



HAL
open science

Benefits, Implications and Ethical Concerns of Machine Learning Tools Serving Mental Health Purposes

Patrícia Gamboa, Cláudia Quaresma, Rui Varandas, Hugo Gamboa

► **To cite this version:**

Patrícia Gamboa, Cláudia Quaresma, Rui Varandas, Hugo Gamboa. Benefits, Implications and Ethical Concerns of Machine Learning Tools Serving Mental Health Purposes. 12th Doctoral Conference on Computing, Electrical and Industrial Systems (DoCEIS), Jul 2021, Costa de Caparica, Portugal. pp.285-294, 10.1007/978-3-030-78288-7_27 . hal-03685944

HAL Id: hal-03685944

<https://inria.hal.science/hal-03685944>

Submitted on 2 Jun 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License

Benefits, Implications and Ethical Concerns of Machine Learning Tools serving Mental Health Purposes

Patricia Gamboa^{1,2}, Cláudia Quaresma¹, Rui Varandas^{1,2}
and Hugo Gamboa^{1,2}

¹ Departamento de Física, LIBPhys, FCT- Universidade Nova de Lisboa, Portugal

² PLUX Wireless Biosignals S.A., Portugal

{pg.neves@campus.fct.unl.pt, q.claudia@fct.unl.pt, r.varandas@campus.fct.unl.pt,
hgamboa@fct.unl.pt}

Abstract. In recent years, Healthcare and Mental Health have been in the spotlight for the creation of tools using Artificial Intelligence (AI). Mental Health technologies have proliferated in the last years, from social robotics to self-help tools and internet-based intervention programs. In particular, machine learning algorithms have been embedded into applications, for example, to provide a probability or classification regarding risk behaviours (e.g., suicide prevention tools). This paper reflects on the use of AI applied in mental healthcare tools, to assist clinicians in diagnosis, differential diagnosis and monitoring of patients with brain-related disorders. Benefits regarding the use of these technologies in mental healthcare will be discussed as well as some limitations and ethical concerns.

Keywords: Mental Health, Artificial Intelligence, Machine Learning, Healthcare technologies

1 Introduction

Artificial Intelligence (AI) has been a research subject for decades. However, only recent advances in computer processing capabilities enabled a pronounced development of AI systems and technological areas such as Virtual Reality, Robotics and Wearable Sensing, supporting the rise of innovative solutions in Healthcare [1].

Healthcare-related technologies relying on digital and internet communication, i.e., eHealth, increased portability and flexibility in terms of time and location (e.g., teleconsultation) and cost-effectiveness [2], allowing typically out-of-reach patients or communities (e.g., chronically ill patients) to be more engaged in their healthcare [3]. In mental health, these technologies tackled some support seeking barriers (e.g., anonymity concerns) and increased awareness for mental health conditions, psychotherapy and psychiatric treatments [4], empowering patients' participation in therapeutic activities beyond appointment hours (e.g., [5]).

Mental health conditions are increasing worldwide, with a 13% rise in these conditions and substance use disorders reported between 2007-2017 [6]. Depression is a primary cause of disability and suicide is the second leading cause of death among young people [6]. New ways of promoting mental health or conducting interventions may be developed and pushed forward to meet the demand (e.g., remote care videocalls or chats, universally adopted in the COVID-19 pandemic [7]).

Solutions using AI can support patients' healthcare and assist mental health professionals identifying and monitoring mental health disorders (e.g., [8], [9]). Machine Learning (ML), a subtype of AI, enables computers to modify or adapt their actions to become increasingly more accurate (hence the word *learning*) (e.g., [10]). ML algorithms can learn from data and as many health-related tasks involve reading from certain type of data (e.g., numerical as in heart rate or image as in radiography), they can provide a probability or classification of having some disorder or condition.

AI can assist in the creation of solutions to enhance citizens' well-being and assist all stakeholders involved in healthcare. In this paper, we reflect on the use of ML for mental health purposes, highlighting its benefits while also unveiling some limitations and ethical considerations that future studies should contemplate.

