HELDA
HELSINGIN YLIOPISTO
HELSINGFORS UNIVERSITET
UNIVERSITY OF HELSINKI

UNIVERSITY OF HELSINKI

https://helda.helsinki.fi

Mapping out the philosophical questions of AI and clinical practice in diagnosing and treating mental disorders

Uusitalo, Susanne

2020-09-30

Uusitalo , S , Tuominen , J & Arstila , V 2020 , ' Mapping out the philosophical questions of AI and clinical practice in diagnosing and treating mental disorders ' , Journal of Evaluation in Clinical Practice , pp. 1-7 . https://doi.org/10.1111/jep.13485

http://hdl.handle.net/10138/334736
https://doi.org/10.1111/jep.13485

acceptedVersion

Downloaded from Helda, University of Helsinki institutional repository.

This is an electronic reprint of the original article.

This reprint may differ from the original in pagination and typographic detail.

Please cite the original version.

Mapping out the philosophical questions of AI and clinical practice in diagnosing and treating mental disorders

Susanne Uusitalo, Jarno Tuominen & Valtteri Arstila

Abstract

How to classify the human condition? This is one of the main problems psychiatry has struggled with since the first diagnostic systems. The furore over the recent editions of the diagnostic systems DSM 5 and ICD-11 has evidenced it to still pose a wicked problem. Recent advances in techniques and methods of artificial intelligence and computing power which allows for the analysis of large data sets have been proposed as a possible solution for this and other problems in classification, diagnosing and treating mental disorders. However, mental disorders contain some specific inherent features which require critical consideration and analysis. The promises of AI for mental disorders are threatened by the unmeasurable aspects of mental disorders, and for this reason the use of AI may lead to ethically and practically undesirable consequences in its effective processing. We consider such novel and unique questions AI presents for mental health disorders in detail and evaluate potential novel, AI-specific, ethical implications.

Keywords: diagnosis, philosophy of medicine, medical ethics, progress

Introduction

Nearly one in seven people globally are afflicted by a mental health disorder.[1] This number is a likely underestimation of the true prevalence. The costs are not only limited to human suffering or decreased quality of life, but also constitute a considerable financial burden. Gustavsson and colleagues have calculated the direct and indirect costs of mental disorders in Europe to account for €462 billion.[2] In the United States the costs are estimated to be higher, in line with the generally higher health care expenditures in the US compared to Europe. Globally, the direct and indirect costs to account for $2.5 trillion in 2010 with an estimated increase to $6 trillion by the year 2030.[3]

Mental health disorders are treatable, and both human and financial costs can be mitigated. To better diagnose, specify and treat mental health disorders requires a multipronged approach ranging from governance and health care system to individual treatment providers. Artificial intelligence (AI) has been proposed to have a significant positive impact on these issues, similarly as it has been proposed to improve performance in other fields of health care.[e.g., 4,5] While we agree with this assessment, we also maintain that mental health differs from other health issues in important ways. Consequently, in order to utilise AI in the context of mental health successfully, some additional hurdles and special cases posed by features inherent in mental health classification, conceptualization and treatment need to be considered.

To explore the philosophical and ethical issues, we focus on three broad themes, namely detection, diagnosis, and treatment, leaving aside research and clinical administration even if they occasionally overlap with our themes. These themes have been prominent also in earlier reviews of AI technologies applied to mental health.[e.g., 6–9] However, as our considerations are philosophical and ethical in nature, they extend and differ to a degree with the main results of the existing literature. Specifically, we consider how these technologies raise novel and unique philosophical questions for detection, diagnosing and treatment of mental health disorders and further evaluate whether these carry novel ethical implications that are AI-specific and need be taken into account in implementation in this kind of setting.

Let us begin, however, by explicating what the technologies included in our considerations are, and why we think their use in mental health might raise novel ethical implications, in particular. First, what exactly is AI for mental health? In this article, we use the term AI broadly as consisting of all computer programs, or systems that include computer programs (e.g., humanoid robots), that have been developed to replicate the intelligent functions of humans (clinicians), and we consider those in relation to mental health care. This conception of AI is purposefully very general, as we do not want to unnecessarily exclude from our consideration the technologies that are likely to have a significant impact on mental health care. Thus, our conception includes various machine learning algorithms (such as supervised and unsupervised learning) that are often used for screening and diagnosing purposes. It also includes complex hybrid cognitive architectures (i.e., which combine symbolic and parallel distributed processing), likely to be more useful in the treatment and therapeutic contexts, and systems that are used either as augmenting the human decision making or in autonomous decision making in screening, diagnostics, and treatment.

