provide is entirely up to you, and that you are free to skip questions that you do not want to answer. If you feel you have suffered an injury, which may include emotional trauma, as a result of participating in this study, please contact the PI of this study as soon as possible.

Voice data collection is a safe non-invasive collection method that poses minimal physical and psychological risk. Privacy risks associated with this study include the risk of your personal information being mistakenly released and/or used for unconsented purposes, including to re-identify you. The

confidentiality of your personal information will be protected, respecting the applicable regulatory requirement(s) and applicable privacy and confidentiality rules.

- We take great care to protect your information and identity. Raw study data that can be used to directly identify you will be separated from the coded research data that is intended to be shared with external researchers (voice recordings, de-identified survey answers, clinical and imaging data, and associated metadata) during collection itself to reduce the risk of re-identification as much as possible. However, please note that there is always a low risk that someone may attempt to re-identify you by comparing your coded research data to other sources of information (e.g., using information from your social media or from public registers).
- Voice is unique to each individual. Although we will ensure that no directly identifiable information is linked to your voice recording, there is still a minimal risk of re-identifying you using your recording. Additionally, voice changes with age, speaking style (for example, scripted or spontaneous), and with different emotions. This means that in a few years, your voice may no longer be a 100% match to the recordings collected for this study which would further prevent your identification.
  Accidental public disclosure of the raw study data or the coded research data could occur due to unintended data breaches occasioned through hacking, researcher misconduct, or other activities outside of the procedures authorized by the study. In such a case, your data may be misused or used for unauthorized purposes.

Participation in this study may involve additional risks that are not known at this time. For example, more advanced technologies capable of identifying you based on your voice recordings or coded research data could be created in the future.

Compensation

Participants will receive compensation for taking part in this study in the form of gift cards.
Participants will be given a $40 gift card for sessions lasting under 90 minutes, and an $80 gift card for sessions lasting over 90 minutes, for a maximum of three sessions and $120.

Costs

There is no financial cost to participate.

Privacy and Confidentiality

This section of the consent form describes the privacy and confidentiality of the raw, unprocessed study data that you contribute to the USF research team, as well as that of the direct identifiers linked to that data. The sharing of coded research data with third parties, after it has been stripped of your direct identifiers, is described in the section below entitled "Sharing of Data for Future Research Purposes."

We will maintain best practices and state-of-the-art safeguards to keep data that is collected from you as part of this research project confidential. We cannot guarantee absolute confidentiality. Your personal information may be disclosed if required by law. Certain people may need to see your study records. These individuals include:

- The study team, including the PIs, collaborating researchers at USF and other universities, study coordinators, study nurses, and all other study staff.
- Certain government and university people who need to know more about the study. For example, individuals who provide oversight on this study may need to look at your records. This is done to make sure that we are doing the study in the right way. They also need to make sure that we are protecting your rights and your safety.
- Any agency of the federal, state, or local government that regulates this study. This includes the Department of Health and Human Services (DHHS) and the Office for Human Research Protection (OHRP).
- The USF Institutional Review Board (IRB) and its related staff who have oversight responsibilities for this study, and staff in USF Research Integrity and Compliance.
- The sponsor for this study, which is the National Institutes of Health (NIH)

To help us protect your privacy, we have obtained a Certificate of Confidentiality from the National Institutes of Health. With this Certificate, the researchers cannot be forced to disclose information that may identify you, even by a court subpoena, in any federal, state, or local civil, criminal, administrative, legislative, or other proceedings. The researchers will use the Certificate to resist any demands for information that would identify you, except as explained below.

The Certificate cannot be used to resist a demand for information from personnel of the United States Government that is used for auditing or evaluation of federally funded projects or for information that must be disclosed in order to meet the requirements of the federal Food and Drug Administration (FDA).

You should understand that a Certificate of Confidentiality does not prevent you or a member of your family from voluntarily releasing information about yourself or your involvement in this research. If an insurer, employer, or other person obtains your written consent to receive research information, then the researchers may not use the Certificate to withhold that information.

The Certificate of Confidentiality does not prevent the researchers from disclosing voluntarily, without your consent, information that would identify you as a participant in the research project under certain circumstances. The investigative team will voluntarily comply with Florida Statutes and federal regulations, which may mandate or permit certain disclosures of protected information by the investigative team to appropriate individuals.

Per USF policy, the original copies of your voice recording will be retained for 5 years after the completion of this study. However, we will plan to keep copies of your voice recording and other research data (stripped of any directly identifying information) for an indefinite period of time.

Sharing of Data for Future Research Purposes

This section of the consent form describes the conditions according to which your coded study data, having been stripped of direct identifiers, will be shared with external researchers for the purpose of performing further studies using that data.

Research data that does not contain your direct identifiers will be shared with external researchers for future research through a secure database. Data that poses a negligible risk of causing individual re-identification will be shared in full open access with the general public. Data that would pose a

heightened risk of re-identification if shared in full open access will be shared through a controlled access mechanism with authorized researchers. This includes certain voice data recordings, demographic information, clinical and imaging data, and associated metadata. This data will be stored for an indefinite duration on cloud servers or another equivalent technological medium. It will be shared with researchers from all categories of organizations for commercial or non-commercial research uses. It can be shared domestically or internationally.