2 Artificial Intelligence, Machine Learning and Mental Health

Social robots and mobile applications for various mental health disorders (e.g., [3], [4]) have been emergent technologies, partly due to advancements in different areas of AI such as Natural Language Processing (NLP). The study of human language to create intelligent systems able to comprehend, break down and separate significance from text and speech [11], played a major role in developing emotional chatbots.

Mobile applications present an opportunity to expand the availability and quality of mental health solutions [12] and promote well-being in cost-effective ways [13]. Though they are growing annually, with more than 10000 targeting anxiety, depression, or emotional well-being [14], their efficacy or effectiveness is largely unknown [15] as few undergo rigorous assessments through Randomized Controlled Trials (RCTs) [16]. In a review of psychosocial wellness and stress management apps, amongst the 1009 analysed, only 47 (4.66%) targeted individuals with psychological disorders and 21 (2.08%) had published feasibility and/or efficacy peer-reviewed studies [17]. Another review on apps targeting anxiety and depression, identified efficacy studies for only 10 apps (6.2%) in 162 stating an evidence-based theoretical framework [18]. Limitations to RCTs studies emphasise they should be independently conducted [15], include large samples [15], well-defined control groups [13], report effect sizes [17] and effectiveness over longer time periods [16].

Despite these limitations, mental health apps can support nonclinical populations by adopting a preventive approach [16]. Furthermore, they can be combined with traditional clinical support for psychotherapeutic homework goals (e.g., symptom tracking and skills practicing [13]) or therapeutic progress assessment [16]). They can be categorised based on their function: self-management, symptom tracking, skills-training, cognition improvement, social support, and passive data collection [19]. Recent comprehensive resources (e.g., PsyberGuide [16]) can overcome the lack of

quality control standards [17] and assist individuals searching for self-help tools or professionals identifying effective solutions to enhance clinical outcomes.

The following table (Table 1) presents few mental health apps that were developed with mental health professionals and integrate evidence-based principles, mostly Cognitive Behavioural Therapy (CBT) as this framework provides larger effects on multiple outcomes [20]. They seem appropriate when combined with professional support, for particular difficulties or disorders such as stress and anxiety (Woebot, Headspace, MoodKit) or Post-Traumatic Stress Disorder (PTSD Coach, PE Coach¹).

Table 1. Mobile apps for mental health purposes

	Self-management	In conjunction with professional support	Evidence-based (e.g., CBT)	RCT	Target
Woebot	✓ *	✓	✓	✓ [21]	a, b
Headspace	✓	✓	✓	✓ [22]	a, b, c
MoodKit	✓	✓	✓	✓ [23]	a, b
PTSD Coach	✓	✓	✓	✓ [24], [25]	d
PE Coach		✓	✓		d

CBT – Cognitive-Behavioural Therapy; RCT – Randomized Control Trial

a) Stress and Anxiety; b) Mood Disorders; c) Sleep; d) Post-Traumatic Stress Disorder

* Only for those experiencing mild stress or “feeling blue”

Some of these apps, such as the emotional chatbot Woebot, use AI, but specific reference to algorithms or features is not available [21]. Recommendations to improve the design, user experience, validation and efficacy of mental health apps can be found in [16]. Also, expectancy effects and individuals’ characteristics that might impact the outcomes should be considered and examined [27].

Other emergent technologies to support clinicians in the diagnosis and monitoring of patients with brain-related disorders [e.g., 9] rely on ML, which focuses on the learning and adaption of AI [10], so their actions are increasingly more accurate.

Machine learning techniques are distinguished in different types but, for the scope of this paper, we will only focus on supervised learning models (SL) and unsupervised learning (UL) models. SL models are built from examples which are labelled and based on a training set, so that the algorithm generalizes to respond correctly to different inputs of the same kind [10]. In other words, these learning algorithms (also known as *classifiers*) learn to identify similarities between inputs that belong to the same class with the purpose of classification [10]. Having this in mind, they are useful tools to enable the screening of large populations for mental health difficulties or disorders, assisting in diagnosis (e.g., [28]), predicting behaviours (e.g., [29]) or facilitating patients’ monitoring.

