Medicinska istaživanja 2022; 55(2):19-29 | DOI 10.5937/medi55-37718

OPEN ACCESS

Medical Research | Published by Faculty of Medicine University of Belgrade

REVIEW

Improvement of the psychiatric care through outsourcing artificial intelligence technologies – where are we now?

Sanja Andrić Petrović [1, 2], Nadja P Marić [1, 2]

[1] University of Belgrade, Faculty of Medicine, Belgrade, Serbia
[2] Institute of Mental Health, Belgrade, Serbia

Recived: 06 May 2022
Revised: 25August 2022
Accepted: 30 August 2022

Check for updates

Funding information:
The authors received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Competing interests:
The authors have declared that no competing interests exist

Correspondence to:
Sanja Andric Petrovic, MD, PhD, Clinical Teaching Assistant
Milana Kasanina 3, Belgrade 11000, Serbia
tel: +381638332290
e-mail: sanjandric@gmail.com

Summary

Currently, the world is entering the fourth industrial revolution – marked by artificial intelligence (AI) powered technologies. The growing ubiquity of AI technologies is already present in many sectors of modern society, but caution still prevails in medicine where their application is far from routine, although it is on the constant rise. Psychiatry has been recognized as one of the disciplines where significant contribution of AI technologies is expected for prediction, diagnosis, treatment and monitoring of persons with psychiatric disorders. Nearly half of the world's population live in countries that have fewer than one psychiatrist per 100 000 inhabitants, which is far below the health needs as the prevalence of psychiatric disorders is within the range of 10-20%. Thus, the question arises – whether AI technologies can help to fill the gap in unmet needs in psychiatry? The main types of autonomous technologies currently applied in psychiatry are machine learning and its subsets deep learning and computer vision, alongside natural language processing and chatbots. The present review will focus on the brief history of the concept, the utility of AI technologies in psychiatry, clinicians' attitudes, ethical dilemmas, clinical and scientific challenges. This review emphasizes that the psychiatric community should not be ignorant but could try to leave the comfort zone and do more to raise the awareness of AI technologies development achievements.

Key words: artificial intelligence; machine learning; psychiatry; mental health; ethics

Cite this article as: Andrić Petrović S, Marić N. Improvement of the psychiatric care through outsourcing artificial intelligence technologies – where are we now? Medicinska istraživanja 2022; 55(2):19-29 DOI: 10.5937/medi55-37718

INTRODUCTION

Currently the world is entering the fourth industrial revolution – marked by artificial intelligence (AI) powered technologies (1). Previous industrial revolutions mostly changed the way we work, while this one will fundamentally alter the way we think and who we are – our identity, sense of privacy, spending patterns, leisure time, the way we develop our careers, meet people and relate to one another. The fourth industrial revolution will mitigate or even eliminate the gap between people and technology, and it is considered that *"in its scale, scope and complexity, the transformation will be unlike anything humankind has experienced before"* (2).

The precise definition and meaning of the AI concept has been the subject of much discussion. It is commonly defined as a scientific discipline that studies and designs computers able to engage in human-like thought processes and behaviors such as learning, reasoning, planning and self-correction (3). What causes a lot of confusion in the existing literature on this topic is when the terms AI and machine learning (ML) are used interchangeably. However, AI is a broader concept referring to creation of autonomous machines that can simulate human thinking capability and behavior, whereas ML represents a subset of AI, increasingly applied in medicine, focused on developing computer systems capable of learning from data (experience) without being explicitly programmed. In other words, the program is not completely predetermined by the given code (i.e. a deterministic algorithm is an algorithm that given a particular input will always produce the same output), but the output can be adjusted based on the input data.

It is important to emphasize that AI is not one technology - but a set of technologies, among which ML, deep learning, natural language processing (NLP), virtual agents/chatbots and computer vision have the most promising applications in medicine. A detailed description of the mentioned AI technologies' technical characteristics exceeds the scope of this review article, which will rather focus on their application in the field of biomedicine - particularly psychiatry, which has been proven to be useful or promising. Psychiatry has been recognized as one of the disciplines where a significant contribution of AI technologies is expected both in the diagnosis and in the treatment of mental disorders. In addition, a brief history of the concept of AI will also be presented, as well as certain ethical dilemmas and challenges, and the application of existing knowledge in everyday practice.

A BRIEF HISTORY OF THE CONCEPT OF AI

The beginnings of the concept of AI date back to 1936 when the brilliant British mathematician and decoder Alan Turing, known as "the father of AI", constructed the so-called Turing machine – a special computer designed to solve a given problem by testing all possible combinations until reaching the correct solution (4). Two decades later, in 1956, the computer scientist John McCarty (who is also considered one of the "founding fathers" of AI) coined the term "artificial intelligence" which he defined as the science and engineering of making automatic computers that simulate the higher functions of the human brain (5).

In the early 1970s AI technologies gradually entered the field of biomedical research and diagnostics, particularly in certain disciplines such as ophthalmology and radiology (6). However, the rest of this decade is a period known as the "winter of AI" – when its development slowed down since previously set high expectations have not been met and funding for this area has been significantly reduced (7).

On the other hand, the 1980s were a period of incredible growth and technological innovation when interest in practical application of AI technologies was renewed. Digital technology has replaced the analog one, and computers and the internet have become the main drivers of social and industrial development (the Third Industrial Revolution began) (8).

The 1997 was historical year when the IBM's chess-playing supercomputer called *Deep Blue* defeated Garry Kasparov – the world chess champion at the time and one of the best chess players ever (9). That was an unprecedented shift in the development of AI technologies that paved the way for further successes.

CHALLENGES IN THE DEVELOPMENT OF AI TECHNOLOGIES

Although we are living in an age of rapid technological changes and development, there are still numerous technical barriers and challenges for the field of AI and related disciplines. Generally speaking, AI-powered technologies are good at things that seem hard for humans, such as solving complex mathematical problems, but they are still not good enough at those things that people find quite simple and natural – like facial recognition or understanding of natural language. Computer scientist Hans Moravec offered a possible evolutionary explanation of the paradox that the easiest things for humans are the hardest ones for AI technologies. In his famous book "Mind children" he stated: *"Encoded in the large, highly evolved sensory and motor portions of the human brain is a billion years of experience about the nature of the world and how to survive in it. The deliberate process we call reasoning is, I believe, the thinnest veneer of human thought, effective only because it is supported by this much older and much more powerful, though usually unconscious, sensorimotor knowledge"* (10). The skills that people find the simplest are the same ones that have been evolving for millions of years, whereby the things that peo-

ple find difficult could be difficult just because they are new. We have been thinking about chess strategy for just over a thousand years, but we have been learning about interaction with the environment since the time our ancestors were single-celled organisms.