To access the controlled-access research data, researchers outside of the study team will need to apply for data access. This requires them to complete an application form and describe their intended use of the research data. It also requires them to agree to respect a number of conditions relating to data use, including commitments not to release the data to third parties without prior authorization, and to not engage in acts that could cause harm to research participants (e.g., re-identification or the use of data to stigmatize).

To this end, the data that you contribute to the study will be double-coded (your identifiers will be replaced with a unique alphanumeric code), and the linkage logs will be kept only by a small number of the members of the original research team.

Publications resulting from the study will not contain any identifiable information. Anyone with the authority to look at your information must keep them secure and not share them with unauthorized third parties. The confidentiality of personal information collected will be protected, respecting the privacy and confidentiality rules in accordance with the applicable regulatory requirement(s).

What if new information becomes available about the study? Will you tell me what my study data says about my present and future health status?

The primary goal of this study is not to generate or report clinically relevant findings based on your data. However, in certain instances, we may wish to contact you to return results or data. For example, new information could affect your willingness to participate in this study. This study may also provide results or uncover incidental findings that could have implications for your health. If you indicate that you wish to be notified of such findings at the end of this consent form, we may contact you if such information becomes available. However, there is no systematic program in place to return results or incidental findings to participants.

If you believe you are having symptoms that may require care, you should contact your primary care physician.

You can get the answers to your questions, concerns, or complaints.

If you have any questions, concerns or complaints about this study, call Yael Bensoussan, M.D. at ((813) 396-0581. If you have questions about your rights, complaints, or issues as a person taking part in this study, call the USF IRB at (813) 974-5638 or contact by email at RSCH-IRB@usf.edu.

Consent to Take Part in Research

I freely give my consent to take part in this study. I understand that by signing this form I am agreeing to take part in research. I have received a copy of this form to take with me. I understand that by signing below, I am consenting for my contact information to be kept in a repository for this study, and that this information may be used to contact me in the future for further data collection if long-term follow-up is required as part of an eventual extension of this grant after the expected 4-year duration, or to be invited to participate in future research studies. My information will not be shared with third parties.

I consent to having all data I submit shared with the researchers of this study and to having my audio recordings, written responses, and other coded research data shared with other qualified researchers as described in this consent form. I further consent to the public sharing of my de-identified and aggregated study data.

I would like to be informed of incidental findings, should the study report this type of information.

- ☐ Yes
- ☐ No

_____ _____

Signature of Person Taking Part in Study
[Authorization]

Date

_____

Printed Name of Person Taking Part in Study

Consent to Take Part in Research

I freely give my consent to take part in this study. I understand that by signing this form I am agreeing to take part in research. I have received a copy of this form to take with me. I understand that by signing below, I am consenting for my contact information to be kept in a repository for this study, and that this information may be used to contact me in the future for further data collection if long-term follow-up is required as part of an eventual extension of this grant after the expected 4-year duration, or to be invited to participate in future research studies. My information will not be shared with third parties.

I consent to having all data I submit shared with the researchers of this study and to having my audio recordings, written responses, and other coded research data shared with other qualified researchers as described in this consent form. I further consent to the public sharing of my de-identified and aggregated study data.

I would like to be informed of incidental findings, should the study report this type of information.

◯ Yes
◯ No

Signature of Person Taking Part in Study
[Authorization]

_____

Date

_____

Printed Name of Person Taking Part in Study

_____

Site Specific Information and Authorization to Collect, Use, and Share your Health Information

Title: Bridge2AI Voice Data Acquisition

Site # ___ SITE000024 _____

Site Name: MIT

Local PI/ Phone number: Satrajit Ghosh / +16173243544

Overview: The information contained in this document is additional information specific to the research site where you are enrolling. None of the information within this document supersedes the information in the attached overall consent. This document provides you with additional information about the local site where you're enrolling in the research. If you have any concerns regarding information within this document please contact the local Principle Investigator, whose contact information is included above.

Emergency Care and Compensation for Injury

If you feel you have suffered an injury, which may include emotional trauma, as a result of participating in this study, please contact the person in charge of the study as soon as possible.

In the event you suffer such an injury, M.I.T. may provide itself, or arrange for the provision of, emergency transport or medical treatment, including emergency treatment and follow-up care, as needed, or reimbursement for such medical services. M.I.T. does not provide any other form of compensation for injury. In any case, neither the offer to provide medical assistance, nor the actual provision of medical services shall be considered an admission of fault or acceptance of liability. Questions regarding this policy may be directed to MIT's Insurance Office, (617) 253-2823. Your insurance carrier may be billed for the cost of emergency transport or medical treatment, if such services are determined not to be directly related to your participation in this study. (If the study is sponsored by a private drug or device manufacturer, delete the previous sentence.)

You can get the answers to your local questions, concerns, or complaints.

If you have any local questions, concerns or complaints about this study, call *Satrajit S. Ghosh* at *617-324-3544*.

You are not waiving any legal claims, rights or remedies because of your participation in this research study. If you feel you have been treated unfairly, or you have questions regarding your rights as a research subject, you may contact the Chairman of the Committee on the Use of Humans as Experimental Subjects, M.I.T., Room E25-143B, 77 Massachusetts Ave, Cambridge, MA 02139, phone 1-617-253 6787.