The use of AI has surged in all areas of modern life, from Netflix recommendation systems to grading student essays, from predicting weather patterns to vehicles with autonomous driving systems. This has, nevertheless, occurred to a lesser degree in the (public) health care—this is not to say that it has not happened, and indeed it has significantly changed, say, the early phases of drug development.[e.g., 10,11] In fact, there is an ever-increasing number of studies and the systems under development show great promise in this sector too.[4,5,12] The highly publicised cases most often involve diagnoses (e.g., how AI can diagnose different types of cancers and eye-diseases or interpret X-ray images as well as an experienced physician). Moreover, the use of AI has also been successful in identifying and prioritizing patients at risk for developing diseases in the (near) future.

The reason why we think AI use in mental health might raise novel ethical implications, or at least stresses the old ones, is the nature of the “problem base” it concerns. Whereas the above-mentioned example of use of AI to detect cancer concerns with well-defined objective parameters (i.e., data is given in pictures, and diagnostic criteria and the treatment options for cancer are clear), this does not apply to many (most?) mental disorders. Quite the contrary, mental health has to do with decidedly subjective and social phenomena, which makes their detection, diagnosis and treatment less clear cut than those of more objectively defined health conditions, which, too, have their own challenges for instance regarding overdiagnoses

(Rogers & Mintzker 2016; Hofmann 2017). In what follows, we will next turn to explicate these problems first in the relation of diagnosis and then turn to treatment.

The challenges of detection and diagnosis for AI

Let us assume that we are able to conceptualise mental disorders in an AI friendly way. Presumably this would lead to adopting of various AI systems that can be used for diagnosing mental health issues. If the success of AI technologies in other areas is of any indication, this would change the detection and diagnosis of mental health dramatically. For example, the technologies would likely to be scalable so that more are more people could be screened cost-effectively. In principle, the diagnosis might improve too, as AI would not be prone to cognitive errors and, say, could base its assessment on data that is difficult for a human to utilise. For instance, AI systems can be taught to process different kinds of data sets, ranging from the medical history of subjects visiting hospitals to their (or some populations') use of social media, thus to improve the early detection of mental health related issues. This, in turn, could lead to either preventing mental disorders from reaching diagnostic thresholds altogether, or mitigating their development by timely and appropriate early interventions. AI could act as a form of "preventive medicine" in this respect far earlier than any attempt to seek help for symptoms takes place. It is worth noting that this is not merely wishful thinking. Quite the contrary, mechanisms for early detection of prodromal psychosis or the onset of a manic phase in Bipolar Disorder (BD) that allow us to prevent, mitigate, and manage the course of illness have already been developed. It is possible to track, for example, smartphone data and assess acoustic features of speech to predict the onset of a manic phase of BD.[e.g., 13]

Above we have illustrated reasons for developing these technologies. They, however, come with a recognised price in other contexts. That price may even rise in mental health context, as we suggest. In this section, due to the limited space we identify only three of them, but go through these three one by one.

To begin with, *the "essence" of mental ailments* is far from agreed upon in disciplines that look into their aetiology, symptoms and potential cure.[14] In fact, classifying mental disorder has been notoriously challenging and yet it has direct consequences on how people are able to live their lives. For instance, the diagnostic category of gender dysphoria (DSM-5) or gender identity disorder (F64 ICD-10) is problematic, as not every individual satisfying these criteria necessarily requires treatment or arguably suffer from mental disorders.[e.g., 15] Moreover, there are several culture-bound disorders in both ICD and DSM which highlight how context and culture affects the way psychological distress is presented and evaluated. Additionally, Hacking separates human kinds from natural kinds, where the former are specified by "looping effects".[16] In short, looping effects are at play when we create classifications which, when taken into use, change the people and behaviour they denote. Several mental disorder categories can be considered to entail looping effects. This interplay makes classification more problematic, as it cannot be detached from its cultural and socio-historical circumstances.

