Title: A Pilot Study: Exploring The Role Of an Artificial Intelligence Powered Mental Health Support Tool For Mental Health Wellness For Physicians In Training

ID: STUDY00003196

NCT ID: Pending

Date: 6/6/2025

Table of Contents

1.0 Protocol Summary

Study Purpose	The purpose of this pilot study is to evaluate the acceptability of an artificial intelligence powered mental health support tool for mental health wellness amongst physicians in training.
Research Procedures	The primary research procedures are: • The primary research procedure will be the utilization of artificial intelligence powered mental health support session through a web browser for physicians in training to address symptoms related to depression and anxiety. A pre and post intervention survey, evaluating depression symptoms with the standardized PHQ-9 questionnaire and anxiety symptoms with the standardized GAD-7 questionnaire will be conducted to assess changes in mental health wellness.
Subject Population	The study will enroll physicians in training, including Internal Medicine residents and Gastroenterology fellows.
Duration of Subject's Participation	• The study includes at least 4 sessions. • The total study duration for each subject is 3 months.

2.0 Background, Rationale

Mental health is an important aspect of trainee's wellbeing, training and success. Mental health concerns may arise as changes in behavior and emotions and include symptoms related to stress, anxiety and depression (Nobleza, 2021). Physicians are an at-risk profession for suicide, and a recent systemic review and meta-analysis found that physicians in the USA are at higher risk for suicide as compared to other countries [Dutheil, 2019]. Depression is a known risk for suicide, with an estimated 20-43% of trainee physicians experiencing depression and symptoms of depression [Mata, 2015]. Psychosocial works factors thought to contribute to depression include lack of a cohesive teamwork, social support, delivering bad news, frequent contact with illness and death [Dutheil, 2019]. In addition, female physicians in training appear to be at a higher rate for depression which may be related to imbalances in work and household obligations [Guille,2017]. Numerous studies have demonstrated a higher prevalence of depression or emotional distress including anxiety, irritability, loneliness and stress among trainees and physicians. These conditions have been associated with adverse outcomes, including patient care errors, medical leaves, career shifts, and even suicide [Collier 2002,Craig 1968, Fahrenkopf 2008, Reuben 1985, West 2006, West 2009, Pospos 2019]. Emotional

distress not only poses health risk to the individual but has also been linked to poor-quality patient care and increased medical errors [Mata 2015, West 2009].

Studies have also shown that only a minority of physicians utilize mental health services [Nobleza, 2021]. Preventative strategies need to include recognition and management of stress and depression symptoms. Physicians in training have limited free time and unpredictable schedules limit ability to schedule appointments with mental health care providers. Artificial intelligence (AI) has been shown to comprehend language and respond with conversational and knowledgeable answers [OpenAI]. The effectiveness of the use of an AI therapist has not been evaluated in the management of mental wellness in regards to symptoms of depression and anxiety in physicians in training.

Artificial intelligence powered mental health support could provide an opportunity for mental health support that is easier to implement within the constrained work schedules of trainees. We propose a self-guided AI powered mental health support tool for physicians in training as an additional resource to address mental health management such as symptoms of depression and anxiety to promote mental health wellness in physicians in training.

3.0 Study Purpose and Objectives

The purpose of this pilot study is to evaluate the acceptablity of an Artificial Intelligence (AI) mental health support system as an additional resource to address mental health wellness amongst physicians in training at Cedars Sinai. The AI program is not meant to be an intervention or to replace mental health resources for the treatment of depression and anxiety diagnosis.

The primary objective is to determine if the trainees find the AI sessions through a web browser to be an acceptable and feasible wellness resource for promoting mental health wellness A pre and post intervention survey evaluating depression symptoms with the standardized PHQ-9 questionnaire and anxiety symptoms with the standardized GAD-7 questionnaire will be conducted as a measure of mental health wellness.

