Artificial intelligence for mental health diagnoses and treatment plans: A protocol for actors and patients.

Prof. Mark Weiser

Dr. Asaf Caspi

Hadasa Afgin

Ver. 2.3

27.03.2024

Background

Artificial intelligence (AI) is becoming prevalent in modern medicine and psychiatry. AI is based upon a wide variety of computer algorithms classified under machine learning (ML). Beginning in the 1960s, a computer program known as ELIZA was developed to emulate psychotherapists' conversational abilities. By simulating human conversation, the machine would help the patient do most of the cognitive work of interpretation. As a result of the program, the field of artificial intelligence has expanded rapidly, initially intended only for research purposes involving natural language processing experiments. In 1971, another computer model was designed to simulate paranoia in a diagnostic psychiatric interview. An attempt was made to characterize the inner structure of paranoid behavior often encountered by clinicians when interviewing paranoid patients (Pham at all, 2022).

Artificial intelligence algorithms were able to identify healthy patients from patients with psychotic disorders with an accuracy of more than 70%. Artificial intelligence has been used in studies to examine speech disorders and differentiate speech samples from patients with psychotic disorders from healthy controls. Algorithms could detect where speech sample incoherence occurred, predict levels of incoherence, and were sensitive to very subtle deviations in long clinical interviews (Ray at all, 2022).

The purpose of the present study is to evaluate this potential, using AI. It will conduct an interview with standardized SP patients (actors) presenting a psychiatric illness, and based on the interview, it will present a differential diagnosis and treatment plan. Immediately afterwards, the actors will be interviewed by a board certified psychiatrist, who will also give a differential diagnosis and a treatment plan. The results of the AI and psychiatrist will be compared. In the second part of the study, AI will examine patients coming for consultation by a psychiatrist in the inpatient units, outpatient units or in the emergency room (ER) in Sheba. The patient will give informed consent, and will be interviewed by AI. The interview will be done either digitally, or will use to voice-to-text and text-to-voice technology that will allow the patient to speak with the computer, instead of typing questions into the computer. Future studies, not included in the current proposal, will use voice analysis, face recognition and computerized emotion recognition as part of AI.

Methods

Participants:

1.Five actors SPs, each time employed by MesserMSR, the Israeli Center for Medical Simulation, trained to simulate symptoms of psychiatric illnesses.

2.Patients (n=200)treated in the Division of Psychiatry in the Sheba Medical Center. In addition, patients who applied for treatment in the psychiatric clinic in the Division of Psychiatry in the Sheba Medical Center but are not yet treated there, and patients treated in the psychiatric emergency room (ER) in Sheba. In additional, a recruitment ad will be published in the media and on social networks, to include patients who will come in specifically to participate in the study.

3.Ten board certified Psychiatrists from the Drora and Pinchas Zachai Division of Psychiatry in the Sheba Medical Center.

Scenarios for actors:

The actors will be trained in one of 5 scenarios: depression, anxiety, PTSD with cannabis abuse, psychosis and control patient with no diagnosis. Each scenario will have 2 levels of severity. The actors will present demographic and medical background, symptoms and signs, including disorders of thought and affect.

Procedure:

Actors or patients will be interviewed by AI. The interview will either be digital or will utilize AI voice to text and text to voice technology, so that the participant will be speaking with the application which will respond vocally. AI After the AI interview the participant will be interviewed by a board-certified psychiatrist, who will determine the differential diagnosis and recommend treatment. Both, the AI and the psychiatrist, interviews will be audiotaped. The entire procedure will be repeated twice, for two days. After that, technological improvements will be made in AI. For the third time, the SPs will play against the AI only to check that the technological changes made didn't harm the quality of the AI.

