

人工知能で、人のこころは癒せるか？

人工知能(自然言語処理)フィードバック機能搭載型のインターネット認知行動療法(iCBT-AI)の抑うつ者に対する世界初の効果検証(無作為統制試験)

Can Artificial Intelligence Heal Human Hearts?: A randomized controlled trial on the effects of internet-based cognitive behavioral therapy with natural language processing on depression

宗 未来*1 関沢洋一*2 竹林由武*3
Mirai So Yoichi Sekizawa Yoshitake Takebayashi

*1 慶應義塾大学医学部 Keio University School of Medicine *2 (独) 経済産業研究所 Research Institute of Economy, Trade and Industry *3 福島県立医科大学 Fukushima Medical University

1. 背景

うつ病は自殺危険因子でもある重大な障害である。また、その予備軍である軽症うつ状態まで含めると、高額な社会的コストにもつながる。うつ病を改善するための取り組みとして、非適応的な考え方や行動習慣を修正する「認知行動療法」が注目されている。しかし、少ないセラピスト数やコスト高などにより、未だ十分な広がりにはみせていない。その対応策として、諸外国ではセラピストに頼る前にインターネットを活用した自習での認知行動療法(iCBT: internet-based Cognitive Behavior Therapy)によって、まず自分で取り組むことが推奨されており、英国をはじめとした諸外国では、軽症うつ病の第一選択として公的医療制度においてもiCBTが導入されている。

しかし、iCBTによる介入では短期的には抑うつ症状は改善するものの、効果が長期的に持続しない、脱落率が高い、社会機能の改善につながらないといった課題が残されている^[1]。専門家のサポートによる支援型iCBTは効果の増強が期待される反面、依然としてコストの問題を無視できない。軽症うつを呈する人々の潜在的な数が極めて多いことを考慮すると、iCBTにおける専門家による関与を最小限にするための対応が必要である。

近年、自然言語処理技術(NLP: natural language processing)を用いた対話エンジンをiCBTに応用して、画面上に架空のセラピストが登場して、エクササイズの実施者に共感の言葉を示したり、不適切な入力に対してアドバイスをしたりするiCBT-AIが民間企業により開発されている。これらの機能によって、誰にも共感してもらえない無味乾燥なエクササイズへの動機づけを高め、また間違った方法で行われていても修正されることなく終えていた従来のエクササイズと異なり適切なエクササイズの遂行を促すことが可能となった。

国際的にもこれまで人工知能機能を搭載したiCBTに対する応用は、2016年4月に米国においてiCBT-AIを用いた研究計画が報告されているだけであり、AIのiCBTへの実際の効果に関しては十分に検証されておらず、無作為統制試験による厳密な検証が求められていると考えられた。

2. 方法

2.1 対象者

本研究は、調査会社のモニターとして登録している人々に募

連絡先: 宗未来, 慶應義塾大学医学部精神神経科学教室
東京都新宿区信濃町 35, 03-3353-1211

集をかけることによって行われた。包有基準は、1)エントリー時の年齢が20~60歳、2)インターネット使用環境下にある、3)スクリーニング時に抑うつ尺度であるPHQ-9(後述)の得点が5点以上、4)日本語による説明書を理解し、同意できる判断能力を有する、とした。除外基準としては1)現在、メンタルヘルスの問題について専門家からなんらかの治療を受けており、主治医の承諾が得られない、2)介入期間中に認知行動療法を専門家から受ける、3)統合失調症に罹患している(医療機関による診断の自己申告)、4)認知症に罹患している(医療機関による診断の自己申告)、5)過去12ヶ月間に何らかの物質依存を有する(喫煙を除く)、とされた。

2.2 研究デザイン

本研究では、1)iCBT-AI群、2)AIを使用しない以外の機能はiCBT-AIと全く同じスペックを有する通常型iCBT群、3)待機群の3群間での併行群間比較により、どの群が最もうつ症状の軽減効果が大きいかなどを検証するための部分盲検化無作為割り付け統制試験である。主要評価指標としては、代表的なうつ評価指標であるPHQ-9(Patient Health Questionnaire)を用い、副次的項目として、参加率、不安(GAD-7)、抑うつ(QIDS-J)、機能(SDISS)、満足度(CSQ8-J)が評価された。

