Writing...
FACTS, OPINIONS & IDEAS

Published in *Artificial Intelligence for Better or Worse* in 2019 by Future Leaders (www.futureleaders.com.au)

# Artificial intelligence and mental health

Grant Blashki and Simon Lock

When we began to reflect on the role of artificial intelligence (AI) for mental health care, immediately there were a whole range of concerns that leapt out at us. How on earth could a non-living digital device be of value to any human being experiencing mental health issues? Isn't another human being absolutely essential to provide the compassion and understanding of fellow human beings' emotions, thoughts and feelings, and intimate troubles of life? Can a person really develop a sense of trust with AI like they might have with another person? And then further layers of trepidation immediately presented themselves to us. What about the duty of care for a person with a mental health issue, and what if the artificial intelligence makes a wrong assessment or provides the wrong advice? Or, even more subtlety, what if over time the AI is developing algorithms and drawing conclusions about people's mental health that are incorrect, biased or even discriminatory. And then there are some monumental privacy and data security concerns that have significant personal autonomy and human rights implications. It is fair to say, we approached this emerging field with a circumspect perspective.

Before we throw out the potential of artificial intelligence to genuinely help people who are experiencing mental health issues, it is well worth reflecting on some of the potential benefits. Currently, countries all around the world are grappling

with the high prevalence of mental health disorders and their enormous impact on people's day-to-day lives, the community, and society as a whole. The truth is that mental health workforces, even when well funded and supported in developed nations, are not able to keep up with the demand. We also know that even in the developed nations there are populations unserved by mental health services, such as remote, socially disadvantaged people or people who simply can't afford support. Digital mental health also has the potential to provide access to those people who are afraid to seek care because of perceived stigma. For many people, the idea of going along to a therapist for psychological care is embarrassing and they are worried about what other people think, and therefore instant access on the phone or via the internet can be a good way to circumvent this barrier. So, some of the enticing aspects of digital mental health are its low-cost, low stigma, scalability and high accessibility.

And, of course, looking through the scientific lens of evidence-based medicine, we owe it to our patients to explore just how effective these new types of interventions might be. Digital mental health interventions have the potential for consistent quality and the tireless delivery of mental health care at a reliable and constant standard, so researchers need to explore what works best for patients.

## Definitions

What are we talking about when we speak of artificial intelligence in the mental health field? It is worth exploring some of the terminologies and what they mean, even though they are often used interchangeably in the digital mental health field.

Let's start by pointing out an important distinction between traditional online therapies and AI mental health interventions per se. There is a plethora of internet-based mental health interventions usually referred to as e-therapies or e-counselling approaches, many of which have been proven to be effective. For example, moodgym is a well-researched cognitive behavioural therapy (CBT) based online therapy that has been proven in several randomised trials to be of great benefit to people experiencing mental health problems such as depression and general psychological distress.[1] Another example is My Compass, which has demonstrated significant improvement in symptoms of depression, anxiety and stress in work and social functioning.[2]

A distinguishing feature of AI-based mental health interventions as opposed to traditional online therapies is their ability for their algorithms to adapt and evolve. This means AI mental health interventions are designed to learn, and to adjust and change based on experience to make better decisions in the future.

Underpinning this type of AI is what is called machine learning. Machine learning is a group of statistical techniques that allow for a computer to improve at tasks with more experience completing them. Machine learning is where most of the development of mental health AI is now occurring, and its basic principles are described in this book in Chapter 4 by Khoa Cao and Luke Oakden-Rayner. Closely related to this is what is called deep learning, which allows for a machine to intake a significant amount of input and then train itself to achieve particular outcomes.

Despite these great advances, AI has so far proved to be most useful for very clear specialised tasks, but much less so when the task requires accounting for a broad perspective, unpredictability, or even common sense. Even so, there is already work underway utilising AI to provide real patient benefits. For the remainder of the chapter, we outline different examples of how artificial intelligence is being applied to the problem of mental illness — its use in early detection, its application as a diagnostic aid, its role as a treatment intervention, and its capacity to prognosticate likely clinical outcomes.