It is useful to compare this with the successful cases of the use of AI in other sectors of health care. Arguably, what is common to all or most of the latter cases is that they concern phenomena in which the items that AI takes into account are well-defined and measurable, and there are established criteria based on which to assess the success. For instance, when AI screens images for eye-diseases, we know what the nature of the relevant data is (pictures showing subjects' eyes, retina, and so forth) and we have classifications for various diseases. Although in some cases we might not have pre-established classifications (i.e., when we use unsupervised learning algorithms), the used data is still well-founded and measurable, and there is consensus as to the relevance (and soundness) of the classifications AI suggests.

This is not obviously so with many mental health issues. Quite the contrary, most of them are decidedly subjective—it is a person who suffers from depression or panic attacks—and the relevant data comes from their subjective reports. This assumes, of course, that the subjects are willing and able to report their subjective reality, which is not a given. If they are not, then the subjective reports might conflict, to some degree, with their external behaviour. How to interpret such situations, when it is uncertain whether the subjects are really suffering from mental disorders (and not merely exhibit a slightly deviating behavioural patterns, which we all do in different contexts). Adding to the complexity of finding automatic solutions for mental health diagnosis, many mental health issues include social components or are intimately related to the social environment subjects belong to. It is an open question how those can be measured and taken into account. However, this is not the case only for AI but for the scientific community and the society as a whole.

What is more, the data for mental health diagnoses comes in several forms, as people report their conditions differently and pay attention to different aspects. The data is typically incomplete in some way (e.g., it may include the number of social encounters, but describe their nature only superficially). Then again, if subjective reports of one's feelings and the nature of social encounters are ignored or dismissed as irrelevant, we risk missing the very central features of the phenomena in question and thus reducing it to something that they are not. We are not arguing that there are or can be no measurable features present in mental disorders, but we are questioning whether these objectively measured correlations suffice for diagnoses in mental health and, maybe even more importantly, for assessing individuals' treatment needs. Leaving aside salient features of subjective experience and social factors runs the risk of simplifying the categories to the extent that the phenomenon is misconstrued.

One particular place in which this problem could emerge relates to the above-mentioned point that AI could aid us in increasing the specificity of mental disorder categories too. Given the high comorbidity of mental disorders (approximately half of people present more than one disorder),[17,18] and the fact that diagnostic concepts in psychiatry rarely meet stringent validity criteria,[e.g., 19] it is likely that an AI classifier would carve diagnoses differently from the current tradition. While this would create obvious problems in the current treatment provision and service structure, it could help understand difficult problems such as differing or even negative treatment responses to certain treatments. Ideally, it would allow a step closer to Kraepelin's dictum "it is necessary to turn away from arranging illnesses in orderly well-defined groups, and to set ourselves the undoubtedly higher and more satisfying goal of understanding their essential structure".[20] However, the ethical concerns would be broad, as such approaches increase the likelihood of finding spurious patterns from noise, and would generate new categories on how to understand human suffering.

The previous considerations do not mean, of course, that it is altogether impossible to use AI to determine and use criteria for diagnosing mental disorders. Rather, the point is only to emphasise the difficulties in accomplishing those functions. But, it is also worth noting that once we finally succeed in it, and even though classification and standardisation of diagnostics has its assets, it also has its downsides: When we develop fixed classes and criteria the chosen features are highlighted and the practices become locked. This would not be a problem had we the right features, but if only a few relevant and salient aspects are chosen, this will lead gradually to bias. The uncontrolled and contested diagnoses and views of mental disorders may have been a problem in their own right,[21] but this kind of "wild practice" would have arguably evened out the biases and prevented strong biases from developing. AI uses standardised measures and this is likely to enhance the biases as the technology makes the detection, diagnosis and treatment more and more effective.

As a final point related to the criteria used for diagnosing mental disorders, we want to emphasise that the detection of early signs of brain change in imaging reinforces the idea that mental disorders are naturalised. The AI can probably detect finer nuances and make more precise readings of the images than a human being doing the same task. This has already been noted at least in cancer screenings.[22] Lynette Reid argues that the way in which the cancer is diagnosed also tells about the concept of disease in question. The critique of overdiagnoses typically involves this kind of dimension (see Rogers & Mintzker 2016, Hofmann 2017). The same applies here. The mental disorder becomes a naturalistic change in the brain structure or function and that may affect the way in which the disorder is understood. Equally problematically, it is known that emphasising the biological aspect of mental disorders reduces clinicians' empathy[23] and increases the attached stigma.[24] Both of these effects negatively affect treatment outcomes. In the worst case scenario, AI can reinforce these effects, or alternatively provide solutions for avoiding them.