AI TECHNOLOGIES IN MODERN MEDICINE

At the mention of AI technologies the first association is most likely some kind of technological innovation that will appear in the distant future, but the fact is that they are already operating all around us – although often behind the scenes. The growing ubiquity of AI technologies is already present in many sectors of modern society, but caution still prevails in medicine where its application is far from routine – given that the risks are also significantly higher than in other fields - but it is still on the constant rise. As the main advantage of AI technologies (particularly ML) is the possibility of rapid analysis of large amounts of data and identifying patterns and regularities in them, the medical disciplines that are currently applying such technologies most successfully are ophthalmology, oncology, dermatology and radiology, where the algorithms of this kind exhibit the same or even better performance as experienced clinicians in assessing images or spotting abnormalities and subtle differences that are otherwise not visible to the human eye (11). The Food and Drug Administration (FDA) recently authorized for the first time an autonomous AI-based diagnostic system designed to detect diabetic retinopathy and macular edema in primary care clinics, with the potential to help prevent vision loss in thousands of people with diabetes annually (12). Although the use of AI technologies in medical practice has increased in recent years, the field of mental health has been slower to accept these innovations, and at the moment they are not routinely applied. The main reason is that psychiatry represents a unique field requiring a more complex integration of cultural and psychosocial factors with medical comorbidities.

DESCRIPTION OF AI TECHNOLOGIES WITH POSSIBLE IMPLEMENTATION IN PSYCHIATRY

About 15% of the world's population lives with some form of mental disorder, with approximately half of them not receiving timely professional help due to treatment-related costs, insufficient number of mental health professionals and the stigma associated with mental illness (13). In recent years, mental disorders top the list of medical conditions in terms of the estimated spending – surpassing even cardiovascular diseases (14).

There are huge unmet needs in mental health care and psychiatry, so the question arises – whether AI technologies can help to bridge this gap? Globally, the median number of mental health workers (including psychiatrists, nurses, psychologists, social workers, and other clinicians) is 9 per 100 000 populations, but there is extreme variation (from below 1 in low-income countries to 72 in high-income countries) (13). Nearly half of the world's population lives in a country with below one psychiatrist per 100 000 (15), but more than half of the world's population owns a smartphone. The increasing use of smartphones and other AI-powered technologies has enabled patients to better monitor their own condition, but also improved their access to mental health care. AI technologies could provide alternative methods that are affordable, available, effective and evidence-based (16). Their application refers to the use of computerized techniques and algorithms for the prediction, diagnosis, treatment and monitoring of mental disorders. AI-powered technologies could have the potential to make significant changes in psychiatry by enabling mental health practitioners to re-define mental illnesses more objectively than currently done in DSM-5 and ICD-11 (17), to identify these illnesses at an earlier stage – when targeted interventions may be more effective (18), and to enable personalized treatment for patients (16).

The main types of AI technologies currently applied in psychiatry **(Figure 1)** are ML and its subsets deep learning and computer vision, alongside NLP and chatbots.

ML has led to numerous breakthroughs in the field of AI, resulting in the outstanding performance of computer systems that in certain domains even surpass human experts (19). Basically, ML represents the ability of a software to improve its performance through exposure to data and gaining experience. A typical ML model first discovers patterns and acquires knowledge from large data sets, and then applies that knowledge to predict future outcomes. For example, these techniques are used in predictive analytics – to anticipate future outcomes

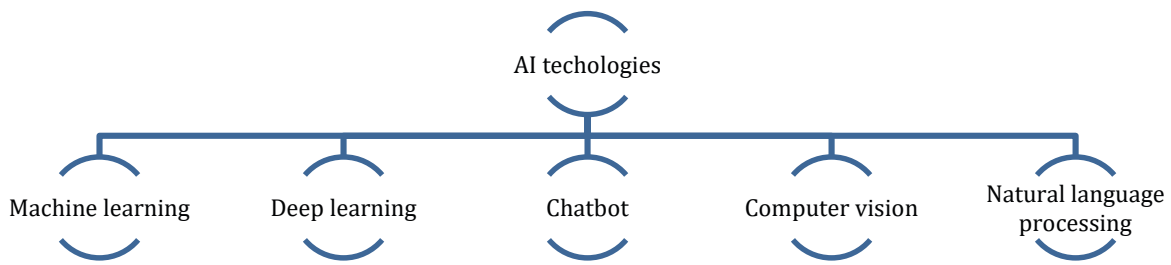

Figure 1. The main types of artificial intelligence (AI) technologies currently applied in psychiatry

based on available data (20). Unlike humans, whose ability to learn is limited by constitutional capacities, access to knowledge and lived experience, ML technologies are able to quickly analyze large amounts of data and spot certain regularities in them - often without *a priori* knowledge of that data, which is their major precedence. Ironically, the lack of classical human intelligence can be an advantage of ML technologies – so they have no prejudices and can detect certain patterns that humans would overlook, and they can also ignore some generally accepted knowledge that is not supported by data (21). Currently there are two main types of ML techniques - supervised and unsupervised. The supervised ML firstly uses the labeled datasets to train the algorithms (the model learns on the set of training data), after which the algorithm is tested on unlabeled data to ensure its accuracy in classifying the target variable (22). On the other hand, unsupervised ML is based on analyzing unlabeled data and discovering hidden patterns in them (unlabeled data are sorted into groups or patterns to identify their underlying structure) (22). Due to the growing availability of data pertaining to an individual's mental health status nowadays, ML technologies are being increasingly applied to improve the understanding of mental health disorders and assist clinicians for improved decision making process (23).

Deep learning is a subset of ML and one of its latest advances which transforms the data through layers of nonlinear computational processing units therefore providing a new paradigm to gain knowledge from complex data (for detailed description see reference 24). Deep learning programs use the so-called artificial neural networks - the computer programs intended to discover latent relationships in raw complex data. This technology has greatly improved the state-of-the-art in speech and visual object recognition, as well as many other domains, thereby bringing breakthroughs in processing images, video, speech and audio data (24).

