Voice Biomarkers predictive of Depression and Anxiety: an Observational study

Principal Investigator:
Mei-Hsin Cheng, PhD
2737 Garber Street, No. 8
Berkeley, CA 94705

Version 1.2
1 September 2021

Statement of Compliance

The study will be conducted in accordance with the International Conference on Harmonisation guidelines for Good Clinical Practice (ICH E6), the Code of Federal Regulations on the Protection of Human Subjects (45 CFR Part 46). All personnel involved in the conduct of this study have completed human subject's protection training.

SIGNATURE PAGE

The signature below constitutes the approval of this protocol and the attachments, and provides the necessary assurances that this trial will be conducted according to all stipulations of the protocol, including all statements regarding confidentiality, and according to local legal and regulatory requirements and applicable US federal regulations and ICH guidelines.

Principal Investigator:

Signed: _____ Date: _____

Mei-Hsin Cheng, PhD
Director of Product

Table of Contents

Abbreviations

AI	Artificial Intelligence
API	Application Programming Interface
CBT	Cognitive Behavioral Therapy
CFR	Code of Federal Regulations
DSM-5	Diagnostic and Statistical Manual, 5th edition
GAD-7	General Anxiety Disorder-7
GCP	Good Clinical Practice
ICF	Informed Consent Form
ICH	International Conference on Harmonisation
IRB	Institutional Review Board
JSON	JavaScript Object Notation
KiVA™	Kintsugi Voice Biomarker API
MDD	Major Depressive Disorder
MFC	Mel-frequency cepstral
MFCC	Mel-frequency cepstral coefficient
MS	Multiple Sclerosis
NER	Named Entity Recognition
PHQ-9	Patient Health Questionnaire-9
PI	Principal Investigator
PTSD	Post-traumatic stress disorder
TBI	Traumatic brain injury

1. KEY ROLES

Principal Investigator: Mei-Hsin Cheng, PhD, Kintsugi Mindful Wellness, Inc.
2737 Garber Street, Ste. 8
Berkeley, CA 94705
mace@kintsugihello.com

Research Coordinator: Alice Kim, Kintsugi Mindful Wellness, Inc.
2737 Garber Street, Ste. 8
Berkeley, CA 94705
alice@kintsugihello.com

Research Coordinator: James Heppner, Kintsugi Mindful Wellness, Inc.
2737 Garber Street, Ste. 8
Berkeley, CA 94705
jim@kintsugihello.com

Research Coordinator: Amanda Edwards, Kintsugi Mindful Wellness, Inc.
2737 Garber Street, Ste. 8
Berkeley, CA 94705
amanda@kintsugihello.com

2. BACKGROUND INFORMATION AND SCIENTIFIC RATIONALE

1. Background Information

With the rise of the COVID-19 global pandemic, society has fundamentally changed. The delivery of healthcare services has shifted dramatically to telehealth from in-person care, and there has been a dramatic jump in the incidence of mental distress in the U.S. (**1**)

Major Depressive Disorder (MDD) is the leading cause of disability worldwide. Depression and anxiety disorders are among the most prevalent of all mental disorders, with an estimated annual prevalence of 9.7% and 18.1% respectively (**2**).

Kintsugi, a digital health technology company based in Berkeley and founded in 2019, has developed a consumer application called Kintsugi that is available on the Apple App Store. This app allows users to journal by speaking, and interact with community members who share similar challenges to promote mental wellbeing. Based upon their preliminary work and user engagement, Kintsugi team has been awarded several National Science Foundation SBIR grants to develop new Artificial Intelligence (AI) technologies in healthcare. Most recently, the team has developed an AI-based algorithm that is able to detect depression and anxiety from short clips of free-form speech. Kintsugi Voice Biomarker API (KiVA™) is the product based on this algorithm.

AI has great potential to augment healthcare in the form of clinical decision support tools for clinicians or as a standalone autonomous solution.