2.2 研究デザイン

ベースラインの質問に回答してもらった者は、性別・年齢・抑うつ重症度によって層別無作為化がなされ、iCBT-AI群、共感機能とアドバイス機能の有無以外には、iCBT-AIと全く同じスペックとなっている通常iCBT群、待機群に分けられた。iCBT-AI群とiCBT群のエクササイズ期間は7週間(平成27年10月5日~11月20日)で、エクササイズ期間が終了した直後に第2回目の評価が行われ(平成28年11月23日~29日)、更に、その3か月後に第3回目評価が行われた(2月22日~28日)。

2.3 解析手法

解析は、まず、PHQ-9の得点が10点以上が大うつ病水準の病態とされていることから、PHQ-9の得点が10点以上と10点未満の二値アウトカムを作り、これを従属変数として、ロジスティック回帰分析により、介入終了時と3か月後フォローアップ時において、共変量調整後に各群間で大うつ病の基準を満たす者のオッズ比に違いがあるかを検証した。次に、PHQ-9得点を加工せずに連続値として扱い、混合効果モデルによる反復測定データ解析法(Mixed Model Repeated Measures, MMRM)を行った。他の副次的評価指標についても同様の解析を行った。

2.3 解析手法

本研究は、経済産業省下のシンクタンクである(独) 経済産業研究所の助成により行われた。

3. 結果

iCBT-AI 群では、iCBT 群に比べてエクササイズにおける参加率が有意に高かった ($p<0.005$) (図参照)。全データでの結果では、通常の iCBT 群では、待機群と比べて、介入期間終了直後にうつ症状が改善する傾向 ($p=0.05$) が認められ、3 カ月後のフォローアップでは有意に低くなったが ($p=0.01$)、iCBT-AI 群では有意な改善は認められなかった。一方で、PHQ-9 の得点が 10 点以上 (大うつ病性障害水準に相当) の基準を満たす者の割合は、介入期間終了直後には通常の iCBT 群、iCBT-AI 群ともに待機群に対して有意差がなかったが、3 カ月後のフォローアップでは、iCBT-AI 群においてのみ、この基準を満たす者の割合が低い傾向 (改善傾向) が認められた (オッズ比 0.67; $p=0.08$)。PHQ-9 で 10 点未満の軽症うつ者に限った下位分析でみると、その減少は有意 (オッズ比 0.35; $p=0.02$) であった (表参照)。以上から、iCBT-AI は非 AI 型 iCBT に比べて短期的にはマイナス効果が予想される一方で、長期的には非 AI 型には認められない将来の重症抑うつ者を減らす可能性が示唆された。

4. 結論

iCBT-AI は、非 AI 型の iCBT に比べて参加率を有意に高めること、短期的な効果は非 AI 型に劣ること、長期的には非 AI 型には認められない将来の重症抑うつ者を減らす可能性があることが示唆された。また、iCBT-AI の参加率が非 AI 型よりも高かったことを考慮すると、iCBT-AI では iCBT に比べて多くの抑うつ者が行えるエクササイズとなる可能性が示唆された。本研究における iCBT-AI と iCBT との差は、NLP による共感機能とアドバイス機能の 2 つしか存在しない。そのため脱落に差が出る原因としても、それらのいずれか、もしくは両者が影響したと推測されるが、それらは今後予定されている二次解析の結果が待たれる。我々の知る限り、諸外国でもこのような人工知能の iCBT も含めた精神医療介入への応用に関する効果検証の報告は認められていない。走り出したばかりの本技術は限界も少なくないが、当研究を足がかりとして今後の更なる技術開発、およびその検証が期待される。