Proposed schedule for actors

Time line ***		AI session 1

		Doctor 1	Doctor 2	Doctor 3	Doctor 4	Doctor 5
	1	Depression Low severity	Anxiety Low severity	PTSD with cannabis abuse Low severity	Psychosis Low severity	Control patient
	2	Depression High severity	Anxiety High severity	PTSD with cannabis abuse High severity	Psychosis High severity	Control patient
		Face to face phase *				
		Doctor 1	Doctor 2	Doctor 3	Doctor 4	Doctor 5
	3	Depression Low severity	Anxiety High severity	PTSD with cannabis abuse Low severity	Psychosis Low severity	Control patient
	4	Psychosis High severity	Depression Low severity	Anxiety Low severity	PTSD with cannabis abuse Low severity	Control patient
	5	PTSD with cannabis abuse High severity	Psychosis Low severity	Depression High severity	Anxiety High severity	Control patient
	6	Anxiety Low severity	PTSD with cannabis abuse High severity	Psychosis High severity	Depression High severity	Control patient
		AI session 2 **				
	7	Depression High severity	Anxiety High severity	PTSD with cannabis abuse High severity	Psychosis High severity	Control patient
	8	Depression Low severity	Anxiety Low severity	PTSD with cannabis abuse Low severity	Psychosis Low severity	Control patient

* We can choose any other balance of disorder and severity between doctors.

** two AI sessions to overcome the question per hour limit.

*** We will set the exact schedule prior to the training. The study should take half a day. 3:30-4 hour day.

For the actor's part of the project, over all there will be 125 encounters, 75 with AI and 50 with a psychiatrist in 3 days. The interviews of the actors and psychiatrist in MSR will be videotaped, so that in the future voice and facial analysis can be integrated into the project.

For the patient's part of the project, 200 patients will be interviewed by AI and then by a psychiatrist. This includes patients treated in the Division of Psychiatry in the Sheba Medical Center. In addition, patients who applied for treatment in the psychiatric clinic in the Division of Psychiatry in the Sheba Medical Center but are not yet treated there, and patients treated in the psychiatric emergency room (ER) in Sheba. In additional, a recruitment ad will be published in the media and on social networks, to include patients who will come in specifically to participate in the study.

For patients arriving for intake at the clinic - the study coordinator will contact them before the appointment, and ask them to arrive an hour earlier for the study. These patients will be paid NIS 100 for their time and travel reimbursement.

For patients who are on the waiting list, the study coordinator will contact them and explain that they be can participate in the study: will be examined by Liv and then receive a psychiatric examination by a psychiatrist. The examination will be performed for research purposes only, the patient will receive a brief letter to the family doctor with the diagnosis and recommendations for treatment .

The first approach to the patient will be made in an SMS, the wording will be (Hebrew below):

"Dear Sir/Madam,

"At the Sheba Medical Center, a study is currently being conducted on the topic: "Artificial intelligence to be used as a diagnostic tool in mental health and recommendations for a treatment plan." And we would like to offer you to participate in it.

If you have no objection, we will contact you by phone in the coming days for a short questionnaire.

Any information provided during the interview will remain confidential and will be used for research purposes only.

Your response to the research will contribute to the understanding of the researched topic and the expansion of general knowledge.

If you are not interested in participating in the study or hearing more details about it, please reply to this message "not interested".

If the patient does not answer because he is not interested. The research coordinator will contact you by phone.

For social media ads, the wording will be (Hebrew below):

"Participants are needed for a research study

In the psychiatric department at Haim Sheba Tel Hashomer Hospital, a groundbreaking study is currently underway on the subject:

"Artificial intelligence to be used as a diagnostic tool in mental health and recommendations for a treatment plan."

As part of this study, we are recruiting men and women over the age of 18 The main purpose of the study is to compare the results of the AI system test with the results of a psychiatrist's examination.

Participation in the study includes arriving at the Sheba Tel Hashomer Hospital for examination by the AI system and a psychiatrist. It will be possible to receive the summary of the meeting with the psychiatrist. Expenses will be reimbursed in the amount of NIS 100.