## AI for early detection of those at risk of mental illness

One of the first applications of artificial intelligence in this field has been the early detection of mental health conditions through analysis of people's online data. One fascinating example is research into Facebook data that helps to predict depression. In 2018, researchers from Pennsylvania looked at almost 700 electronic Facebook records, with the consent of participants in the study, and found that there were correlations between the use of certain types of language in their posts and the presence of depressive disorder. Notably, themes of loneliness, hostility, rumination and self-reference were associated with increased risk of depression.[3] The researchers also looked at the frequency of posts, the length of posts and demographic information. For many patients, picking up the early signs of depression relapse is a good way to put in place a treatment plan early, so the potential of social media type cues may be of benefit, though of course it raises major ethical, consent and privacy issues.

Twitter posts, even though limited to only 140 characters, also provide some fascinating insights into mental wellbeing.

Reece and colleagues, in 2017, developed sophisticated models for analysing Twitter posts to detect early signs of depression, and also post-traumatic stress disorder. Analysis and comparison of more than 250,000 tweets by over 200 participants, half mentally healthy and half experiencing mental illness, have shed some light on the potential of social media posts to be a prompt for early detection of mental conditions.

It's not just text which provides early cues on social media for mental health issues. A study from 2017 by Reece and Danforth showed that users of Instagram with clinical depression (as defined by a high score on a depression scale) were more likely to post pictures that were more blue, dark and grey, and also less likely to post images with faces in them. The study, involving more than 160 participants and over 40,000 images, found that mentally healthy users were more likely to select predefined filters with warmer and brighter colours, and users with depression were more likely to use black and white filters. One of the intriguing aspects of this approach is that the model was able to detect depression at rates better than a general practitioner.[4]

While these systems are not able to diagnose mental health issues with 100% accuracy, they do provide us with a useful first-point of call to identify early signs of mental health issues and provide people with the most appropriate support. How or if we would want to do this at scale raises important societal privacy and human rights issues (see Chapter 6 in this book by Solomon et al.), and from the public's perspective having your posts analysed this way has worrying Orwellian overtones. At the very least it is a reminder that we reveal more about ourselves on our social media posts than we imagine.

## AI as a diagnostic tool for mental illness

Beyond the use of AI for early detection, it has a role as a diagnostic tool. The use of diagnostic tools in the mental health area is also one of controversy and debate. Currently it is the case that in routine clinical care, psychometric scales are often used to inform the diagnostic process. So, for example, clinicians often use standardised validated scales such as the K10 (Kessler Psychological Distress Scale), the PHQ 9 (The Patient Health Questionnaire 9) or the DASS (Depression Anxiety and Stress Scale) to collate responses of participants to provide a score to estimate the likelihood of mental health conditions such as depression or anxiety. These scales are enormously helpful though they also have limitations in their narrow conception of mental distress and their imperfect sensitivity and specificity to detection of mental conditions.

Artificial intelligence opens the possibility of novel strategies to identify mental health conditions including the rapidly improving technology to analyse voice. One example is Cogito, which grew out of AI systems designed to monitor call centre interactions, and which monitors and analyses the conversation between an operator and a customer to make recommendations for operators regarding their next interaction with the caller. For instance, it flags if the caller is sounding stressed and prompts the operator to try to ask questions about the caller problem to reduce that stress. From this technology Cogito has gone on to develop a specific tool aimed to detect depression and post-traumatic stress disorder (PTSD), called "Companion", and they have worked with the US Department of Defence and the Defence Advanced Research Program Agency (DARPA) to pilot

the program with veterans who have suffered from PTSD. A 2017 study using this tool combines both analysis of voice and phone data such as texting, duration of calls and geolocation, and demonstrates an impressive capacity to accurately detect symptoms of depression and PTSD.[5]

## AI for monitoring progress

Another area where technology has been helping people with mental health issues is the area of monitoring symptoms and clinical progress over time, and again it is a technology that raises important privacy and autonomy issues. Clinicians have for many decades utilised mood tracking tools to help monitor patient progress, but AI brings a much more comprehensive (and perhaps intrusive) approach to tracking patient trajectories.