Computer vision is another AI technology that uses artificial neural networks and ML principles to enable computers and systems to derive meaningful information from visual inputs (images, videos), and make recommendations or take actions based on such information. The model learns key features of the images (edges, curves, etc.) by imposing a grid-like structure on them, after which it is capable of recognizing similar objects. Computer vision can quickly surpass human visual capabilities as it is able to rapidly analyze large amounts of visual data and point imperceptible alterations or defects in them.

NLP is an AI-technology concerned with the computer programs intended to process and analyze large amounts of natural human language data. By extracting the word features (syntax, grammar, semantic meaning) it aims to comprehend human language from transcribed speech. Given that speech and text represent the main real-world data sources in psychiatry, natural language processing holds great potential in mental health research and care (22).

Chatbots (also known as virtual assistants or conversational agents) are AI-powered software programs that simulate human discussion by creating a bidirectional information exchange. Over the past two decades advances in NLP and deep learning have contributed to the development of sophisticated chatbots that are able to analyze verbal/textual inputs and respond appropriately using human audio or textual language (25).

AI TECHNOLOGIES IN PREDICTION OF MENTAL DISORDERS

The effectiveness of AI-based techniques in predicting various psychiatric disorders has been examined in recent years (18). That could be of great importance since the timely implementation of targeted interventions might possibly change the course and outcome of the disorder, and in the most favorable scenario – even prevent symptom exacerbation.

It is generally acknowledged that social cognitive deficits – such as impaired emotion recognition characterize various psychiatric disorders, among which schizophrenia, depression-anxiety and autism spectrum disorders were the most broadly studied to date (26-30). Growing evidence suggests that deficits in social cognition have an important role in the initiation and maintenance of the symptoms and underlie poor functional outcomes in patients with mental health impairments (31). What has been less studied in the literature so far is whether the impaired ability to recognize emotions can be an indicator of susceptibility or increased risk of developing certain mental disorders (the so-called trait-dependent risk marker). The results of a recent study showed that healthy individuals with higher levels of neuroticism – largely innate disposition to experience distress, anxiety and depression (trait-dependent risk marker associated with many forms of psychopathology), have poorer recognition of happiness on other people's faces. That finding indicates that biased processing of positive emotions might represent a useful marker of general susceptibility to psychopathology that could easily be obtained using computerized neuropsychological tests, thus enabling early detection of individuals at risk and timely application of preventive interventions (32).

Prediction of transition to psychosis in persons at clinical high risk has also become the cornerstone of modern preventive psychiatry and there is a need to formulate a more accurate prognostic estimate at the individual level (33). IBM has developed an automated ML speech classifier - comprising decreased semantic coherence and reduced usage of possessive pronouns, that had an approximately 80% accuracy in predicting psychosis onset in high-risk youths (34). This halved the false-negative rate

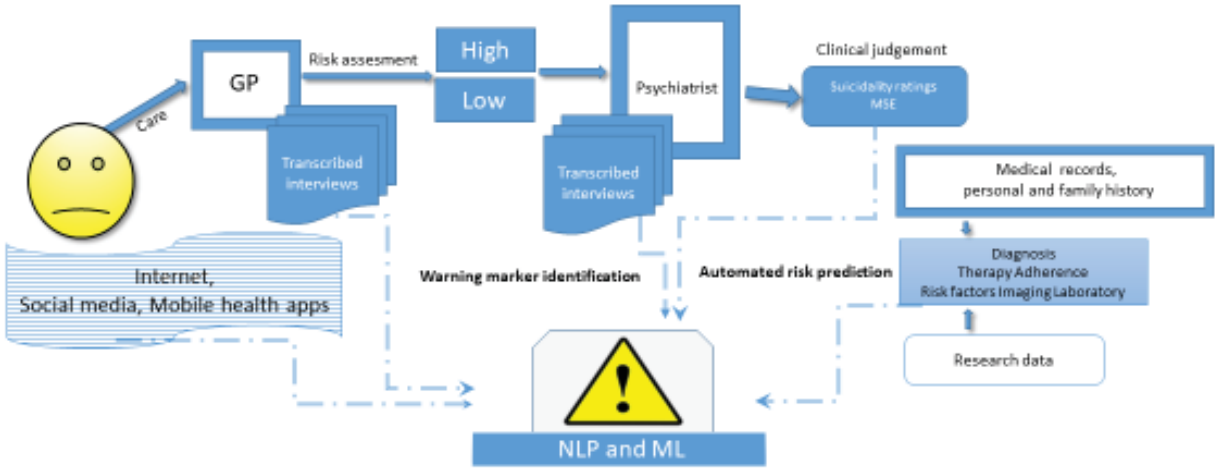

Figure 2. How AI technology might be leveraged to predict a suicide risk by analyzing linguistic patterns
(Abbreviations: GP - general practitioner; MSE – mental state examination; NLP - natural language processing; ML - machine learning; Apps – applications)

in comparison to the clinicians' assessment – which mainly relies on the clinical interview and the patients' motivation to accurately describe their (psychotic) experience.

Another important application of ML algorithms is in predicting suicide in individuals at risk (i.e. with previous suicide attempt or self-injury; see Figure 2). For example, ML algorithms based on linguistic and acoustic characteristics were able to differentiate psychiatric patients who are suicidal from those who are not, as well as from a healthy control group, with an accuracy of 85% (35). By applying ML to electronic health records it is possible to predict future suicide attempts in patients with the history of previous self-harm with the accuracy of 80% for the next two years, and in the next week with the accuracy of 92% (36).

Social media can also become a space where prediction and detection of mental disorders can take place outside the traditional framework. The rationale for that is the mere nature of social networks – where in real-time a huge amount of data that can be linked to the emotional state of the user is available. For example, since 2017 Facebook has been implementing a suicide risk screening around the world (except in the European Union due to stricter data privacy laws) that relies on deep learning algorithms. This new proactive approach scans all users' posts for patterns of suicidal thoughts, and when indicated contacts their friends or local emergency services. A recent study showed that applying these methods enabled prediction of future depression with considerable accuracy up to three months before its overt onset (37). The language predictors of depression have been found to include certain emotional (sadness), interpersonal (loneliness, hostility) and cognitive processes (preoccupation with the self, rumination) (37).