The second central issue we want to highlight here concerns the fact that some of the AI technologies bring about *issues of privacy in detection and diagnoses*. Often they relate to some kind of mass surveillance for public health concerns and their use is balanced with the assumed benefits gained. The tracking of the movement of individuals in the age of COVID-19 illustrates the issue well—the benefits of tracking movement for the public health outweigh the harms for the individual in question. There are of course ethical challenges related to such surveillance, but by and large they are not novel and rather highlight old ones.

The detection of behaviour indicative of mental health conditions (pre-diagnosis) for example in social media, however, often deals with questions of privacy and surveillance for the individuals. What is ethically relevant is how much the individual in question can influence her exposure to risks in these kinds of practices without loss of social benefits that justify this kind of exposure (Rogers, Entwistle & Carter 2019). In modern-day preventive medicine such as screenings are typically optional: you may always opt out, but AI does not automatically provide this kind of option, as it does not typically require consent for data mining for detection purposes especially in "public domains" such as social media and discussion platforms. The different venues from which the data is collected or processed in do fall under different kinds of data regulations, such as in the European Union the General Data Protection Regulation. However, unless the social media service user actually is aware that the AI processes the public data for early detection, the consenter may easily or is even likely to miss this function. Indeed, the fact that users do not recognise how much individualising data they give away (even though they have consented to it) is one of the reasons why Evan Selinger and Woodrow Harzog have argued that face recognition methods should be banned.[25] This starkly contrasts with health care, where it is the duty of the staff to ensure that the patients have understood what they are consenting to.[26]

Finally, before we turn to discuss AI and the treatment of mental health, a final issue that we hinted above needs to be mentioned: The current development of AI systems for diagnosis purposes is likely to make the screening cost-effective at large scale and, if the issues related to diagnosis criteria can be met, the accuracy of the diagnosis can be expected to be at least comparable to those done by health care professionals. Furthermore, it can be argued that as the detection becomes finer and finer, the issue of classifying disorders as instances that require medical treatment increase even more even if some of these cases could turn out to be overdiagnoses, i.e. diagnoses of conditions that would not otherwise have become clinically significant (Reid, Carter, Hofmann & Rogers 2018). As a result, it is only to be expected that increased detection would further burden the treatment sector considerably, especially if the resources of helping and critical review of diagnoses do not grow at the same speed.

Of course, the projected situation is hardly untypical in the health care sector in general. What could be uncommon though, is the extent at which this might occur and its impact. On the one hand, it has been estimated that over 70 % of global population requiring treatment for mental health do not receive it, and those who do, receive it either too late or in arguably ineffective modalities, such as alternative or complementary medicine.[27] On the other hand, it has been suggested that long waiting times in mental health negatively affect treatment engagement, symptom reduction and functional recovery.[28,29] This means that tackling these health issues as early as possible would be particularly beneficial. What we are thus facing is a dilemma: in the not so distant future, presumably we can use AI to detect mental disorders better and earlier than before. Yet, before AI-based treatment options are equally well-developed, this could only make the future treatment more difficult and thus discourage us from deploying those technologies.

Philosophical and ethical considerations of AI application for treatment

Diagnosis is the beginning of the treatment, but before considering the potential of AI in the means of treatment, it is worthwhile to focus on the outcome of diagnosing, i.e. the diagnosis itself. Does it matter if a health care professional or an AI makes the diagnosis? As discussed above, it may well be that AI is more reliable in taking into account all the symptoms and scanning through the other possibilities and it is not prone to cognitive errors or struggle with boredom or being tired.[30] At the same time, it seems to lack the beginning of the therapeutic relationship with a health care person. It is not a relationship in a strict sense, as it is not a relation between individuals. Nonetheless, this need not be a negative thing. In fact, it may be beneficial for individuals who struggle with human contact or are afraid of being judged and stigmatised. It should, however, be kept in mind that social withdrawal is a core feature of various mental disorders and is in general often considered to exacerbate the issues. One of key concerns is how to avoid increasing social isolation and withdrawal.