Voice biomarker technologies provide a unique way to scale access to mental healthcare by assisting clinicians with triage especially in high-volume environments where administrative burdens are high. Researchers have attempted to find objective methods to increase the accuracy of depression and generalized anxiety disorder diagnoses, and voice-based biomarkers are attractive because of their non-invasive nature and ability to provide passive monitoring.

2. Scientific Rationale

Voice biomarkers detect a variety of health conditions, emotions, and diseases and provide non-invasive ways to collect patient feedback. Transforming voice intonations into voice biomarkers would allow us to predict the existence of disease, monitor progression or deterioration of disease or chronic condition, predict hospitalization and mortality, focus on patients in need, and as a result, optimize health care outcomes and cost.

It has been known for the last 100 years that depression and anxiety both likely affect vocal acoustic properties. In 1921, Emil Kraepelin, characterized depressed patient's voices as having a lower pitch, lower volume, lower rate of speech, more monotony of prosody as well as more hesitations, stuttering, and whispering **(3)**. Various groups of researchers, such as Zhang et al (2020), Espinola et al. (2020), Wang et al (2019), Smiljanic et al. (2019) have tried to

characterize the most salient vocal acoustic (paralinguistic) features in depression, to develop a possible biomarker for the condition. **(4-11)**

Mechanistically, it is possible that the neural circuitry involved in the pathophysiology of mood and anxiety disorders impinge upon the neural circuit involved in speech production, affecting qualities that include rate, prosody, speech latency and other paralinguistic features. Thus, acoustic features of speech may be one of the more readily accessible biomarkers for these conditions.

Given this understanding, we sought to develop a passive vocal biomarker instrument for depression and anxiety screening that could markedly expand access as well as standardize the quality of screening in primary care settings.

Numerous research groups have studied human voice and its acoustic features for its suggested correlation with psychiatric disorders. **(12-17)** General understanding in the field of vocal analysis about vocal features that construct perceptually relevant amplitude and frequency representations led us to use mel-frequency cepstral coefficients (MFCCs) for our experiments. Mel-frequency cepstral coefficients (MFCCs) are coefficients that collectively make up a Mel-frequency cepstral (MFC). MFCCs in turn are a representation of the short-term power spectrum of a sound. **(18)** Features that we extract represent phonemes (distinct units of sound) as the shape of the vocal tract. Sounds generated by humans are filtered by the shape of the vocal tract that includes tongue, teeth, lips, throat, et cetera and determines the sounds produced. In addition to MFCCs, other characteristics such as fundamental frequency, frequency variability, shimmer, jitter and pause length are also considered in our modeling efforts.

2.3 Potential Risks and Benefits

2.3.1 Potential Risks

Potential risks associated with this study include:

- Breach of confidentiality of personal information
- Participation may bring up emotionally charged content which could temporarily impact participants' mood

The probability of breach of confidentiality will be minimized by storing all data in a secure, password protected Google Cloud Platform storage. Researchers will also be using Named Entity Recognition (NER) and number detection scripts to strip out any person names, organizations, locations, medical codes, quantities, medical record numbers, monetary values, and percentages. The PI and the Research Coordinators will be the only people with access to the Google Cloud Platform bucket. After the study concludes, participants' First and Last name, email and phone number will be deleted.

The probability of discomfort related to probing for emotionally charged content will be minimized by the Research Coordinator reminding the subjects to answer the questions at their own comfort level. All the participants will be emailed a list of resources before

the study begins. List of resources will include phone numbers of various services that can be contacted immediately if a participant experiences distress or discomfort.

2.3.2 Known Potential Benefits

While there may be no direct benefit from participation in the study, possible indirect benefits of voice journaling include:

- Reduction in anxiety
- Reduction in stress

3. OBJECTIVES

The objective of this study is to collect data to improve the sensitivity and specificity of KiVATM's predictions on audio files. A web-based audio recording tool will be used to record voice sample and ask participants the PHQ-9 and GAD-7 questions. Participants will then be assessed by mental health professionals with the Hamilton Depression Rating Scale (HAM-D) (**19, 20**) and Hamilton Anxiety Rating Scale (HAM-A) (**21**) scores.