AI TECHNOLOGIES IN THE DIAGNOSIS OF MENTAL DISORDERS

Psychiatry generally lacks specific and reliable biomarkers indicating mental health disorders – that would enable accurate distinction between psychopathology and mental health. For decades and even centuries psychiatric diagnosis has largely relied on the patient's self-reported symptoms, medical history and clinical observations, often leading to subjective and insufficiently reliable symptom assessment (38). Despite recent advances in research in genetics and neuroimaging that are slowly paving the way for improved diagnostics, there is widespread disappointment with the overall pace of progress in detecting and treating mental disorders. AI technologies in psychiatric diagnostics implies the use of advanced computerized techniques and algorithms such as automated speech analysis and ML to assess the mental state of patients beyond self-reporting and clinical observation (39).

In recent decades, neuroscientists and practitioners have increasingly recognized the importance of computerized neurocognitive test batteries and their use has largely become a part of the everyday clinical practice of modern psychiatric centers, especially the academic ones. In such a way, objective assessment of cognition has become an important part of both psychiatric diagnostics and monitoring the course of the disorder, the effectiveness of therapy, and the neuro-rehabilitation progress (40, 41).

One of the promising AI-based approaches in the diagnosis of mental disorders is the use of latent semantic analysis (LSA) – an automated tool for NLP. The basic principle on which LSA relies is that words that are close in meaning will occur in similar pieces of sentences/text (the so-called distributional hypothesis) (42). By applying the LSA in combination with the analysis of speech structure the research team from the National Institute of Mental Health (USA) was able to detect very subtle deviations in speech between probands with schizophrenia, their

unaffected first-degree relatives and unrelated healthy controls and to successfully distinguish the mentioned groups from each other (43). Studies have also shown that ML and NLP algorithms trained to analyze the choice of words and their order in sentences can distinguish between genuine and elicited suicide notes with even greater accuracy than mental health professionals, indicating that they are able to detect signs of distress well (44).

With the aid of computer vision it has recently become possible to detect ADHD or autism spectrum disorder based on the analysis of a person's behavior from a video with high accuracy (of as much as 96%), which might significantly improve the diagnosis of these disorders (45). AI technologies could be a helpful complement to the clinical assessment of psychiatric disorders, which would reduce the number of false-positive and false-negative diagnoses.

The neuroscience technology company Cambridge Cognition recognized voice audios as a rich biomarker of both content (what is said) and acoustic speech properties (how it is said). They recently released the NeuroVocalix - an automated tool for verbal cognitive assessments that brings a fuller understanding of patient functioning through an additional layer of understanding to cognitive assessment (46). One of the major uses of this ML technology is suggested to be in the objective voice-based measurement of pain. The NeuroVocalix is designed to elicit and automatically analyze pain-related signals in human voice therefore providing objective means of measuring pain for use in clinical trials as a more accurate assessment of a drug's effects and eventually in clinical practice where it will support appropriate prescribing.

AI TECHNOLOGIES IN MENTAL HEALTH MONITORING

Early detection and prevention of relapse can have a significant impact on the outcome of psychiatric disorders. Various applications are available for active monitoring of mood, sleep, physical activity and other areas closely related to an individual's mental state, and in recent years the development of AI technologies in combination with different sensors and smartphone applications allowed improved monitoring of psychiatric patients in the community (47). For example, a number of outgoing calls, sent messages, variability in typing dynamics (errors, pauses), walking distance, rate of speech speed and voice modulation have been identified as good predictors of impending depressive or manic episodes (48).

Speech and language analysis have emerged as the two most useful applications of AI technologies in psychiatry, since speech and mental wellbeing are closely related. So monotonous speech may indicate depression, rapid pressured speech may be a sign of mania, and disorganized speech is characteristic of psychotic disorders. When these alterations are sufficiently pronounced – a clinician can spot them, but ML and NLP algorithms can be trained to recognize them even when they are too subtle for humans to spot them.

AI TECHNOLOGIES IN THE TREATMENT OF MENTAL DISORDERS

Overall, the treatment of mental disorders consists of pharmacological and non-pharmacological interventions. Both approaches have already been in the focus of the technology innovations with a lot of examples and ongoing debates. Just a few examples of possible interventions will be mentioned here, followed by the ethical considerations, since the use of treatments based on AI technologies is raising many ethical questions and dilemmas.

Non-pharmacological interventions

Emotion recognition technology finds its application in modern psychiatric practice. The pilot study conducted by the researchers from Stanford university showed that wearing the Google Glass, so-called "smart glasses" with a built-in camera associated with software that detects facial expressions and categorizes emotions, might help children with autism to better understand emotional states and facial expressions of others (49), which is one of core deficits associated with these disorders (26).

Woebot is another digital health technology created by clinical psychologist dr. Alison Darcy (50). It is a virtual assistant or chatbot that simulates conversations that patients have with their psychotherapists – asks the user about his/hers thoughts and mood, "listens" to how he/she feels, learns about him/her, and provides evidence-based psychotherapy based on cognitive-behavioral techniques (CBT). For example: "How are you feeling today?"; "I am panicking"; "Oh, no. I'm sorry. Breathe along with me for a minute, and then we'll talk more about it, OK?" (51). Users can reach the *Woebot* at any time and the chatbot promptly responds to their messages. Although experts agree that it is only a robot that cannot provide a human connection, they also agree that it has many advantages (50).

The advantages of chatbots are that they are affordable and easy to use. Furthermore, they make CBT more accessible to modern generations who are accustomed to being online all the time. Sessions do not have to be scheduled in advance, they are much more frequent than with a human therapist, and chatbots are always available for their users (51). The initial studies examining their effectiveness have yielded promising results. The first randomized controlled trial examining the preliminary efficacy of the *Woebot* for college students who self-identify as having symptoms of anxiety and depression showed that this therapeutic method appears to be a feasible, en-

gaging and effective way to deliver CBT that significantly reduced their symptoms after just two weeks of use. The participants reported that they did not feel condemned while "talking" with a chatbot, and they were more willing to express themselves and to share potentially unpleasant information about themselves without shame, which has been recognized as a significant advantage of virtual therapists (50).