The secondary objectives of this study are to:

- Evaluate the prevalence of depression and anxiety in training physicians
- Determine what work-related and home-related factors physicians in training believe impact their mental health.
- To describe the acceptance and feasibility of AI powered mental health support sessions for physicians in training. If our study shows these sessions to be efficacious and feasible in this pilot cohort, we plan to recruit other training physicians in a larger study here at Cedars Sinai.
-

4.0 Study Population

Physicians in training at Cedars Sinai.

4.1 Inclusion Criteria

- Individuals 18 years old or older are included.
- Training physicians in Internal Medicine residency and Gastroenterology fellowship at Cedars Sinai.
- Study is limited to Cedars-Sinai employees who are known to speak English.
- HRP-514.2 academic HR compliance addendum review is attached (Appendix A)

4.2 Exclusion Criteria

- Pregnancy
- Active treatment of severe and uncontrolled anxiety or depression with a health care professional
- Unwilling and or unable to participate
- Non-English speakers.

4.3 Subject Identification, Recruitment, and Consent

Subjects will be initially identified in the following ways:

- The study population will be identified from the internal medicine residency and gastroenterology fellowship roster at Cedars Sinai. All subject's will be invited to participate in this study. Subjects will be recruited through an e-mail sent to the group roster and through didactic announcements.

4.3.1 Consent Description

Informed consent will take place via e-consent through REDCap. Participants will be invited via e-mail to participate in the study. The email will be sent to the group email for Internal medicine residents and the group email for Gastroenterology fellows for inclusion and limiting perceived coercion as no one person will be emailed directly. If trainees have questions or concerns, they will be directed to contact a team members who is not involved in trainee education or employment.

Potential Risks

Artificial Intelligence mental health support: The AI system developed here at Cedars trained system here trained the test system to behave with beneficence and act according to professional standards for a psychotherapist, there is potential that the artificially intelligent therapist response to the participant's communications in a manner perceived to be offensive, incorrect, or potentially traumatic. These adverse responses may cause the participant emotional or psychological harm.

Questionnaire risks: The questionnaires could make participants feel uncomfortable or embarrassed.

Potential Benefits

We expected little to no immediate benefit to be incurred by exposure of participants to one 20-minute session of AI powered mental health support. However, those who complete the 4 sessions over the course of the 4 weeks may experience mental health benefits. The long-term potential benefit of this investigation is evaluating the safety and feasibility of AI-powered mental health support tool over the course of 4 weeks. In doing so, the overall benefits and limitations of this approach will be better understood for future investigations and potential implementation. Furthermore, the information yielded from this investigation will inform further research and development of AI powered mental health support, which, if proven efficacious and safe, has the potential to greatly increase accessibility to high-quality mental health support to physicians in training with limited free time.

ASSESSMENT OF POTENTIAL RISKS AND BENEFITS

The risk of the artificially intelligent therapist generating adverse responses/advice during the session is rationalized by the nature of the investigation being a feasibility study of the implementation of an AI-powered mental health support tool. We anticipate that the built-in safeguards which minimize the ability of the model to generate derogatory, unprofessional, or dangerous responses will greatly minimize the chances of this occurring. In internal pilot testing within our lab, which has included over 100 mock conversations to date, we have not yet experienced any perceived adversarial, derogatory, unprofessional, or harmful response or advice. Nonetheless, an important purpose of this feasibility study is to confirm safety and appropriateness of the system. Furthermore, the risk of emotional or psychological adverse events as a result of the interaction with the AI model will also be minimized by providing participants with access to a hot line (Appendix E) to licensed mental health workers who are available at all times free at charge to trainees at Cedars Sinai. At the beginning of each session, participants will be provided instruction on how to access the licensed mental health works should an adverse event occur. If participants feel they need immediate assistance or report concerning suicidal thoughts or ideation, XAIA can send an immediate text and email to our social worker who is not involved in trainee education and employment. Our social worker can then contact the trainee for immediate intervention and provide immediate psychiatric resources available to trainees. In addition, a psychiatrist who is not involved in the proposed participants training, will be available for any urgent psychiatric attention.