From studies using ML for mental health, conducted in the last 5 years, we highlight research with clinical purposes [8], [9], [29], [30].

In clinical assessment tasks performed by mental health professionals, the analysis and prediction of behaviours is often a complex task. By coupling automated speech analysis with SL techniques, researchers [30] were able to detect alterations in speech

¹ This app was developed exclusively for therapy and the study available is not a RCT [26].

data (semantics and syntax) to predict psychosis. ML techniques may be further explored to create new tools for the diagnosis, prognosis and in treatments' decision-making of patients with Schizophrenia [30].

The identification of suicide risk is also a common task in clinical assessment. Using SL models namely decision tree techniques, researchers [29] were able to predict with high levels of accuracy if a patient belongs to a suicide risk group (vs no-risk), with a low number of variables (gathered from a large pool of suicide risk assessment items). These techniques apply if-then rules, mutually exclusive and sequentially applied, in a top-down process [29] which enabled the identification of variables playing a part in the "risk status" [28]. A brief assessment tool to track and detect suicide risk is also feasible, which would benefit the recognition of suicide risk behaviour over time and more timely responses in managing these situations [29].

On the other hand, UL models are developed using unlabelled examples so that these algorithms identify inputs that have something in common and are categorized together [10]. With no information about the correct outputs, the algorithm must discern similarity between different inputs for itself [10] but it needs reasonable amounts of data to learn from, which also increases computational time and storage [10].

Unsupervised algorithms that employed Deep Learning (DL) - namely by applying clustering k-means receiving the features extracted from a Convolutional Neural Network followed by an autoencoder for dimensionality reduction - were used to help diagnose disorders using image-based data as input such as X-rays or histologic samples [31]. These algorithms are able to detect abstract patterns or trends from image data, subsequently allowing generation of predictions on new data [9]. As several psychiatric and neurological symptoms are best explained with network-level analysis of changes in physiological structures, these algorithms are relevant to research neuroimaging correlates of brain-related disorders [9] as they are able to extract discriminative features. Several diagnostic studies reviewed by [9] corroborate the application of DL techniques to data from structural MRIs to detect Alzheimer's Disease (AD) and Mild Cognitive Impairment (MCI) with high levels of accuracy [9] and also in conversion-to-illness studies - from MCI to AD [9]. Taken together, these studies highlight DL techniques as valuable in the diagnostic and prognostic of brain-related diseases using biomarkers or neuroimaging. They may also prove useful in patients' treatment, assisting clinicians in decision-making and improving long-term clinical outcomes of several disorders. Nonetheless, the identification of complex nonlinear relationships from functional brain data, needs larger samples to compensate for the great number of variables to be estimated [9].

3 Contribution to Applied Artificial Intelligence Systems

In the following section we highlight the main advantages of ML techniques applied to mental health solutions. Important aspects regarding the construction of AI systems in this area, namely ethical standards and constraints and limitations that future studies should overcome, are also mentioned.

3.1 Benefits of Machine Learning Tools in Mental Health

AI approaches with ML techniques can increase prevention and intervention in mental health. eHealth allows professionals to reach diverse populations, improve their trust and openness to clinical assistance [3], track relevant parameters or support self-monitoring tasks (e.g., [13],[16]). Exposure of sensitive or potentially stigmatizing information, such as through AI-emotional chatbots, is also facilitated, with patients reporting feeling of less embarrassment than in face-to-face appointments [3].

Tools with ML algorithms allow innovative diagnostic paradigms of brain-related disorders by offering other measures besides self-reported symptoms, enabling better timing in the identification of disorders [8], better distinction between them (e.g., [9]) and prediction of relevant clinical behaviours, such as suicidal [29] or psychotic symptomatology [30]. Compared to traditional techniques (manual approaches of data collection and data analysis), they are also rapid and less expensive, easily adapted to research/clinic environments [28]. Compared to statistical inference, traditionally used to perform data analysis in psychometrics, ML data analysis is model agnostic and primarily focused on prediction (instead inference), enabling generalization/replication of results to unseen data. Therefore, these methods present a complementary approach that can also be used for psychometric purposes pointing to deeper biologically meaningful findings [28].