The details of the subjects will be kept confidential

Those interested in participating are welcome to send a text/WhatsApp message

Tel: 052-6667023

In addition, all participants in the study will be asked to fill out a satisfaction questionnaire (attached Appendix 1). And the examining psychiatrists will be asked to fill out a form of diagnosis, severity and recommendations for drug and psychotherapeutic treatment (attached in Appendix 2). This interview will be audiotape.

The primary outcome of the study will be a comparison of the interviews, differential diagnoses and recommendations for treatment assigned by the AI with those assigned by the board certified psychiatrist. the rates of agreement will be the outcome of interest.

All actors, patients, and psychiatrists participating (The psychiatrists who participate in the actors part) will give informed consent.

security and privacy:

1-Actors in MSR: The interviews will be recorded and stored in the Simboost server managed by MSR.

2-Patients in the Division of Psychiatry and ER: after giving informed consent, the investigators will enter the patients ID number into a conversion table, so that each patient will receive a code number which will identify the patient for AI.

The conversion table and the audiotape interviews will be stored in the Sheba secured server, which be accessed only by the research team.

Statistical Analysis Plan

Data

Actors who simulate one of four psychiatric illnesses (anxiety, depression, PTSD, or Mania) or a control condition that is not considered psychotic) are interviewed by a psychiatrist or a computerized AI program (called Liv), and for each interview the following is recorded:

(i) Interviewer (psychiatrist or Liv)

(ii) If psychiatrist, identity number

(iii) True diagnosis

(iv) True severity (mild/severe)

(v) Medication indicated

(vi) Psychiatric treatment indicated

(vii) Primary diagnosis made

(viii) Secondary and Tertiary diagnoses made

(ix) Severity (mild/moderate/severe)

(x) Medications recommended

(xi) Psychiatric treatment recommended

The codes 0, 5 (no diagnosis) and 9 (adjustment disorder) are taken to be equivalent diagnoses in the analyses described below.

Analyses

Note that in the analyses below, all estimated percentages will be accompanied by a 95% confidence interval.

     (a) Diagnosis:

The overall results will be summarized as:

1. % agreement of the primary diagnosis with gold standard over all conditions: (i) for Liv; (ii) for psychiatrists.

2. % agreement between Liv and psychiatrists; this calculation will need a special method – see paragraph (f) below.

3. The results will be displayed also as cross-tabulations of the 5 conditions according to gold standard versus Liv, according to gold standard versus psychiatrists, and according to psychiatrists versus Liv.

4. Analyses 1, 2, and 3 above will be repeated under the rule that agreement with the gold standard is achieved if one of the diagnoses (primary, secondary or tertiary) agrees with the gold standard, and agreement with the psychiatrists is achieved if one of the Liv diagnoses (primary, secondary or tertiary) agrees with the primary diagnosis of the psychiatrist.

(b) <u>Diagnosis according to severity of condition</u>:

The following analyses will be conducted:

5. As in 1 and 2 above, but for severe conditions only.

6. As in 1 and 2 above, but for mild/moderate conditions only.

7. As in 3 above, but for severe conditions only.

8. As in 3 above, but for mild/moderate conditions only.

9. As in 4 above, but for severe conditions only.

10. As in 4 above, but for mild/moderate conditions only.

(c) <u>Diagnosis and Severity</u>:

Note that throughout, **mild and moderate severity assessments will be considered equivalent**. If the psychiatrist or Liv classifies a patient with mild severity as having moderate severity, this will not be counted as an error.

Severity assessments will be evaluated in two different ways:

First, they will be included as part of the definition of a correct classification: e.g. a patient with severe anxiety will be considered correctly classified only if the psychiatrist or Liv classifies them as having severe anxiety. A classification of mild/moderate anxiety will be counted as an error. Analyses using this type of evaluation are listed below.

11. As in 1 and 2 above.

12. As in 3 above, but here the table will be a 9 x 9 cross-classification, with the following categories: mild anxiety, severe anxiety, mild depression, severe depression, mild PTSD, severe PTSD, mild Mania, severe Mania, control.