5.0 Study Design and Procedures

5.1 Schedule of Events

Legend

- **R** = Research item/procedure done only for research purposes and their costs are covered by the study. You are not responsible for the costs of these procedures.

Procedures	Screening Visit	Visit #1	Visit #2	Visit #3	Visit #4	3-month follow-up
Informed Consent	R					
Questionnaires	R				R	R
VR sessions		R	R	R	R	

5.2 Study Design and Duration

This is a prospective pilot study with internal medicine residents and gastroenterology fellows to evaluate the role of AI mental health support tool as a wellness resource.

Participants will be invited via e-mail to participate in the study.]The invitation e-mail will describe the study and ask that those who meet exclusion criteria do not participate in this pilot study. The e-mail will also include a file with information of the study for participants to review including how study members are blinded and the de-coding process required for participants to be contacted by our social worker should they score ≥15 on PHQ-9 and or ≥10 on GAD-7. Those who choose to participate will be asked email our team member who is not involved in trainee education who will send the participants a REDCap link to the consent, survey and a code to be used in the survey.

Participants will be asked to complete a set of questionnaires (Appendix B, C and D) to evaluate their mental health and factors that contribute to mental health as well as four AI powered mental health support sessions over a four-week period. This will conclude the active phase of their participations. Participants will then be sent a questionnaire at 3 months from study initiation to complete a questionnaire evaluating their mental health and if they would be interested in additional AI powered mental health support sessions.

The questionnaires will be administered before the AI sessions and at the end of the study period to evaluate for changes in mental health scores. The questionnaires will include the validated PHQ-9 for a standardized depression score and the GAD-7 for a standardized anxiety score to evaluate symptoms related to depression and anxiety. In addition, a set of questions will focus on demographics, work and home environment to evaluate factors that may contribute to mental health prior to the AI sessions and a set of questions related to the AI sessions at the end of the study period. The questionnaire will be administered via REDCap.

To maintain subject safety and confidentiality the following work-flow will be followed.

1. Participants who choose to participate in the study will be given a redcap link for the consent form, survey and a code via e-mail. In order to complete the consent and survey, participants will have to enter the code provided to them.
2. Our social worker will be the only study members with a list of the subject's email and code.
3. Our social worker will review the survey responses on redcap.
4. If a participant survey has as a score on the PHQ9 >15 and or GAD>10, our social worker will contact the participant for evaluation and provide additional resources as needed.

5. During the sessions, XAIA will be programed to send a text message to our social worker from participants who report actionable comments such as suicidal thoughts and concerns. Our social worker will be available to receive the messages from XAIA and contact the trainee for immediate intervention and connect the trainee to additional resources as needed. In addition, a psychiatrist physician will be available for any urgent psychiatric attention.

The AI powered sessions will be administered via a web browser such as a computer, laptop, tablets or cellphone. Participants can complete the session in the privacy of their home or in a private location at work such as a call-room. Participants will be asked to complete one 20-minute session per week during the 4-week study period for a total of four sessions.

The AI program used will be the XAIA software program. XAIA was developed at Cedars Sinai. Participants will be sent an email with the link to use the XAIA software. This includes a tutorial on how to use XAIA. Typical interactions with XAIA involve a process of XAIA initially probing for topics with general inquiries "So [name], how can I help?", then asking more in-depth questions about a particular topic identified by the participant, and then offering support and/or coping methods about the identified problem or topic. Once XAIA is done speaking, the participant will answer the questions and press the “done speaking” button. XAIA then processes the messages and then replies. This process repeats until the participant informs XAIA that they wish to end the session and can click the “end session” button.