SL techniques may be particularly well-suited to disentangle the role of personal, situation-specific and sociocultural variables on the onset, development, maintenance and remission of psychopathology [32]. On the other hand, DL techniques applied to large sets of image-based data, in particular, might allow the recognition of biomarkers relevant for identifying conversion-to-illness or discrimination amongst different stages of brain-related disorders [9].

3.2. Limitations and Ethical Considerations of Machine Learning Tools in Mental Health

Despite the positive outcomes these technologies bring to mental health, their misuse or ethical implications need to be taken seriously. The temptation to replace them by clinical services, further exacerbating existing health inequalities, should be prevented [3]. We highlight relevant aspects to take into consideration when developing tools with AI and ML for mental health purposes.

Systems built by multidisciplinary teams and incorporating diversity of views. Collaboration/training between clinicians and developers designing these technologies needs to be strengthened and a common language to describe digital technologies for health should be adopted [33]. Since AI algorithms require human involvement (e.g., defining features, pre-classifying training data, adjusting thresholds and parameters), biases also need to be acknowledged and properly addressed (i.e., the statement that algorithms will classify more ‘objectively’ should not be taken at face value [34]).

Systems built on scientific evidence. Tools to support mental health should be carefully designed, with evidence-based frameworks and well-planned RCTs (e.g., [3], [16]). Each disorder’s symptomatology should be carefully considered. In a study about the psychosis’ monitoring, adverse events or symptoms were largely neglected [35]. Algorithms may result in decision-making that lacks precision or damages the

patient (e.g., compulsory measures) when data misrepresents the reality. Therefore, rigorous data collection and analysis of data used to train the algorithms is needed to avoid the AI system reproducing that bias [36]. Data embedded into algorithms usually exclude contextual variables important in mental health assessment (e.g., interpersonal, socioeconomic or environmental variables [36]), hence, an effort should also be made to realize how to add this into systems.

Systems built with Regulations and Ethical Standards. Despite AI being embedded in systems for the last decades, the area is lacking strong ethical and regulatory frameworks [37]. Just as other medical devices, authors (e.g., [3]) advocate they ought to undergo scrutiny through rigorous risk assessment and regulatory oversight, partly due to biased algorithms that can exclude or harm in unintended ways yet contributing to social inequalities. Clear standards on confidentiality, privacy and data security by intelligent agents should be developed. Furthermore, AI in mental health applications should be bounded by the same ethical code of human mental health professionals. Though they seem to be aligned with the *principle of beneficence*, other principles (e.g., *principle of nonmaleficence*) need to be further explored, such as in cases of robots' malfunctions [3]. General guidelines initiatives for research in AI are emerging (e.g., Good AI), still not specifically related to mental health [3]. These are critical since the ethical considerations this area deals with are not transversal to other areas (e.g., compulsory measures, suicidal assessment). Early identification of ethical issues by researchers, and developers should be encouraged in the design and creation of the next generation of AI systems for mental health [3].

Systems with supervision and monitoring mechanisms. Systems should be supervised frequently by humans and mechanisms for rapid identification of bias and errors implemented (e.g., signals of attentive hacking or unauthorized monitoring [3]). Some AI algorithms, namely DL, lack transparency. The output of the algorithm rarely has any concrete sense of how or why a particular classification has been arrived at from inputs [36]. Indeed, the inputs themselves may be entirely unknown or known only partially [9] and so models can have an excellent performance despite using irrelevant features (e.g., orientation of the neuroimage) instead of relevant information. Their lack of an explicit model might also make it difficult to directly relate to existing biological knowledge [28]. Consequently, clarifying the inner process of an algorithm displaying the "*logic trails to show how any conclusion has been reached, so decisions can be scrutinized and challenged*" [38], leading to algorithms' transparency may, in the field of mental health, allow a deeper understanding of brain-related diseases.

Systems built with all stakeholders. Despite AI advances in mental health, key stakeholders such as patients, caregivers and families, are still excluded from the discussions about AI in mental health [35]. Indeed, stakeholder participation and social debates [38] are an ethical imperative for AI development. Therefore, including all stakeholders involved in AI solutions for mental health should be encouraged.