One example from Singapore is Cogniant, which monitors behaviour using phone data and informs the clinician on progress. Their tools use a number of datapoints from the patient's phone and digital devices to monitor the patient's daily routine and activities. This information can then be interpreted by the clinician during their next session or can trigger other events such as an early check-in call or escalation to emergency support if required.[6]

Another app, the Sophie app developed in Australia, is a keyboard app that overlays the child's screen and uses AI to analyse the language used by that child to identify mental wellbeing, and also detects predatory, bullying or other inappropriate communications. Parents can then be informed if their child is at risk. It is true to say that 50% of mental health conditions in adults begin under the age of 14 years old, and it is promising that this tool may be helpful for early detection.[7]

These type of monitoring interventions raise privacy and consent issues. While the Cogniant product does require consent from the users to use the service, the Sophie app consent comes from the parent, not the child. So, while a child may not wish to have all of their communications monitored, they have little ability to change the situation because of their age and legal status. There is also the possibility of users deliberately altering their responses because of the knowledge that they are being monitored. It is not uncommon when people know they are being monitored to modify their behaviour to achieve the outcome they desire or what they believe their clinician may want to see. This therefore can affect the accuracy of the data being analysed by these types of systems.

Another concern in relation to monitoring is data ownership. Currently most of these products are being developed by private companies and many of their terms and condition statements advise that they retain ownership of the data that is collected. As many of the example products discussed in this chapter are utilising a patient's health or mental health information, this ownership exposes users to risks that their personal information may be used or sold to other parties without their knowledge, and so this is part of a much broader discussion around data ownership and protection of personal data that must be considered when discussing AI.

## AI for management of mental health conditions

One of the fastest growing areas of development of AI in mental health has been the emerging role of AI as a management tool for psychological and behavioural issues. Many

companies have been exploring this area and it is projected to be an enormously lucrative growth area and, not surprisingly, some of the biggest companies in the world are getting on board. From the smorgasbord of new interventions, we have focused on a few relevant categories and examples: the chat bots, the smart speaker/digital assistants, and the virtual humans.

Chatbots provide a very accessible mental health support tool, usually on a smart phone, with simulated text-based conversation and guided education using videos, animations and quizzes to assist the user. Advantages include convenience, reliability, capacity to send reminders and low cost. Some preliminary evidence suggests that they can be effective for common mental disorders such as depression and anxiety.

Woebot is a popular chatbot that predominantly uses CBT approaches. The user has the sensation that they are having an SMS-type conversation with a person, and the interaction is often humorous and engaging. Based on the responses, the AI directs the user to one of a number of educational subprograms in the app, such as Practising Gratitude, Setting Goals or Finding your Strengths. A 2017 randomised controlled trial showed substantial improvements in depression and anxiety in young people using Woebot.[8]

Another chatbot Wysa also uses a CBT approach and offers an opportunity for users to discuss worries, learn CBT skills or embark on one of 40 conversational coaching tools around topics such as loss, conflict or anxiety, for example. Wysa also offers people an opt-in for a paid premium program to access a human coach.[9]

A third company that has been exploring this field has been X2Ai, who developed an AI called Tess. This psychological chat bot provides a range of personalised mental health services to user and can be accessed through existing communication channels such as a text messaging (SMS), Facebook messenger and Slack. It utilises evidence-based psychological approaches, including CBT, mindfulness, and acceptance and commitment therapy. A randomised controlled trial reported in 2018 showed significant improvements in depression and anxiety scales from students using Tess as compared to students just receiving education.[10]

Turning now to the digital personal assistants, the three big players are Amazon's Alexa, Google Home and Apple's Siri. There is intense competition regarding personal assistants and smart speakers and ultimately which company is going to be able to dominate the voice recognition field, with the extraordinary data opportunities. In this context there has been the emergence of a number of preliminary attempts to provide mental health treatments through these mediums.

There are some potential advantages of using voice recognition in the mental health field. Voice is such a direct way for people to communicate emotions. Many people now have smart speakers or a personal assistant on the smart phone and so the technology is widely available. Much more can be gleaned from voice than from text alone, such as intonation, pace and volume of speech, and AI is gradually refining its capacity to interpret these nuances of voice.