Ultimately, the results of this study will allow us to create a technology that can scale access to mental healthcare by increasing the rates of depression and anxiety screening and routing patients to the appropriate level of care.

4. STUDY DESIGN

This is a prospective, fully remote observational e-clinical study to improve the sensitivity and specificity of KiVA™ predictions against clinician-rendered HAM-D and HAM-A scores.

The duration of the study enrollment is anticipated to continue until ~1,000 participants are enrolled. We will extend the enrollment period until the necessary number of participants is enrolled.

5. STUDY POPULATION

5.1 Selection of the Study Population

This study will enroll ~1,000 men and women over the age of 18 via social media.

5.2 Inclusion Criteria

- Adult males and or females over the age of 18 at the time of informed consent. Adults with a known history of depression and/or anxiety and adults without a known history of depression and/or anxiety will be included in this study.
- Access to a laptop, smartphone or tablet with a functioning microphone
- Stated willingness to comply with all study procedures and availability for the duration of the study
- Fluency in English

5.3 Exclusion Criteria

- Visual impairment that would make it difficult for the participant to see the web-based survey
- Motor impairment that would make it difficult for the participant to use the web-based survey

6. RECRUITMENT AND RETENTION

The study research coordinators will recruit participants from social media, such as Reddit and Facebook.

Participants will receive a link to our web-based audio survey from social media advertisement or message to complete the study steps.

Potential participants will be offered an Amazon Gift card, PayPal or Visa card of $ 5 or $10 for participating in the first part of the study. Potential participants who qualify for the second part of the study will be offered an Amazon Gift card, PayPal or Visa card of $50. The gift card will be distributed via email after the participant completes the study. Partial completion of the study requirements will not be paid.

Additionally, participants will be able to download the Kintsugi voice journaling app at no cost once they enroll in the study. The usual cost for the app is $19.99.

For the second part of the study, participants will be selected from the pool of participants that completed the first part. The Research Coordinator will email those participants who scored greater than 20 on their PHQ-9 questionnaire with an offer to participate in the second part of the study.

7. STUDY PROCEDURES/EVALUATIONS

7.1 PART I

Participants will follow the link from the social media advertisement or message in order to participate in the study.

They will review the Informed Consent Form (ICF). The ICF provides a detailed explanation of the risks and benefits of the study. By accepting the ICF, the participant agrees to enroll in the study.

We have a different informed consent form and different reward rate depending on the social media platform that the participant comes from. The reason for this is because participants from social media such as Reddit have a higher prevalence of depression than participants from social media such as Facebook.

The participant will be asked to take the online survey. He or she will answer the following questions aloud:

How are you feeling today? Please speak for 1 minute about your day, if there was anything special about it and overall, how you felt.

After the recorded section, participants will complete two self-assessments: PHQ-9 and GAD-7. The Patient Health Questionnaire and Generalized Anxiety Disorder tests have 16 questions in total. These assessments will immediately follow the recorded section.

The Research Coordinator will email a $5 Amazon Gift card, PayPal or Visa card to participants from social media such as Facebook and a $10 Amazon Gift card, PayPal or Visa card to the participants from social media such as Reddit.

Completion of this first part will take about 10-minutes.

We collect the participant email addresses because the Research Coordinator provides the gift cards via email. The information is stored on a password protected location only accessible to the PI and the research coordinators. When the study is over, the Personal Health Information (the email addresses) will be deleted.

7.2 PART II

The Researcher will review PHQ-9 scores of all the participants and select the subset that scored greater than 20. She or he will email the subset of participants about the opportunity to participate in the next part of the study.

After receiving an email where the participant expresses interest in the next part of the study, the Research Coordinator will send an email to the participants containing:

1. Instructions on how to set up an appointment with a mental health professional with appropriate experience
2. List of resources to contact if the participant were to experience distress

After participants set up an appointment through a scheduling program, they'll receive a confirmation email with a link to a web-conferencing tool. The participant communication is included as "Participant Communication for Part II of Study Procedure".