More significantly, as it concerns most people who have been impacted by mental health issues, the trust people have for an AI system to make the diagnosis may have two sides. The reliability of its ability to scan through great amounts of data may be there, but at the same time it does not have that “touch” that professionals could be argued to have stemming from years of experience in practice. The AI might not possess this kind of ability to detect the “x-factor”, something that is hard to pinpoint and articulate, but it is certainly something that we immediately recognise when we come across it. Of course, not all health care personnel have this kind of x-factor and the diagnoses the AI produces are all done in the same way which can be further seen as making the health care system more reliable and trustworthy. The other side of the coin here could be that the reliability of the technology undermines or threatens the prestige of the personnel. This kind of polarisation requires further premises to become notable, as it is not self-evident that this kind of juxtaposition takes place. The AI system and the personnel can just as easily be regarded as features of the same institution, both of which contribute to the reliability and trustworthiness of the providers of treatment.

The famous dodo-bird verdict states that the actual psychotherapeutic framework has a smaller effect than the so-called common factors of psychotherapeutic treatments have. The common factors include therapeutic alliance, empathy, goal consensus and collaboration, positive regard and affirmation, mastery, congruence or genuineness, mentalization and emotional experience.[31] Most, if not all, of these features are very hard to maintain in non-human interactions. Furthermore, the personal relationship aspect of mental health treatment goes even beyond the psychotherapeutic setting. A re-analysis of a broad US’s National Institute of Mental Health antidepressant study found the effect of the prescribed drug to be lower than the effect of who was doing the prescription, that is, the psychiatrist.[32] With regard to AI, this raises the question of what exactly is the treatment? Is it possible that in developing treatment providing AI we focus on an aspect of treatment that is in fact doing little work, and fail to consider the human contact aspect of successful treatment? While effort has mostly concentrated on creating intelligent pattern recognition AI’s or copying specific manualized therapeutic techniques (for instance from Cognitive Behavioral Therapy), there are also alternatives that have attempted to account for the aspect of physical presence. One example of such is the PARO, the therapeutic seal, which has been pilot tested in dementia homes and with people suffering from chronic pain.[e.g., 33]

In counselling, AI agents have also been introduced in mental health settings. These agents can be virtual such as avatars in virtual space or robots,[30,34] and they raise issues regarding for instance the therapeutic relationship that is regarded as one the most important aspects in effective therapy, as mentioned above. There are several other issues, too. First of all, there is the quality of the relationship. If the patient does not falsely believe that the AI agent is a human being (which is another challenge[30]), the patient may, on the one hand, feel that the AI agent cannot feel shame or other feelings and thus fails to understand the patient's ordeals. On the other hand, the patient, however, may find it easier to confide to an AI agent just because it is an AI agent and not a human being.

Also, virtual reality treatments have been developed as adjunct treatments for conventional counselling and psychotherapy for phobias,[35–38] eating disorders,[39] and post-traumatic stress disorder.[40,41] These have usually taken the form of Virtual Reality Exposure Therapy (VRET), and constituted, for example, a virtual Afghanistan or Iraq in treating military personnel with PTSD.[42] Treatments that can be offered regardless of location and material limitations enable access to mental health services for individuals previously deprived of counselling or therapy due to their remote location, limited mental health care resources available or other reasons such as shame.[e.g., 7]

The utilisation of mobile and sensor data to detect changes in behaviour indicative of mental health conditions is one of the functions of AI technologies such as machine learning.[8,34] Questions of epistemic authority arise when the reports of the individual are in conflict with the sensory data. This is important especially in mental health context as mental disorders typically cast doubt to the reliability of the individual to be competent to report such issues. In some disorders such as pathological gambling, it is a diagnostic criterion in the DSM-5 categorization that the individuals feel the need to hide their actions (and potentially are ready to lie about them). The articulated concern of ours realises once again: is it more likely that the assessment of the state of the individual's condition relies on the measurable and naturalised data provided by the AI technology rather than the person reporting their experience of how well rested or stressed they currently are. Also, the questions of responsibility rise, as the health care professionals are not necessarily keen on taking more responsibility of following this kind of data/supervising this kind of system. There is a risk of the situation converging according to the Goodhart's Law "when a measure becomes a target, it ceases to be a good measure". For example, using the common Beck Depression Inventory as a way to evaluate treatment efficacy instead of severity assessment may itself alter the treatment towards the items, which are easier to change (e.g. eating behaviour) than other more serious symptoms (e.g. suicidality, hopelessness). Similarly, this weighing of symptoms is something that is quite transparent for humans, but could prove wickedly opaque for AI. Overall, this brings us back to the question of how to diagnose mental disorders: In so far as the measurable data that AI uses for diagnosis purposes does not cover all relevant information, there is the threat that measuring the success with such AI system will lead to biases and new problems—resembling the curious and problematic case in which the behavior of psychiatric patients changed dramatically in almost overnight when Hospital Romero in Buenos Aires began to diagnose them with DSM-classification.[43]