Second, **within the set of correctly assigned primary diagnoses**, we will do the following analyses:

13. % agreement of assigned severity scores with the gold standard (i) for Liv and (ii) for psychiatrists.

14. % agreement of assigned severity scores by Liv versus that assigned by the psychiatrists.

15. The results will be displayed also as 2 x 2 cross-tabulations of the 2 severities according to gold standard versus Liv, according to gold standard versus psychiatrists, and according to psychiatrists versus Liv.

(d) Medications and psychiatric treatments:

Agreement of the recommendations provided by Liv and by the psychiatrists will be compared with the gold standard. As with the second approach to evaluating severity assessments, the comparisons will be made only **within the set of correctly assigned primary diagnoses.**

There will be only one primary gold standard medication (psychiatric treatment), and the recommendation by Liv or the psychiatrist will be judged correct only if it includes this primary medication (psychiatric treatment).

Thus the two main analyses will be:

16. % agreement with primary gold-standard medication of (i) Liv's recommended medication and (ii) psychiatrists' recommended medication.

17. % agreement with primary gold-standard psychiatric treatment of (i) Liv's recommended treatment and (ii) psychiatrists' recommended treatment.

(e) Decision thresholds:

On the basis of certain of the above statistics, we specify two conditions that must be satisfied for Liv to proceed to the Phase II study that involves real patients arriving for diagnosis and treatment at Sheba Medical Center.

> A. The % agreement of the psychiatrists' primary diagnosis with the gold standard is >85%.

If it falls below this value, then it will be concluded that the "simulation" was not sufficiently realistic and some adjustments to the training of the actors will need to be made.

B. The % agreement of Liv's primary diagnosis with the psychiatrists' primary diagnosis >85%. If it falls below this value, then it will be concluded that adjustments are required to the Liv program.

Reference

McNaughton, N., Ravitz, P., Wadell, A., & Hodges, B. D. (2008). Psychiatric education and simulation: a review of the literature. *The Canadian Journal of Psychiatry*, *53*(2), 85-93.

Pham, K. T., Nabizadeh, A., & Selek, S. (2022). Artificial Intelligence and Chatbots in Psychiatry. *The Psychiatric quarterly*, *93*(1), 249–253

Ray, A., Bhardwaj, A., Malik, Y. K., Singh, S., & Gupta, R. (2022). Artificial intelligence and Psychiatry: An overview. *Asian journal of psychiatry*, *70*, 103021.

Appendix 1

שאלון שביעות רצון

Dear participant,

In order to provide the best care to our patients, we would appreciate it if you filled out the following questionnaire. The questionnaire is not mandatory, but your answers to the following questions will help us make improvements to the computerized system you used.

Indicate the extent to which you agree with the following statements:

strongly disagree	Disagree	neutral	Agree	Strongly agree	
					LIV's assessment was thorough and complete
					I was told what to do if my symptoms worsened
					LIV seemed to be genuinely interested in me
					LIV seemed to understand my problems
					LIV treated me with respect

					LIV seemed to know what she was doing
					I felt I could trust LIV
					LIV checked with me to see if I had any questions
					My general satisfaction level is high
					I will recommend LIV to a friend or relative
					After the assessment with LIV I have greater hope for improvement
					LIV was empathetic to me
					The login and identification process was easy to use
					The use of the button that activates the recording was convenient and friendly
					LIV had no errors in understanding what I said
					There were no significant male/female gender errors
					There were no words that LIV made up and I didn't understand
					There were questions that I felt were direct/intrusive
					LIV asked me enough about topics that were important to me
					The conversation flow was continuous and comfortable
					I left the conversation with a good experience

Appendix 2

Intake Summary- Differential Diagnosis

Diagnosis:	
	1._____
	2._____

	3._____
Severity:	• Mild- Moderate • Severe
Hospitalization:	Yes / No
Medication:	Yes / No
Medication:	1._____ 2._____ 3._____ 4._____
Psychotherapy:	• None • Dynamic • CBT • Supportive • Psycho-education • Other: _____