4 Conclusions

AI and ML techniques present great advantages for mental health purposes including the possibility to enrich our knowledge on the onset and development of brain-related disorders, the monitoring of clinically relevant symptomatology and the identification of at-risk patients (e.g., suicide or psychotic behaviours). Despite allowing cost-effective and timely solutions, AI systems in mental health should not be an excuse for reducing the provision care by trained mental health professionals [3]. Moreover, these solutions do not come without research challenges, ethical concerns and a general lack of understanding of societal implications (e.g., [39]). In order to overcome some of these challenges, AI systems applied to this sensitive area should be built by multidisciplinary teams and incorporate diversity of views, including all stakeholders in their creation [38]. Moreover, they should take into account scientific evidence, incorporate solid regulatory/ ethical standards, be supervised regularly, have mechanisms to signal possible inaccuracies or faults [3].

Further studies on the long-term effects of these technologies on individuals, societies and communities are also required. Ensuring the quality of life and respect for patients should be an essential aspect of innovative technologies applied to mental health. However, there has been a lack in recognizing concerns surrounding trust, privacy, and autonomy. Research on patient acceptance and treatment outcomes of most of disseminated AI applications in mental health fields is also lacking (e.g., [16],[3]). The impact of some of these technologies in human relationships, in human conditions and disorders or attitudes regarding caregiving, is still a grey area (e.g., [39]). For example, patients, specially from vulnerable groups, may become very attached to or dependent on these applications, altering social values and healthcare systems, changes in human communication and social interactions [3].

On the other hand, by possibly narrowing our understanding of disorders with applications tackling some acknowledged determinants but leaving behind the bio-psycho-social factors involved in mental health disorders, the widespread of AI systems in the area can contribute to exacerbate “trends of reductionism”, i.e., oversimplification of complex subjects, in mental health [3].

Therefore, it may be argued that while promising, these technologies require further research from a sound core of specialists from both areas - AI and Mental Health - working towards the mitigation of ethical concerns and long-term consequences regarding the integration of AI systems in the area of mental health and in including all stakeholders into the development and enhancement of these systems.

Acknowledgments. This work was partly supported by Fundação para a Ciência e Tecnologia, under PhD grants PD/BDE/150304/2019 and PD/BDE/150672/2020.

References

1. Lovejoy, C., Buch, V., Maruthappu, M. : Technology and mental health: The role of artificial intelligence. *European Psychiatry*, 55, 1--3 (2019). <https://doi.org/doi:10.1016/j.eurpsy.2018.08.004>