図 : エクササイズ (コラム法) の利用状況

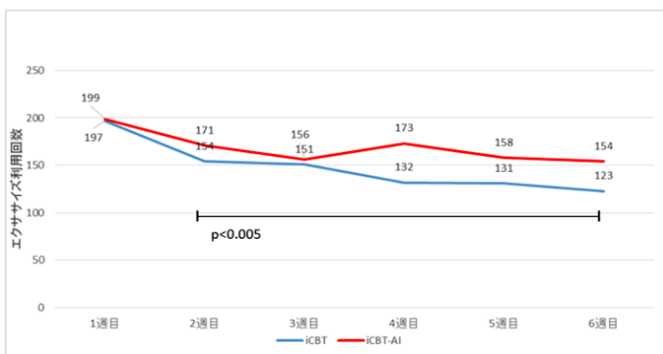


表 : ロジスティック回帰分析の結果 (うつのPHQ-9得点による重症度別)

開始時の得点		オッズ比(95%信頼区間);p値		
		iCBT群 vs iCBT-AI群	iCBT群 vs 待機群	iCBT-AI群 vs 待機群
5点未満	介入終了時	0.77(0.11 to 5.45);0.79	3.27(0.27 to 39.37);0.35	6.09(0.53 to 70.59);0.15
	3ヶ月フォローアップ時	0.97(0.13 to 7.45);0.98	1.29(0.17 to 9.89);0.81	1.15(0.13 to 10.26);0.90
5点以上 10点未満	介入終了時	0.67(0.29 to 1.54);0.35	1.06(0.47 to 2.39);0.89	1.37(0.66 to 2.84);0.40
	3ヶ月フォローアップ時	1.89(0.73 to 4.93);0.19	0.79(0.36 to 1.75);0.57	0.35(0.14 to 0.86);0.02
10点以上	介入終了時	0.81(0.42 to 1.57);0.54	0.59(0.31 to 1.10);0.09	0.71(0.37 to 1.35);0.29
	3ヶ月フォローアップ時	1.10(0.60 to 2.01);0.77	0.86(0.49 to 1.53);0.61	0.81(0.46 to 1.43);0.46

(注1) 性別、年齢、ベースラインのPHQ-9、GDA-7の数値を調整している。各時点毎に算出。

(注2) 赤字は5%水準で有意なもの。

5. 英語原稿

Background: In spite of recent high expectations of internet-based cognitive behavioral therapy (iCBT), iCBT still holds limitations including effect sustainability, function improvement, and dropout. Therefore, we focused on iCBT-AI in which automatic feedback by text or animation expressing empathy or indicating instruction is given to users with respect to their input after analysis conducted by a natural language processing (NLP) system, which is one area of artificial intelligence (AI). Since there is no evidence of iCBT using AI technology so far, we evaluated its effectiveness.

Methods: 1,187 participants recruited from the website were randomly assigned into three groups; iCBT-AI, conventional iCBT without AI, and waitlist as control. Those allocated to interventional arms were encouraged to perform each exercise at least once a week for seven weeks. The primary outcome was moderate-to-severe depression defined as a PHQ-9 score of 10 or higher. Intention-to-treat analyses were performed.

Results: The dropout rate was significantly lower in iCBT-AI than iCBT ($p<0.005$). At post-intervention, there were no significant differences between the interventional arms and the control group although only the iCBT-AI arm indicated a marginally significant lower odds ratio compared to the control group at the three-month follow-up (0.67 (95% confidence interval 0.43 to 1.04), $p=0.08$). In the secondary analysis stratifying participants by the severity of depression, only iCBT-AI users with PHQ-9 between 5 and 10 (corresponding to mild depression) showed a significantly lower odds ratio of having moderate-to-severe depression (0.35 (0.14 to 0.86), $p=0.02$) at the three-month follow-up.

Conclusion: Although iCBT-AI has no significant short-term antidepressant effect, iCBT-AI seems to have an exclusive potential to reduce moderate-to-severe depression in the future. Further research is required.

UMIN Clinical Trials Registry (ID: UMIN00001922)

参考文献

1. So, M., et al., *Is computerised CBT really helpful for adult depression? A meta-analytic re-evaluation of CCBT for adult depression in terms of clinical implementation and methodological validity.* BMC Psychiatry, 2013. **13**(1): p. 1-14.