Machine learning has been applied to provide personalised and timely treatment or interventions.[8] The promises of customised treatment with AI systems ease the challenges posed, for instance, by different cultural and ethnic backgrounds.[7] The objective of this kind of treatment is filled with good intentions: to provide the best care that particular individual and steer away from the idea that one model fits all,[44] particularly in mental health in which there are all kinds of environmental and social factors as well as the aspects of the individual which affect the disorder and consequently the treatment. Personalised medicine relies on biomarkers and other naturalised factors and, in this respect, customised or personalised mental health treatment runs the risk of reducing the condition to the biological or neurological variables. As we have suggested above, this is hardly the best way to construe mental disorders.

Furthermore, in the times of the COVID-19 pandemic, it could be asked what do the AI systems take to be the framework in producing and processing data for treatment. The societies have undergone dramatic changes in a couple of months' period and the patients are facing new challenges in their everyday lives. Are the systems flexible enough to accommodate a whole new infrastructure to the behaviours? For example, some behaviours necessitated by pandemics, such as physical distancing, are also symptomatic of various mental disorders during standard times. Due to the way mental disorders are conceptualized as non-normative behaviours and ways of being, there is a risk that AI would not accurately react to a sudden contextual shift.

Concluding remarks

The promises of AI in mental health lay in increasing timely assessment and access to treatment. For instance, AI can form a part of initial screening and act as an adjunct assessment. Machine learning could ideally optimise diagnostic criteria, and—in the best scenarios due to their scalability—AI systems could make treatment (especially therapies) more accessible to those who would benefit from them. We have highlighted the point that in various occasions such promises are threatened by the unmeasurable (or difficult-to-measure) aspects of mental disorders, and for this reason the use of AI may lead to ethically and practically undesirable consequences in its effective processing.

In this article we have not emphasised the issues of just health care system, for example equal access and just allocation of resources, but this marginality is merely due to the limitations of space, not due to lack of importance. We have focused solely on the AI issues connected to the clinical practice. Moreover, most of the ethical issues of AI in mental health raise issues that have been discussed before in different domains. The most notable difference between human-provided mental health care and that of AI seems to be the question of responsibility. Cases such as, for instance, malfunctions are an issue more controversial in relation to AI than to health care personnel with many codes of conduct that provide professional ethics to the practitioners.[e.g., 45]

References

1. Institute for Health Metrics and Evaluation (IHME). Global health data exchange. http://www.healthdata.org/. Published 2018.

2. Gustavsson A, Svensson M, Jacobi F, et al. Cost of disorders of the brain in Europe 2010. *Eur Neuropsychopharmacol*. 2011;21(10):718-779. doi:10.1016/j.euroneuro.2011.08.008

3. World Health Organization. *Global Status Report on Noncommunicable Diseases 2010*. Geneva; 2011.

4. Topol E. *Deep Medicine: How Artificial Intelligence Can Make Healthcare Human Again*. New York: Basic Books; 2019.

5. Yu K-H, Beam AL, Kohane IS. Artificial intelligence in healthcare. *Nat Biomed Eng*. 2018;2(10):719-731. doi:10.1038/s41551-018-0305-z

6. Fiske A, Henningsen P, Buyx A. Your robot therapist will see you now: Ethical implications of embodied artificial intelligence in psychiatry, psychology, and psychotherapy. *J Med Internet Res*. 2019;21(5):1-12. doi:10.2196/13216