2. Oh, H., Rizo, C., Enkin, M., Jadad, A.: What is eHealth: a systematic review of published definitions. *Journal of medical Internet research*, 7, 1 (2005). <https://doi.org/10.2196/jmir.7.1.e1>
3. Fiske, A., Henningsen, P., Buyx, A.: Your Robot Therapist Will See You Now: Ethical Implications of Embodied Artificial Intelligence in Psychiatry, Psychology, and Psychotherapy. *Journal of Medical Internet Research*, 21(5), e13216 (2019). <https://doi.org/10.2196/13216>
4. Aboujaoude, E., Starcevic, V. (Eds.): *Mental Health in the Digital Age: Grave Dangers, Great Promise*. Oxford University Press (2015)
5. Taylor, C. B., Luce, K. H.: Computer- and Internet-Based Psychotherapy Interventions. *Current directions in psychological Science*, 12, 1, 18--22. (2003).
6. World Health Organization, <https://www.who.int/health-topics/mental-health>
7. Vieta, E., Pérez, V., Arango, C.: Psychiatry in the aftermath of COVID- 19. *Revista de psiquiatria salud mental*, 13, 2, 105--110 (2020). <https://doi.org/10.1016/j.rpsm.2020.04.004>.
8. Richter, T., Fishbain, B., Markus, A., Richter-Levin, G., Okon-Singer, H.: Using machine learning-based analysis for behavioral differentiation between anxiety and depression. *Scientific Reports: Nature Research*, 10 (2020). <https://doi.org/10.1038/s41598-020-72289-9>
9. Vieira, S., Pinaya, H., Mechelli, A.: Using deep learning to investigate the neuroimaging correlates of psychiatric and neurological disorders: methods and applications. *Neurosci. Biobehav. Rev.*, 74, 58--75 (2017). <https://doi.org/10.1016/j.neubiorev.2017.01.002>
10. Marsland, S. *Machine Learning: An Algorithmic Perspective*. CRC Press, NY: (2015)
11. Clark, K., Althoff, T.: How to help someone feel better: NLP for mental health. The Stanford Language Processing Group (2016), <https://nlp.stanford.edu/blog/how-to-help-someone-feel-better-nlp-for-mental-health/>
12. Chandrashekar, P.: Do mental health mobile apps work: evidence and recommendations for designing high-efficacy mental health mobile apps. *mHealth*, 4, 3 (2018). <http://dx.doi.org/10.21037/mhealth.2018.03.02>
13. Neary, M., Schueller, S.M.: State of the Field of Mental Health Apps. *Cogn Behav Pract*, 25(4), 531--537 (2018). <http://doi.org/doi:10.1016/j.cbpra.2018.01.002>
14. Baumel, A., Muench, F., Edan, S., Kane, J.M.: Objective User Engagement With Mental Health Apps: Systematic Search and Panel-Based Usage Analysis. *J Med Internet Res*. 25, 21(2019). <https://doi.org/10.2196/14567>
15. Lui, J. H. L., Marcus, D. K., Barry, C. T.: Evidence-Based Apps? A Review of Mental Health Mobile Applications in a Psychotherapy Context. *Professional Psychology: Research and Practice* (2017). <http://doi.org/10.1037/pro0000122>
16. Bakker, D., Kazantzis, N., Rickwood, D., Rickard, N.: Mental Health Smartphone Apps: Review and Evidence-Based Recommendations for Future Developments. *JMIR Ment Health*, 3, 1: e7 (2016). <https://doi.org/10.2196/mental.4984>
17. Lau, N., O'Daffer, A., Colt, S., et al.: Android and iPhone Mobile Apps for Psychosocial Wellness and Stress Management: Systematic Search in App Stores and Literature Review. *JMIR Mhealth Uhealth*. 22, 8(5), e17798 (2020). <http://doi.org/10.2196/17798>.
18. Marshall, J.M., Dunstan, D.A., Bartik, W.: Apps With Maps-Anxiety and Depression Mobile Apps With Evidence-Based Frameworks: Systematic Search of Major App Stores. *JMIR Ment Health*. 7(6), e16525 (2020). <http://doi.org/10.2196/16525>
19. National Institute of Mental Health. *Technology and the Future of Mental Health Treatment* (2017).
20. Linardon, J., Cuijpers, P., Carlbring, P., Messer, M., Fuller-Tyszkiewicz, M.: The efficacy of app-supported smartphone interventions for mental health problems: a meta-analysis of randomized controlled trials. *World Psychiatry* (2019) 18:325--36. <http://doi.org/10.1002/wps.20673>