Pharmacological interventions and hybrid-medicines

In 2017, upon the FDA approval of Otsuka's application for digital aripiprazole (ABILIFY MYCITE®) i.e. a combination of the medicine and medical device, psychiatry became the discipline with the first-ever digital drug. Aripiprazole is an antipsychotic drug which has been registered for schizophrenia and bipolar disorder, severe psychiatric conditions frequently associated with therapeutic non-compliance, whose course might be (more or less) improved by the antipsychotic medication. The rationale behind adding the ingestible sensor is to transmit a signal to the web-based portal for professionals and caregivers. When the drug-device combination is exposed to gastric acid in the stomach, the technology will allow for real-time information about medication ingestion and this method has been expected to increase medication adherence, to improve health outcomes and decrease the health care costs (52).

Upon the FDA registration, this innovation raised quite a few questions and ethical debates. For example, whether such a medicine improves medication adherence and improves quality of life was not sufficiently studied and how autonomy of the patients, their privacy and human rights will be protected was unresolved. Not surprisingly, the US approval was not followed by the European approval. In mid 2020, European Medical Agency (EMA) expressed "certain major objections" towards the registration of ABILIFY MYCITE® for schizophrenia and bipolar disorder "which cannot be fully addressed" (53). It could be that "schizophrenia was a difficult indication because patients whose symptoms include paranoia and delusions were likely to reject it", as it was stated in the comment published by "Nature" (54). However, it seems that digital medicines might prove more acceptable to promote adherence and improve the outcomes in non-psychiatric indications (diabetes mellitus, hypertension, hepatitis C virus, etc.).

ETHICAL CHALLENGES

In the field of the non-pharmacological interventions, virtual therapists are anonymous, immune to the prejudices that human therapists might have and more accessible, which could be seen as an advantage compared to human therapists. Also, there is no stigma associated with revealing psychiatric symptoms to a clinician, the cost-effectiveness could be better and so on (51). However, besides the potential advantages, AI-based techniques also have certain limitations. First of all, it is possible to encounter certain technical difficulties. Chatbots are not always able to understand the user and his/her intentions because they do not have their own mind but follow a predefined script. Some experts therefore suggest that they should always be used in collaboration with a human therapist – to ensure something doesn't get missed. Also, therapy chatbots do not include empathy of a human therapist, and for that reason patients are sometimes not sufficiently motivated to keep on their treatment. Another significant shortcoming is that most studies on the effectiveness of these interventions have been conducted by their developers, who have conflict of interest and personal financial gain from the results (39).

One of the important topics of the modern age is privacy, as tracking and sharing personal information has become an integral part of a new way of connecting with people. The so-called 'digital phenotyping' provides a novel, nontraditional route to yield inferences about patients' health status, but also presents a novel challenge to orthodox boundaries of traditional medical expertise (55). In 2018 the European Union enacted its 'General Data Protection Regulation' (GDPR) aimed at ensuring citizens have control of their data, and provide consent for the utilization of their sensitive personal information. However, this model was not accepted everywhere and some countries (for example the US) still have considerably weaker data privacy rules.

A fundamental requirement for the responsible use of AI technologies in biomedicine, including psychiatry, is that the prediction or diagnosis of mental disorders are accurate and reliable, and do not in any way increase the risk to patients (56). In 2017 it was reported that an estimated excess of 10 000 apps related to mental health was available for download. However, the majority of them have not been tested by randomized controlled trials, and many may even provide harmful 'guidance' to users (57). The problem of responsibility is very important, i.e. who is responsible for decisions and actions of AI technologies – their designers, users, or both? One of the crucial differences between intelligent machines and humans is that the former cannot accept responsibility for their own actions or have moral consequences. The machines will not experience humiliation, stress or pain associated with reprimands, loss of professional privileges, license or other legal sanctions.

AI-powered machines might be seen to have advantages over human therapists from the ethical point of view as they are objective, unbiased, have no personal problems that might interfere with their professional competence, and are not subject to other ethical pitfalls

such as inappropriate relationships with patients or their family members. By outsourcing some aspects of medical care to machine learning, physicians could be freed up to invest more time in higher quality face-to-face doctor-patient interactions (58). Although many are concerned about the lack of emotions and empathy in intelligent machines, some also claim that even though it is important to be sensitive to other people's emotions, it is not necessary for a machine to have its own emotions in order to act in accordance with ethical principles. The software could be properly trained to take into account someone's mental state and potential suffering when making a decision in a particular ethical dilemma (59). Being emotional can even interfere with the ability to make the right ethical decisions and act properly in crisis situations because people tend to be blinded by their strong emotions.

However, clinician's contextual, comprehensive assessment of a patient's condition cannot be easily replicated by AI technology. Technology-mediated care is biased due to several reasons, ranging from "decontextualization" to social biases and technical biases. According to a recent report by EU Steering Committee for Human Rights (60), certain restrictions on the use of personal health data in AI-driven analyses could interfere with essential data linkages. Also, social bias could appear by examining the outputs of AI systems for unequal distributions across demographic groups, or when existing biases in society have been learned by an AI system. Critical examination of existing inequalities must be taken into account in order to better control the aforementioned bias.

Last but not least, AI systems might be generating new forms of discrimination. The so-called "digital divide" is a huge ethical challenge - a phenomenon describing a gap between those who have access to digital technology (smart phones, computers, internet) and those who do not, which can cause inequality and deepen the existing gap in the community. Such a gap could be a so-called "digital Berlin Wall". This is why the United Nations considered a "global commitment for digital cooperation" as the key recommendation (61). Such recommendations could be seen as intrusive, and in parallel like as a very humanistic. The obvious reason for such global initiatives, such as the cooperation suggested by the UN, is the fact that in a large number of places around the world people in need for traditional mental health care still cannot access it (13).