Completion of this first part will take about 15-minutes.

The participant will review the Informed Consent Form (ICF). The ICF provides a detailed explanation of the risks and benefits of the study. By accepting the ICF, the participant agrees to enroll in the next part of the study.

The participant will be asked to take the online survey. He or she will answer the following questions aloud:

1. *How are you feeling today? Please speak for 1 minute about your day, if there was anything special about it and overall, how you felt.*

2. *What is your biggest source of stress and why?*

3. *What or who are you grateful for today?*

After three recorded sections participants will complete two self-assessments: PHQ-9 and GAD-7. The Patient Health Questionnaire and Generalized Anxiety Disorder tests have 16 questions in total. These assessments will immediately follow the recorded section.

Enrollment will be on a rolling basis and last however long it takes to recruit the required number of participants.

7.3 III

Immediately after completing Part II, participants will receive a clinical assessment from a mental health professional (MHP). A web-conferencing tool link to their appointment will be provided in the confirmation email. Participants will be reminded that this is only a clinical assessment and will not involve treatment of any kind.

MHPs will fill out the Hamilton Depression Rating Scale (HAM-D) and Hamilton Anxiety Scale (HAM-A), keep a copy of the spreadsheet for their records and for subsequent data monitoring activities. Along with the spreadsheet, MHPs will send the concomitant medication log and clinical notes to the research coordinators. Some medications affect voice biomarkers and we want to be able to separate the effect of depression and anxiety vs medication from the voice biomarkers. The information is stored on a password protected location only accessible to the PI and the research coordinators. When the study is over, the Personal Health Information (the email addresses) will be deleted and coding will be used to match the records going forward.

Completion of this part will take about 30-minutes. The participant does not get a copy of their assessment and will not be able to see their assessment results.

Once participants have been evaluated by the MHP using the HAM-D and HAM-A the Research Coordinator will email the $50 Amazon Gift card, PayPal or Visa cards.

Kintsugi staff and any data analysts will be blinded to the HAM-D and HAM-A scores.

The MHPs are either licensed mental health professionals or PhD students who have relevant experience assessing mental health patients as part of their course work. MHPs are remunerated at their market rate for conducting the clinical assessment. The MHPs do not provide any treatment to the participants. The MHPs provide a list of additional resources for the participant if they experience distress or discomfort (see page 2 of Participant Communication for Part II of Study Procedure attachment). MHPs are instructed not to establish any clinical relationship with the participant.

The MHPs are instructed to set all of their appointments with participants to private so that no one else can see these appointments. The MHPs have access to email addresses of the participants they assess because it's necessary to set up the video appointments. .

8. PARTICIPANT WITHDRAWAL

Participants can withdraw from the study at any time. If a participant decided to withdraw, a participant is encouraged to:

1. Call or email the Research Coordinator.
2. Let the research coordinator know about their intention to withdraw.

If a participant doesn't follow the steps to complete the study 2 weeks after the enrollment, they will be considered as withdrawn.

9. STATISTICAL CONSIDERATIONS

9.1 Study Outcome Measures

9.1.1 Primary Outcomes

- Hamilton Depression Rating Scale.
  - Participants will be interviewed by MHPs using the clinically validated HAM-D tool.
- Hamilton Anxiety Rating Scale.
  - Participants will be interviewed by MHPs using the clinically validated HAM-A tool.

9.2 Data Analysis Plan

The data collected during this study will be aggregated and for training a machine learning model.

10. ETHICS/PROTECTION OF HUMAN PARTICIPANTS

The investigator will ensure that this study is conducted in full conformity with the principles set forth in The Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research, as drafted by the US National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research (April 18, 1979) and codified in 45 CFR Part 46 and/or the ICH E6.

10.1 Institutional Review Board

The protocol, informed consent form, recruitment materials, and all participant materials will be submitted to the IRB for review and approval. Approval of both the protocol and the consent form will be obtained before any participant is enrolled. Any amendment to the protocol will be reviewed and approved by the IRB before the changes are implemented in the study

10.2 Participant Confidentiality

Participant confidentiality and privacy is strictly held in trust by the PI and staff. The study protocol, documentation, data, and all other information generated will be held in strict

confidence. No information concerning the study, or the data will be released to any unauthorized third party without prior written approval of the sponsor.