21. Fitzpatrick, K.K., Darcy, A., Vierhile, M.: 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 Health*, 4(2), e19 (2017). <https://doi.org/10.2196/mental.7785>
22. Champion, L., Economides, M., Chandler, C. The efficacy of a brief app-based mindfulness intervention on psychosocial outcomes in healthy adults: A pilot randomised controlled trial, 13(12), e0209482 (2018). <https://doi.org/10.1371/journal.pone.0209482>
23. Bakker, D., Kazantzis, N., Rickwood, D., Rickard, N.: A randomized controlled trial of three smartphone apps for enhancing public mental health. *Behaviour Research and Therapy*. 109, 75--83 (2018). <https://doi.org/10.1016/j.brat.2018.08.003>
24. Possemato, K., Kuhn, E., Johnson, E., Hoffman, J.E., Owen, J.E., Kanuri, N., De Stefano, L., Brooks, E.: Using PTSD Coach in primary care with and without clinician support: a pilot randomized controlled trial. *Gen Hosp Psychiatry*. 38, 94--8 (2016). <http://doi.org/10.1016/j.genhosppsy.2015.09.005>
25. Kuhn, E., Kanuri, N., Hoffman, J.E., Garvert, D.W., Ruzek, J.I., Taylor, C.B.: A randomized controlled trial of a smartphone app for posttraumatic stress disorder symptoms. *J Consult Clin Psychol*. 85(3), 267--273 (2017). <http://doi.org/10.1037/ccp0000163>. PMID: 28221061.
26. Reger, G.M., Skopp, N.A., Edwards-Stewart, A., Lemus, E.L.: Comparison of Prolonged Exposure (PE) Coach to Treatment as Usual: A Case Series With Two Active Duty Soldiers. *Military Psychology*, 27:5, 287--296, (2015). <http://doi.org/10.1037/mil0000083>
27. Firth J, Torous J, Nicholas J, et al.: The efficacy of smartphone-based mental health interventions for depressive symptoms: a meta-analysis of randomized controlled trials. *World Psychiatry*, 16(3):287--298 (2017). <https://doi.org/10.1002/wps.20472>
28. Orrù, G., Monaro, M., Conversano, C., Gemignani, A., Sartori, G.: Machine Learning in Psychometrics and Psychological Research. *Frontiers in Psychology*, 10, 2970 (2020). <https://doi.org/10.3389/fpsyg.2019.02970>
29. Morales, S., Barros, J., Echávarri, O., García, F., Osses, A., et al.: Acute Mental Discomfort Associated with Suicide Behavior in a Clinical Sample of Patients with Affective Disorders: Ascertaining Critical Variables Using Artificial Intelligence Tools. *Frontiers in Psychiatry*, 8 (2017). <https://doi.org/10.3389/fpsy.2017.00007>
30. Corcoran, C.M., Carrillo, F., Fernández-Slezak, D., Bedi, G., Klim, C., et al.: Prediction of psychosis across protocols and risk cohorts using automated language analysis. *World Psychiatry*, 17, 1, 67--75 (2018). <https://doi.org/10.1002/wps.20491>
31. Xie, J., Liu, R., Luttrell VI, J., Zhang, C.: Deep Learning Based Analysis of Histopathological Images of Breast Cancer. *Frontiers in Genetics*, 10 (2019). <https://doi.org/10.3389/fgene.2019.00080>
32. Coutanche, M.N., Hallion, L. S.: Machine Learning for Clinical Psychology and Clinical Neuroscience. In Wright, A.G.C., Hallquist, M.N. (eds.), *The Cambridge Handbook of Research Methods in Clinical Psychology*. Cambridge, UK (2019). <https://doi.org/10.31234/osf.io/7zswH>
33. World Health Organization Classification of digital health interventions. Geneva: World Health Organization (2018).
34. Burrell, J.: How the machine 'thinks': Understanding opacity in machine learning algorithms. *Big Data & Society* (2016). <https://doi.org/10.1177/2053951715622512>
35. Carr, S.: 'AI gone mental': engagement and ethics in data-driven technology for mental health, *Journal of Mental Health*, 29, 2, 125--130 (2020). <https://doi.org/10.1080/09638237.2020.1714011>
36. Zuiderveen Borgesius, F.: Discrimination, artificial intelligence, and algorithmic decision-making. Council of Europe, Directorate General of Democracy (2018).

37. Gamble, A.: Artificial intelligence and mobile apps for mental healthcare: a social informatics perspective. *Aslib Journal of Information Management*, 72, 4, 509--523 (2020). <https://doi.org/10.1108/AJIM-11-2019-0316>
38. EU High-Level Expert Group on AI. Ethics Guidelines for Trustworthy AI, <https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai>
39. Cresswell, K., Cunningham-Burley, S., Sheikh, A.: Health Care Robotics: Qualitative Exploration of Key Challenges and Future Directions. *J Med Internet Res*, 20 (7): e10410 (2018). <https://doi.org/10.2196/10410>