Conversations are processed through a HIPAA-compliant server: audio is recorded, transcribed by speech-to-text AI, and responded to by LLM (GPT-4). The LLM output is sent through an “Appropriateness Classifier”—a stand-alone AI to detect potentially dangerous or unhelpful responses. If triggered, the LLM is again queried, and its response is again analyzed. Otherwise, the output is released to a text-to-speech AI, which finally, together with sentiment analysis meta-data controlling XAIAs expressions, plays to the user. The system’s development involved iterative testing with therapists role-playing clinical scenarios, leading to continuous refinement of its psychotherapeutic communication. The program compiles and records data through audio responses. The responses from the user are recorded in-app, encrypted, and sent to a speech-to-text AI model. the resulting text transcript is sent to the LLM, which replies with text that is then converted to speech and played back to the user. Both text from the user and the LLM are stored on HIPAA-compliant servers. All audio data is not stored. Investigators have no access to the encrypted data.

The duration of the study session is 4 weeks and total duration of the study is three months.

The study will be conducted at one site only, Cedars-Sinai or participants may conduct the sessions at the privacy of their own home.

5.3 Description of Study Procedures

The study involves the completion of questionnaires to be done at the start and at the end of the study period and the completion of four sessions over the course of four consecutive weeks.

The overall study will last four weeks but we may contact participants up to 3 months after for follow up.

Questionnaires/surveys:
This study involves the administration of questionnaires to evaluate scoring changes for depression and anxiety symptoms as a measurement for mental health wellness before and after the intervention. This information is routinely collected as part of clinical care with the PHQ-9 and GAD-7 questionnaires. The questionnaires will be administered at baseline and at the end of the 4-week study period. As part of the initial questionnaire, there will be a set of questions that aim to determine demographics as well as work-related and home-related factors that contribute to mental health will be administered prior to the AI powered sessions. At the completion of the 4-week session a questionnaire with a set of questions evaluating the use and acceptability of the AI mental health support will be administered. Finally, participants will be sent a questionnaire including the PHQ-9 and GAD-7 along with a set of questions evaluating their interest in participating in additional sessions at 3 months. The follow-up questionnaires will also provide information on active resources for mental health available to trainees at Cedas Sinai. In total, the questionnaires should take no longer than 20 minutes to complete at the beginning and end of the study period and no longer than 10 minutes at follow-up. The questionnaire will be administered electronically via REDCap. The study will not collect personal identifying information. The questionnaire will also contain a link with information on how to obtain immediate mental health assistance including a hot line as provided by GME to all trainees at Cedars Sinai. This includes access to a mental health professional who is on call for any psychiatric emergencies.

We will be conducting following questionnaires.

- Patient Health Questionnaire -9 (PHQ-9) – Will be administered at the beginning and end of the study as part of Appendix B, C and D.
    - This questionnaire is validated.
    - The questionnaire is available in languages other than English.
- General Anxiety Disorder- 7 (GASD-7) – will be administered at the beginning and end of the study as part of Appendix B, C and D.
    - This questionnaire is validated.
    - The questionnaire is available in languages other than English.
- Work and Home Environment – will be administered only at the beginning of the study as part of Appendix B.
    - This questionnaire is not validated.
    - The questionnaire is not available in languages other than English.
- Artificial Intelligence Session Assessment – will be administered only at the end of the sessions as part of Appendix C.
    - This questionnaire is not validated.
    - The questionnaire is not available in languages other than English
- Post Artificial Intelligence Session Follow-up – will be administered at 3 months after study initiation as part of Appendix D.
    - This questionnaire is not validated
    - The questionnaire is not available in languages other than English

Behavioral Intervention:
This study involves a behavioral intervention. The behavioral intervention is the completion of a self-guided AI mental health support sessions by each participant. Participants will be asked to complete the session once per week for four consecutive weeks. Each session should take 20 minutes for completion. This intervention will take place at the participant's home or at Cedars in a location with access to a web browser via computer, tablet or cellphone. Those who choose to complete the session at Cedars-Sinai, may choose to use a call room dedicated for trainees. The intervention will be implemented by participants themselves.