Authorized representatives of the Sponsor or representatives of the IRB may inspect all documents and records required to be maintained for the participants in this study. The Sponsor will retain the data gathered from this study indefinitely.

Participant confidentiality will be maintained by deleting the participants' Personal Identifiable Information (Email, Phone number, First and Last Name) once the study concludes.
Audio recordings will be stored on a password-protected Google Cloud Platform bucket indefinitely. Researchers will also be using Named Entity Recognition (NER) and number detection scripts to strip out any person names, organizations, locations, medical codes, quantities, medical record numbers, monetary values, and percentages after the study concludes. NER will be implemented using the NLTK library of Python.

The following documents will be retained after the study for a period of at least six years on a password-protected Google Cloud Platform bucket:

- Study protocol
- Informed Consent forms
- Participant recruitment ads
- PI and Research Coordinator's CVs
- PI and Research Coordinator's training certificates
- Amendments to the protocol
- Documentation of approval of amendments by IRB
- De-identified participant enrollment log
- Final report to IRB

The above listed electronic data will be deleted after 6 years. Data gathered from the study (audio recordings, HAM-D/HAM-A scores) will be stored indefinitely.

11. DATA HANDLING AND RECORD KEEPING

Research records will be kept in compliance with ICH E6, Section 4.9 and regulatory requirements for the protection of confidentiality of participants. The Research Study staff will permit authorized representatives of regulatory agencies to examine (and when required by applicable law, to copy) research records for the purposes of quality assurance reviews, audits, and evaluation of the study safety, progress and data validity.

Audio recordings will be saved on a password-protected Google Cloud Platform bucket that is only accessible to authorized personnel.

The audio recordings and HAM-D/HAM-A score data will be maintained for at least three years following the completion of the study. Participant confidentiality will be maintained by deleting the participants' Personal Identifiable Information (Email, Phone number, First and Last Name) and using codes once the study concludes.

12. PUBLICATION POLICY

We anticipate submitting to a peer-reviewed journal, such as Nature's NPJ or Karger Digital Biomarkers.