PSYCHIATRISTS' OPINIONS ABOUT THE IMPACT OF AI ON PSYCHIATRIC PRACTICE

A recent global study (22 countries) explored mental health professionals' opinions on the future of psychiatry and the likelihood that future autonomous technologies - namely AI/ML, would replace them in performing some of their key responsibilities (62). Four major domains related to AI/ML were surveyed in a large group of professionals from 22 countries: 1. patient-psychiatrist interactions; 2. the quality of patient medical care; 3. the profession of psychiatry; and 4. health systems. Half of the respondents deemed that AI/ML usage will significantly change their job, but the majority was skeptical that it would be able to perform complex psychiatric tasks in the same way or better than human therapists. The respondents generally agreed that AI technologies could fully replace humans only in administrative tasks such as collecting and updating medical records and data consolidation. Only 4% of the respondents thought that AI technologies would make their work redundant, and the vast majority thought it was unlikely that AI/ML would ever be able to provide empathetic care in the same way or better than the average psychiatrist, to assess someone's mental status or aggressive behavior, or to formulate an individual treatment plan. The conclusion is that AI technologies should not be meant to replace doctors, but to complement and augment human capabilities through some kind of hybrid of the traditional way of working and improving psychiatric care by outsourcing the AI technologies - to help and enable doctors do their job in a more efficient and human way (62).

What could be a concern for health workers is the possibility for continuous technological progress to enable the creation of autonomous technologies that would surpass human beings in terms of social and intellectual abilities. An intriguing concept of a super-clinician, an autonomous system that would incorporate different AI technologies, has been recently proposed (56). It would use facial recognition technology to confirm a patient's identity, advanced sensors to observe and analyze nonverbal behavior (facial expression, voice modulation, etc.) and to assess certain parameters that the human eye cannot see – such as detecting changes in body temperature with infrared cameras. The system would be able to access and analyze in real time all available patient data from the electronic health record. It could be trained to use all known therapeutic approaches, and with the use of predictive analytics the intelligent system could know exactly how and when to apply the best treatment or intervention, thereby making it personalized. However, what is inherent for helper professions would be lost – empathy, authentic care and compassion, as well as physical contact through handshakes, comforting lying of a hand on the shoulder of a person who is suffering, and so on.

CONCLUSION

Much has been done so far in the implementation of AI/ML technologies in psychiatry and this review paper aimed to summarize these findings. Experts in the field are divided in their opinions – from the belief that catastrophic consequences for humanity can occur if there is

no control over the development of AI technologies, to the view that such an opinion is paranoid and that any form of their development should be supported due to its boundless potentials (63). We can not foresee which scenario is likely to emerge at this moment, but time will provide us with an epilogue of these debates. However, the assumption is that the role of the doctor can never be completely replaced and that the future of medicine will probably become a "team game" of humans and AI technologies. They will likely be outsourced to help and support clinical decision making, to enable doctors to focus on the human aspects of medicine instead of being stuck in administrative tasks, and allow them more time to truly communicate effectively with their patients and be more compassionate and humane to them (56). We agree with Blease et al. (62) that lacking adequate education on ML technology and its potential to impact the lives of patients, as well as many related ethical issues, future psychiatrists will be ill-equipped to steer mental health care in the right direction. The psychiatric community should not be ignorant but should leave their comfort zone and do more to raise the awareness of AI technologies development achievements among current and future specialists.

Conflict of interest

None to declare

Contributors

SAP analyzed the literature, searched for references and wrote the first draft of the manuscript. NM revised the manuscript, provided figures and other important intellectual content. Both authors wrote and approved the final version of the manuscript.

References

1. Xu M, David JM, Kim SH. The fourth industrial revolution: opportunities and challenges. International Journal of Financial Research 2018; 9(2): 90-5. doi:10.5430/ijfr.v9n2p90.
2. Schwab K. The fourth industrial revolution. NY: Currency; 2017.
3. Luxton DD. Chapter 1 – An introduction to artificial intelligence in behavioral and mental health care. In: Luxton DD (Ed.) Artificial intelligence in behavioral and mental health care. US: Academic press; 2016. p. 1-26.
4. Turing AM. Computing machinery and intelligence. Mind 1950; LIX (236): 433-60.
5. McCarthy J, Minsky ML, Rochester N, Shannon CE. A proposal for the Daartmouth summer research project on artificial intelligence, August 31, 1955. AI magazine 2006; 27(4). doi: https://doi.org/10.1609/aimag.v27i4.1904.
6. Moraru AD, Costin D, Moraru RL, Branisteanu DC. Artificial intelligence and deep learning in ophthalmology - present and future (Review). Exp Ther Med 2020; 20(4): 3469-73.doi: 10.3892/etm.2020.9118.
7. Kulikowski CA. Beginnings ofartificial intelligence in medicine (AIM): computational artifice assisting scientific inquiry and clinical art – with reflections on present AIM challenges. Yearb Med Inform 2019; 28(1): 249-56. doi: 10.1055/s-0039-1677895.
8. Duarte AYS, Sanches RA, Dedini FG. Assessment and technological forecasting in the textile industry: from first industrial revolution to the industry 4.0. Strategic Design Research Journal 2018; 11(3): 193-202. doi: 10.4013/sdrj.2018.113.03.
9. Campbel M, Hoane Jr AJ, Hsu FH. Deep blue. Artificial intelligence 2002; 134(1-2): 57-83.
10. Moravec H. Mind children: the future of robot and human intelligence. London, England: Harvard University Press; 1988.
11. Haenssle HA, Fink C, Schneiderbauer R, Toberer F, Buhl T, Blum A et al. Man against machine: diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists. Ann Oncol 2018; 29(8): 1836-42. doi: 10.1093/annonc/mdy166.
12. Abramoff MD, Lavin PT, Birch M, Shah N, Folk JC. Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices. NPJ Digit Med 2018; 1: 39. doi: 10.1038/s41746-018-0040-6.
13. World Health Organization, 2018. Mental health atlas 2017. Available at: https://www.who.int/publications/i/item/9789241514019.
14. Roehrig C. Mental disorders top the list of the most costly conditions in the United States: $201 billion. Health affairs 2016; 35(6): 35(6): 1130-5. doi: 10.1377/hlthaff.2015.1659.
15. https://apps.who.int/mediacentre/news/notes/2015/finances-mental-health/en/index.html
16. Graham S, Depp C, Lee EE, Nebeker C, Tu X, Kim HC, et al. Artificial intelligence for mental health and mental illnesses: an overview. Curr Psychiatry Rep 2019; 21(11): 116. doi: 10.1007/s11920-019-1094-0.
17. Cho G, Yim J, Choi Y, Ko J, Lee SH. Review of machine learning algorithms for diagnosing mental illness. Psychiatry Investig 2019; 16(4): 262-9. doi: 10.30773/pi.2018.12.21.2.
18. D'Alfonso S. AI in mental health. Current Opinion in Psychology 2020; 36: 112-7. https://doi.org/10.1016/j.copsyc.2020.04.005
19. Webb ME, Fluck A, Magenheim J, Malyn-Smith J, Waters J, Deschenes M et al. Machine learning for human learners: opportunities, issues, tensions and threats. Educational Technology Research and Development 2020; 69: 2109-30.
20. Peterson ED. Machine learning, predictive analytics, and clinical practice: can the past inform the present. JAMA 2019; 322(23): 2283-4. doi: 10.1001/jama.2019.17831.
21. Khanzode KCA, Sarode RD. Advantages and disadvantages of artificial intelligence and machine learning: a literature review. Iaeme publication, 2020. https://doi.org/10.34218/IJLIS.09.01.20.04
22. Renn BN, Schurr M, Zaslavsky O, Pratap A. Artificial intelligence: An interprofessional perspective on implications for geriatric mental health research and care. Front Psychiatry 2021; 12:734909. doi: 10.3389/fpsyt.2021.734909.
23. Su C, Xu Z, Pathak J, Wang F. Deep learning in mental health outcome research: a scoping review. Transl Psychiatry 2020; 10(1): 116. doi: 10.1038/s41398-020-0780-3.
24. LeCun Y, Bengio J, Hinton G. Deep learning. Nature 2015; 521:436-44.
25. de Cock C, Milne-Ives M, van Velthoven MH, Alturkistani A, Lam C, Meinert E. Effectiveness of conversational agents (virtual assistants) in health care: Protocol for systematic review. JMIR Res Protoc 2020; 9(3): e16934. doi: 10.2196/16934.
26. Andric S, Maric Bojovic N. How social cognition deficits affect psychopathology – a neuroscientific approach. Medical Youth, 2015; 66(2): 25-30.
27. Andric S, Maric NP, Mihaljevic M, Mirjanic T, van Os J. Familial covariation of facial emotion recognition and IQ in schizophrenia. Psychiatry Res 2016; 246: 52-7. doi: 10.1016/j.psychres.2016.09.022.