The behavioral intervention will be implemented in the following manner:

- Participants will use the link provided by our social worker to gain access to XAIA. They will then follow the 2 minute tutorial provided directly by the link to use the software on a web browser. Participants can use a computer, laptop, tablet or cellphone with a web browser to access the XAIA software.
- Participants will be asked to spend 20 minutes during each session. They may choose to spend longer than 20 minutes but no more than 60 minutes.
- Participants will be asked to complete 1 session per week during the 4-week study phase.
- Participants will be asked to complete the sessions between the hours of 0700 and 2000 in order to have immediate access to our social worker and psychiatrist if needed.
- The email with link provided to access the XAIA software will include information for access to immediate mental health assistance such as a hotline provided by GME to all trainees at Cedars Sinai (Appendix E). This includes access to a mental health professional who is on call for any psychiatric emergencies.

6.0 Data Collection and Management

6.1 Data Procurement

- **Identification/Access/Abstraction**
  - ☐ Members of the study team will require access to the clinical data source (e.g., electronic medical record) to identify eligible data/specimens and to conduct data abstraction or gain access to specimens.[1]
  - ☐ Electronic Information Systems (EIS) Department will identify and/or abstract applicable data or specimens (e.g., RISCC website, request query of Deep6 through EIS).
  - ☐ Separate registry or repository will identify, abstract, and/or provide specimens and/or data to the study team.
  - ☒ Other:

Members of the study team will only require access to data provided by the questionnaires completed by participants. No personal or identifying information will be used in this study.

[1] Clinical records can only be accessed by study team members who are listed on the CS-IRB application and are IRB Certified.

- **Source(s) of Data/Specimens**:
  No medical record data will be accessed as part of this research.
  All data collected is from the survey/questionnaire.

6.2 Time Period of Data under Review

- One year
- Information will be kept for 2 years.

6.3 Data Elements

- The following data points will be collected:
  - *Age*
  - *Sex / Gender Identity*
  - *Race/ Ethnicity*

The information to be accessed and reviewed is that which is minimally necessary to achieve the goals of this research.

HIPAA Identifiers
☐ **No HIPAA Identifiers will be collected for this study (or select the identifiers from the following list).** *The investigators will not attempt to re-identify subjects from the collected data.*
☐ Medical Record Number
☐ Name
☐ Address (all geographic sub-divisions smaller than state, including street address, city county, and zip code)
☐ All elements (except years) of dates (including birthdate, admission date, discharge date, date of death, and exact age if over 89)
☒ Telephone Numbers
☐ Fax Numbers
☒ Email Address
☐ Social Security Number
☐ Health plan beneficiary number
☐ Account Number
☐ Certificate or license number
☐ Vehicle identifiers and serial numbers, including license plate numbers
☐ Device identifiers and serial numbers
☐ Web URL
☐ Internet Protocol (IP) Address
☐ Finger or Voice Print
☐ Photographic Image (photographic images are not limited to the face)
☐ Any other characteristic that could uniquely identify the individual

6.4 Confidentiality and Security of Data

- **Secure Storage**: Data will be housed in a HIPAA-compliant secure storage system, like REDCap within the Cedars-Sinai network with access restricted to approved members of the research team.
- **Limited Access:** Private identifiable information will be accessible only to IRB approved study team members with current IRB training.
- **Unique ID Numbers and Coding:** Each participant will be assigned a unique ID number. The linking list will be kept secure. Direct identifiers listed in section 8 will be separated from the study materials (data and/or specimens) as soon as possible. During data abstraction, no personal information is required.
- **Destroying Identifiers:** The identifiers and the linking list will be destroyed as soon as scientifically possible and maintained only as long as necessary to abstract, analyze and verify data.
- **Retention/Destruction of Study Materials:** Study data/materials will be kept and/or destroyed according to applicable policy.

7.0 Data and Safety Monitoring

7.1 Data and Safety Monitoring Plan

The study will be monitored by the PI to ensure appropriate study conduct, including obtaining proper access to data, compliance with the HIPAA Privacy Rule, compliance with Cedars-Sinai policy, and adhering to the plans outlined in the protocol for all study procedures, abstracting and recording data, data security and maintenance, and data accuracy and integrity. Any adverse events, deviations, protocol exception requests, potential unanticipated problems involving risks to subjects or others, or other events will be submitted to the IRB in accordance with IRB reporting policy.