13. LITERATURE REFERENCES

1. Czeisler MÉ , Lane RI, Petrosky E, et al. Mental Health, Substance Use, and Suicidal Ideation During the COVID-19 Pandemic — United States, June 24–30, 2020. DOI: http://dx.doi.org/10.15585/mmwr.mm6932a1
1. ADAA. Facts & statistics: Anxiety and Depression Association of America. Retrieved March 05, 2021, from https://adaa.org/about-adaa/press-room/facts-statistics
2. Low DM, Bentley KH, Ghosh SS. Automated assessment of psychiatric disorders using speech: A systematic review. *Laryngoscope Investig Otolaryngol*. 2020;5(1):96-116. Published 2020 Jan 31. doi:10.1002/lio2.354
3. Mundt JC, Vogel AP, Feltner DE, Lenderking WR. Vocal acoustic biomarkers of depression severity and treatment response. Biol Psychiatry. 2012 Oct 1;72(7):580-7. doi: 10.1016/j.biopsych.2012.03.015. Epub 2012 Apr 26. PMID: 22541039; PMCID: PMC3409931.
4. Zhang L, Duvvuri R, Chandra KKL, Nguyen T, Ghomi RH. Automated voice biomarkers for depression symptoms using an online cross-sectional data collection initiative. Depress Anxiety. 2020 Jul;37(7):657-669. doi: 10.1002/da.23020. Epub 2020 May 7. PMID: 32383335.
5. Espinola, C.W., Gomes, J.C., Pereira, J.M.S. *et al.* Detection of major depressive disorder using vocal acoustic analysis and machine learning—an exploratory study. *Res. Biomed. Eng.* (2020). https://doi.org/10.1007/s42600-020-00100-9
6. Wang, J., Zhang, L., Liu, T. *et al.* Acoustic differences between healthy and depressed people: a cross-situation study. *BMC Psychiatry* 19, 300 (2019). https://doi.org/10.1186/s12888-019-2300-7
7. He L, Cao C. Automated depression analysis using convolutional neural networks from speech. J Biomed Inform. 2018 Jul;83:103-111. doi: 10.1016/j.jbi.2018.05.007. Epub 2018 May 29. PMID: 29852317.
8. Ozkanca Y, Göksu Öztürk M, Ekmekci M, N, Atkins D, C, Demiroglu C, Hosseini Ghomi R: Depression Screening from Voice Samples of Patients Affected by Parkinson's Disease. Digit Biomark 2019;3:72-82. doi: 10.1159/000500354
9. Yi H, Smiljanic R, Chandrasekaran B. The Effect of Talker and Listener Depressive Symptoms on Speech Intelligibility. J Speech Lang Hear Res. 2019 Nov 18;62(12):4269-4281. doi: 10.1044/2019_JSLHR-S-19-0112. PMID: 31738862; PMCID: PMC7201326.
10. Stassen HH, Bomben G, Günther E. Speech characteristics in depression. Psychopathology. 1991;24(2):88-105. doi: 10.1159/000284700. PMID: 1886971.
11. Takaya Taguchi, Hirokazu Tachikawa, Kiyotaka Nemoto 3, Masayuki Suzuki 4, Toru Nagano 4, Ryuki Tachibana 4, Masafumi Nishimura 5, Tetsuaki Arai 3 Major depressive disorder discrimination using vocal acoustic features https://pubmed.ncbi.nlm.nih.gov/28841483/
12. Lu-Shih Alex Low; Namunu C. Maddage; Margaret Lech; Lisa B. Sheeber; Nicholas B. Allen Detection of Clinical Depression in Adolescents' Speech During Family Interactions https://ieeexplore.ieee.org/document/5629355
13. Jingying Wang, Lei Zhang, Tianli Liu, Wei Pan, Bin Hu,Tingshao Zhu Acoustic differences between healthy and depressed people: a cross-situation study https://bmcpsychiatry.biomedcentral.com/articles/10.1186/s12888-019-2300-7
14. Tuka Alhanai, Mohammad Ghassemi, and James Glass Detecting Depression with Audio/Text Sequence Modeling of Interviews http://groups.csail.mit.edu/sls/publications/2018/Alhanai_Interspeech-2018.pdf

15. Wei Pan, Jonathan Flint, Liat Shenhav, Tianli Liu, Mingming Liu, Bin Hu, Tingshao Zhu Re-examining the robustness of voice features in predicting depression: Compared with baseline of confounders
16. Daniel M. Low MSc, Kate H. Bentley PhD, Satrajit S. Ghosh PhD Automated assessment of psychiatric disorders using speech: A systematic review https://onlinelibrary.wiley.com/doi/full/10.1002/lio2.354
17. Min Xu; et al. (2004). "HMM-based audio keyword generation" (PDF). In Kiyoharu Aizawa; Yuichi Nakamura; Shin'ichi Satoh (eds.). Advances in Multimedia Information Processing – PCM 2004: 5th Pacific Rim Conference on Multimedia. Springer. ISBN 978-3-540-23985-7.
18. https://www.json.org/json-en.html
19. Hamilton, M. (1960). A rating scale for depression. *Journal of Neurology, Neurosurgery & Psychiatry, 23*, 56-61. http://dx.doi.org/10.1136/jnnp.23.1.56
20. Trajković, G., Starčević, V., Latas, M., Leštarević, M., Ille, T., Bukumirić, Z., & Marinković, J. (2011). Reliability of the Hamilton Rating Scale for Depression: A meta-analysis over a period of 49 years. *Psychiatry Research, 189*(1), 1-9. http://dx.doi.org/10.1016/j.psychres.2010.12.007
21. Hamilton, M (1959).The assessment of anxiety states by rating