28. Andric Petrovic S, Jerotic S, Mihaljevic M, Pavlovic Z, Ristic I, Soldatovic I, Maric NP. Sex differences in facial emotion recognition in health and psychotic disorders. Cogn Neuropsychiatry 2019; 24(2): 108-22. doi: 10.1080/13546805.2019.1582411.
29. Shanok NA, Aron Jones N, Lucas NN. The nature of facial emotion recognition impairments in children on the autism spectrum. Child Psychiatry & Human Development 2019; 50: 661-7.
30. Velthorst E, Mollon J, Murray RM, de Haan L, Germeys IM, Glahn DC et al. Cognitive functioning throughout adulthood and illness stages in individuals with psychotic disorders and their unaffected siblings. Mol Psychiatry 2021; 26(8): 4529-43. doi: 10.1038/s41380-020-00969-z.
31. Fett AK, Viechtbauer W, Dominquez MD, Penn DL, van Os J, Krabbendam L. The relationship between neurocognition and social cognition with functional outcomes in schizophrenia: a meta-analysis. Neurosci Biobehav Rev 2011; 35(3): 573-88. doi: 10.1016/j.neubiorev.2010.07.001.
32. Andric S, Maric NP, Knezevic G, Mihaljevic M, Mirjanic T, Velthorst E et al. Neuroticism and facial emotion recognition in healthy adults. Early Interv Psychiatry 2016; 10(2): 160-4. doi: 10.1111/eip.12212.
33. Maric NP, Andric Petrovic S, Rojnic-Kuzman M, Riecher-Rossler A. Implementation of early detection and intervention services for psychosis in Central and Eastern Europe: Current status. Early Interv Psychiatry 2019; 13(5): 1283-8. doi: 10.1111/eip.12805.
34. Corcoran CM, Carrillo F, Fernandez-Slezak D, Bedi G, Klim C, Javitt DC et al. Prediction of psychosis across protocols and risk cohorts using automated language analysis. World Psychiatry 2018; 17(1): 67-75. doi: 10.1002/wps.20491.
35. Pestian JP, Sorter M, Connolly B, Bretonnel Cohen K, McCullumsmith C, Gee JT et al. A machine learning approach to identifying the thought markers of suicidal subjects: A prospective multicenter trial. Suicide Life Treat Behav 2017; 47(1): 112-21. doi: 10.1111/sltb.12312.
36. Walsh CG, Ribeiro JD, Franklin JC. Predicting risk of suicide attempts over time through machine learning. Clinical Psychological Science 2017; https://doi.org/10.1177%2F2167702617691560
37. Eichstaedt JC, Smith RJ, Merchant RM, Ungar LH, Crutchley P, Preotiuc-Pietro D et al. Facebook language predicts depression in medical records. Proc Natl Acad Sci U S A 2018; 115(44): 11203-8. doi: 10.1073/pnas.1802331115.
38. Sartorius N, Maric N. Koliko nam je kategorija u klasifikaciji psihijatrijskih poremecaja potrebno. Medicinski podmladak 2017; 68(2): 1-7.
39. Fakhoury M. Artificial intelligence in psychiatry. Adv Exp Med Biol 2019; 1192: 119-25. doi: 10.1007/978-981-32-9721-0_6.
40. Andric S, Maric NP. Computerized neuropsychological testing – a review of findings in schizophrenia. Engrami 2012; 34(3): 57-66.
41. Andric Petrovic S, Maric NP. Klasifikacija psihoticnih poremecaja - aktuelnosti. Engrami 2019; 41(1): 85-100/
42. Landauer TK, Dumais ST. A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. Psychological Review 1997; 104(2): 211-40. https://doi.apa.org/doi/10.1037/0033-295X.104.2.211
43. Elvevag B, Foltz PW, Rosenstein M, DeLisi LE. An automated method to analyze language use in patients with schizophrenia and their first-degree relatives. J Neurolinguistics 2010; 23(3): 270-84. doi: 10.1016/j.jneuroling.2009.05.002
44. Pestian J, Nasrallah H, Matykiewicz P, Bennett A, Leenaars A. Suicide note classification using natural language processing: a content analysis. Biomed Inform Insights 2010; 2010(3): 19-28. doi: 10.4137/bii.s4706.
45. Mengi M, Malhotra D. Artificial intelligence based techniques for the detection of socio-behavioral disorders: a systematic review. Archives of Computational Methods in Engineering 2021; https://doi.org/10.1007/s11831-021-09682-8. https://www.cambridgecognition.com/products/neurovocalix/
46. https://www.cambridgecognition.com/products/neurovocalix/
47. Lovejoy CA, Buch V, Maruthappu M. Technology and mental health: the role of artificial intelligence. Eur Psychiatry 2019; 55: 1-3. doi: 10.1016/j.eurpsy.2018.08.004.
48. Place S, Blanch-Hartigan D, Rubin C, Gorrostieta C, Mead C, Kane J et al. Behavioral indicators on mobile sensing platform predict clinically validated psychiatric symptoms of mood and anxiety disorders. J Med Internet Res 2017; 19(3): e75. doi: 10.2196/jmir.6678.
49. Daniels J, Schwartz JN, Voss C, Haber N, Fazel A, Kline A et al. Exploratory study examining the at-home feasibility of wearable tool for social-affective learning in children with autism. NPJ Digit Med 2018; 1:32. doi: 10.1038/s41746-018-0035-3.
50. Fitzpatrick KK, Darcy A, Vierhile M. Delivering cognitive behavioral therapy to young adults with symptoms of depression and anxiety using a fully automated conversational agent (Woebot): a randomized controlled trial. JMIR Ment Health 2017; 4(2): e19. doi: 10.2196/mental.7785.
51. Xu B, Zhuang Z. Survey on psychotherapy chatbots. Concurrency and Computation Practice and Experience 2020. https://doi.org/10.1002/cpe.6170.
52. Cosgrove L, Karter JM, McGinley M, Morrill Z. Digital Phenotyping and Digital Psychotropic Drugs: Mental Health Surveillance Tools That Threaten Human Rights. Health Hum Rights 2020; 22(2): 33-39.
53. https://www.ema.europa.eu/en/documents/withdrawal-letter/withdrawal-letter-abilify-mycite_en.pdf
54. Digital drug fortunes falter. Nat Biotechnol 38, 117 (2020). https://doi.org/10.1038/s41587-020-0425-3
55. Insel TR. Digital phenotyping: Technology for a new science of behavior. JAMA 2017; 318(13): 1215-6. doi: 10.1001/jama.2017.11295.
56. Luxton DD, Leigh Anderson S, Anderson M. Chapter 11 – Ethical issues and artificial intelligence technologies in behavioral and mental health care. In: Luxton DD (Ed.) Artificial intelligence in behavioral and mental health care. US: Academic press; 2016. p. 1-26.
57. Torous J, Weiss Roberts R. Needed innovation in digital health and smartphone applications for mental health: Transparency and trust. JAMA Psychiatry 2017; 74(5): 437-8. doi: 10.1001/jamapsychiatry.2017.0262.
58. Topol E. Deep medicine: how artificial intelligence can make healthcare human again. London: Hachette UK, 2019.
59. Awad E, Dsouza S, Kim R, Schulz J, Henrich J, Shariff A et al. The moral machine experiment. Nature 2018; 563(7729): 59-64. doi: 10.1038/s41586-018-0637-6.
60. Steering Committee for Human Rights in the fields of Biomedicine and Health. (2022, June 7). The Impact of Artificial Intelligence on the Doctor-Patient Relationship. https://rm.coe.int/inf-2022-5-report-impact-of-ai-on-doctor-patient-relations-e/1680a68859
61. https://www.un.org/en/un75/impact-digital-technologies
62. Blease C, Locher C, Leon-Carlyle M, Doraiswamy M. Artificial intelligence and the future of psychiatry: qualitative findings from a global physician survey. Digit Health 2020; 6:2055207620968355. doi: 10.1177/2055207620968355.
63. Müller, V.C., Bostrom, N. (2016). Future Progress in Artificial Intelligence: A Survey of Expert Opinion. In: Müller, V.C. (eds) Fundamental Issues of Artificial Intelligence. Synthese Library, vol 376. Springer, Cham. https://doi.org/10.1007/978-3-319-26485-1_33