7.2 Quality Control and Quality Assurance

- The PI will delegate a study team member to conduct QC/QA activities. The PI *will be responsible for the evaluation of data quality and how frequently this will be done quarterly*. Data will be evaluated for adherence with the protocol and for accuracy in relation to source documents.

8.0 Sample Size and Statistical Considerations

8.1 Sample Size

This is a single-site study at Cedars-Sinai. The sample size of this pilot study is 20 subjects.

8.2 Statistical Sample Size Justification

A statistical sample size is not required because this study involves:

- A pilot study with at most 20 subjects.

8.3 Statistical Analysis Methodologies

No statistical methods description is required given that the study does not require a sample size justification.

9.0 References

1. Nobleza, D., Hagenbaugh, J., Blue, S. et al. Resident Mental Health Care: a Timely and Necessary Resource. Acad Psychiatry 45, 366–370 (2021)
2. Dutheil F, Aubert C, Pereira B, Dambrun M, Moustafa F, Mermillod M, Baker JS, Trousselard M, Lesage FX, Navel V. Suicide among physicians and health-care workers: A systematic review and meta-analysis. PLoS One. 2019 Dec 12;14(12)
3. Mata DA, Ramos MA, Bansal N, et al. Prevalence of depression and depressive symptoms among resident physicians: a systematic review and meta-analysis. *JAMA*. 2015;314(22):2373-2383.
4. West CP, Tan AD, Habermann TM, Sloan JA, Shanafelt TD. Association of resident fatigue and distress with perceived medical errors. JAMA. 2009;302(12):1294–1300
5. Pospos S, Tal I, Iglewicz A, et al. Gender differences among medical students, house staff, and faculty physicians at high risk for suicide: A HEAR report. Depress Anxiety. 2019;36(10):902-920.
6. Guille, C., Frank, E., Zhao, Z., Kalmbach, D. A., Nietert, P. J., Mata, D. A.,Sen, S. (2017). Work-family conflict and the sex difference in depression among training physicians. JAMA Internal Medicine, 177, 1766–1772
7. Collier, V. U., McCue, J. D., Markus, A., & Smith, L. (2002). Stress in Medical Residency: Status Quo after a Decade of Reform? Annals of Internal Medicine, 136, 384-390. http://dx.doi.org/10.7326/0003-4819-136-5-200203050-00011
8. Craig, A. G., & Pitts, F. N. (1968). Suicide by Physicians. Diseases of the Nervous System, 29, 763-772.
9. Fahrenkopf, A. M., Sectish, T. C., Arger, L. K., Sharek, P. J., Lewin, D., Chiang, V. W., Landrigan, C. P. et al. (2008). Rates of Medication Errors among Depressed and Burnt out Residents: Prospective Cohort Study. British Medical Journal, 336, 488-491. http://dx.doi.org/10.1136/bmj.39469.763218.BE
10. Reuben, D. B. (1985). Depressive Symptoms in Medical House Officers. Effects of Level of Training and Work Rotation. Archives of Internal Medicine, 145, 286-288. http://dx.doi.org/10.1001/archinte.1985.00360020114019
11. West, C. P., Huschka, M. M., Novotny, P. J., Sloan, J. A., Kolars, J. C., Habermann, T. M., & Shanafelt, T. D. (2006). Association of Perceived Medical Errors with Resident Distress and Empathy: A Prospective Longitudinal Study Journal of the American Medical Association, 296, 1071-1078. http://dx.doi.org/10.1001/jama.296.9.1071
12. West, C. P., Tan, A. D., Habermann, T. M., Sloan, J. A., & Shanafelt, T. D. (2009). Association of Resident Fatigue and Distress with Perceived Medical Errors. Journal of the American Medical Association, 302, 1294-1300. http://dx.doi.org/10.1001/jama.2009.1389
13. OpenAI. ChatGPT: Optimizing Language Models for Dialogue. OpenAI. Published November 30, 2022. https://openai.com/blog/chatgpt