The potential for surveillance and misuse of data using personal assistant technologies is a topic understandably of

great concern to the public. The capacity to assess people's mental status without their consent or to overhear private conversations is of great concern and has implications for data misuse in, for example, targeting advertising, making insurance assessments, and assessing employment prospects and other aspects of people's lives. Ultimately, how secure and protected people feel their data is going to be will determine how trusting they are to engage with personal assistants to help them with their mental health.

Another AI development is virtual humans, which are digital images of face, and often body too, rendered on screen to mimic a human counsellor. Virtual humans are tools which operate in much the same way as a chatbot, but the interface is a computer-rendered image of a human being that is able to communicate with users through verbal and non-verbal means. An example of this is Sim Sensei, which was created by University of Southern California, and researchers found that virtual human interfaces increased a participant's willingness to disclose information about themselves.[11] In this situation, the use of virtual psychologists or AI chatbots may reduce those barriers for people to seek or attend support services that help them with their issues; however, there is still a long way to go to make these technologies feel natural for the user.

## AI for prognostication

Another application for AI is to help forecast clinical outcomes. These tools utilise datapoints from a large selection of patients to help predict the likely outcomes for individual patients and have been shown in some cases to be even more effective than doctors at getting these predictions correct. A recent study

explored the likely outcomes for people with mental health issues such as depression and psychosis, and using clinical and radiological imaging data showed that AI was able to correctly predict the social outcomes for patients at one year with 83% reliability compared to clinicians' ability to predict these outcomes correctly only 70% of the time.[12] This suggests that in the future, clinicians may well work much more closely with AI to better refine their prognostic determinations.

## Conclusion

Artificial intelligence has some useful applications in the mental health field. At this stage, it is especially of value in early detection, diagnosis, and treatment and assessing prognosis. Society is understandably and appropriately concerned about the way in which data is collected and used, and there is still a long way to go with regard to the ethical issues around consent, privacy and the use of data. However, the ubiquitous uptake of technologies such as smart phones and smart speakers means that there will be more platforms on which AI mental health care can be delivered. As these interventions continue to be refined, it is likely that they will grow to supplement face-to-face care and provide ongoing support and monitoring of mental health issues as part of an overall management plan. In the longer term, we will likely see digital mental health apps and programs supplement and even replace many of the tasks traditionally provided by human beings. How ethical and privacy issues are managed will be crucial to gaining acceptability and public trust.

## Endnotes


1 Twomey C & O'Reilly G (2017). Effectiveness of a freely available computerised cognitive behavioural therapy programme (MoodGYM) for depression: meta-analysis. *Australian & New Zealand Journal of Psychiatry*, *51*, 260–269.

2 Proudfoot J et al. (2013). Impact of a mobile phone and web program on symptom and functional outcomes for people with mild-to-moderate depression, anxiety and stress: a randomised controlled trial. *BMC Psychiatry, 13*, 312.

3 Eichstaedt JC et al. (2018). Facebook language predicts depression in medical records. *Proceedings of the National Academy of Sciences*, *115*, 11203–11208.

4 Reece AG & Danforth CM (2017). Instagram photos reveal predictive markers of depression. *EPJ Data Science*, *6*, 15.

5 Real-time conversational guidance and analytics. *cogito*. https://www.cogitocorp.com/

6 Automated Patient Engagement. Cogniant. https://www.cogniant.co/

7 Sophie app. http://www.sophie app.io/

8 Fitzpatrick KK et al. (2017). Delivering cognitive behavior therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): a randomized controlled trial. *JMIR Mental Healt*h, *4*, e19.

9 Your 4am friend and AI life coach. *Wysa*. https://www.wysa.io/

10 Fulmer R et al. (2018).Using psychological artificial intelligence (Tess) to relieve symptoms of depression and anxiety: randomized controlled trial. *JMIR Mental Health*, *5*, e64.

11 DeVault D et al. (2014). A virtual human interviewer for healthcare decision support. In *Proceedings of the 2014 International Conference on Autonomous Agents And Multi-Agent Systems* (pp. 1061–1068). International Foundation for Autonomous Agents and Multiagent Systems.

12 Koutsouleris N et al. (2018). Prediction models of functional outcomes for individuals in the clinical high-risk state for psychosis or with recent-onset depression: a multimodal, multisite machine learning analysis. *JAMA Psychiatry, 75,* 1156–1172.