TEHNIKE VESTAČKE INTELIGENCIJE ZA POBOLJŠANJE PSIHIJATRIJSKE PRAKSE - GDE SMO SADA?

Sanja Andrić Petrović [1, 2], Nadja P Marić [1, 2]

Sažetak

Svet je zakoračio u četvrtu industrijsku revoluciju koju obeležavaju tehnologije veštačke inteligencije (VI). Tehnologije VI su vrlo prisutne u brojnim sektorima savremenog društva, međutim u oblasti medicine preovlađuje oprez. Primena ovih tehnologija nije postala rutinska, ali jeste u konstantnom porastu. Psihijatrija je jedna od disciplina gde bi se mogao očekivati značajan upliv tehnologija VI. Oko polovine svetske populacije živi u zemljama sa jednim psihijatrom na 100 000 stanovnika, što je značajno ispod zdravstvenih potreba jer se prevalencija psihijatrijskih poremećaja kreće u rangu 10-20%. Stoga, postavlja se pitanje kako bi tehnologije VI mogle da pomognu stanovništvu i struci? Glavne tehnologije koje trenutno nalaze primenu u psihijatriji kao što su: mašinsko učenje i duboko učenje, kompjuterski vid, obrada prirodnih jezika i četbotovi, mogli bi da nađu primenu u detekciji, dijagnostici, tretmanu i daljem praćenju osoba sa psihijatrijskim smetnjama. U aktuelnom revijalnom radu biće predstavljeni kratak istorijat koncepta VI, trenutne mogućnosti primene tehnologija VI u psihijatriji, kao i stavovi kliničara, etičke dileme i izazovi za struku i nauku. Ovim radom se usmerava pažnja na postojeće informacije o dostignućima tehnologija VI u oblasti mentalnog zdravlja i bolesti, kako bi se lakše pratio razvoj oblasti i formirao aktivan odnos prema primeni u praksi.

Ključne reči: veštačka inteligencija; mašinsko učenje; psihijatrija; mentalno zdravlje; etika

Primljen: 06.05.2022. I **Revizija:** 25.08.2022. I **Objavljen:** 30.08. 2022

Medicinska istaživanja 2022; 55(2):19-29