7. Luxton DD. An Introduction to Artificial Intelligence in Behavioral and Mental Health Care. In: Luxton DD, ed. *Artificial Intelligence in Behavioral and Mental Health Care*. Elsevier; 2016:1-26. doi:10.1016/B978-0-12-420248-1.00001-5

8. Shatte ABR, Hutchinson DM, Teague SJ. Machine learning in mental health: A scoping review of methods and applications. *Psychol Med*. 2019;49(9):1426-1448. doi:10.1017/S0033291719000151

9. Silverman BG, Hanrahan N, Huang L, Rabinowitz EF, Lim S. Artificial Intelligence and Human Behavior Modeling and Simulation for Mental Health Conditions. In: Luxton DD, ed. *Artificial Intelligence in Behavioral and Mental Health Care*. Elsevier; 2016:163-183. doi:10.1016/B978-0-12-420248-1.00007-6

10. Murphy RF. An active role for machine learning in drug development. *Nat Chem Biol*. 2011;7(6):327-330. doi:10.1038/nchembio.576

11. Wale N. Machine learning in drug discovery and development. *Drug Dev Res*. 2011;72(1):112-119. doi:10.1002/ddr.20407

12. Hosny A, Parmar C, Quackenbush J, Schwartz LH, Aerts HJWL. Artificial intelligence in radiology. *Nat Rev Cancer*. 2018;18(8):500-510. doi:10.1038/s41568-018-0016-5

13. Cummins N, Matcham F, Klapper J, Schuller B. Artificial intelligence to aid the detection of mood disorders. Barh D, ed. *Artif Intell Precis Heal*. 2020:231-255. doi:10.1016/b978-0-12-817133-2.00010-0

14. Canino G, Alegría M. Psychiatric diagnosis – is it universal or relative to culture? *J*

Child Psychol Psychiatry. 2008;49(3):237-250. doi:10.1111/j.1469-7610.2007.01854.x

15. Reed GM, First MB, Kogan CS, et al. Innovations and changes in the ICD-11 classification of mental, behavioural and neurodevelopmental disorders. *World Psychiatry*. 2019;18(1):3-19. doi:10.1002/wps.20611

16. Hacking I. The looping effects of human kinds. In: Sperber D, Premack D, Premack AJ, eds. *Causal Cognition*. Oxford: Oxford University Press; 1996:351-383. doi:10.1093/acprof:oso/9780198524021.003.0012

17. Maj M. 'Psychiatric comorbidity': an artefact of current diagnostic systems? *Br J Psychiatry*. 2005;186(3):182-184. doi:10.1192/bjp.186.3.182

18. Roca M, Gili M, Garcia-Garcia M, et al. Prevalence and comorbidity of common mental disorders in primary care. *J Affect Disord*. 2009;119(1-3):52-58. doi:10.1016/j.jad.2009.03.014

19. Jablensky A. Psychiatric classifications: Validity and utility. *World Psychiatry*. 2016;15(1):26-31. doi:10.1002/wps.20284

20. Jablensky A. The nosological entity in psychiatry: a historical illusion or a moving target? In: Kendler KS, Parnas J, eds. *Philosophical Issues in Psychiatry II*. Oxford: Oxford University Press; 2012:77-94. doi:10.1093/med/9780199642205.003.0014

21. Stengel E. Classification of Mental Disorders. *Bull World Health Organ*. 1959;21:601-663.

22. Reid L. Truth or spin? Disease definition in cancer screening. *J Med Philos*. 2017;42(4):385-404. doi:10.1093/jmp/jhx006

23. Lebowitz MS, Ahn W. Effects of biological explanations for mental disorders on clinicians' empathy. *Proc Natl Acad Sci*. 2014;111(50):17786-17790. doi:10.1073/pnas.1414058111

24. Angermeyer MC, Holzinger A, Carta MG, Schomerus G. Biogenetic explanations and public acceptance of mental illness: Systematic review of population studies. *Br J Psychiatry*. 2011;199(5):367-372. doi:10.1192/bjp.bp.110.085563

25. Selinger E, Hartzog W. The Inconsentability of Facial Surveillance. *Loyola Law Rev*. 2019;66:101-122.

26. Appelbaum PS. Assessment of Patients' Competence to Consent to Treatment. *N Engl J Med*. 2007;357(18):1834-1840. doi:10.1056/NEJMcp074045

27. Wang PS, Aguilar-Gaxiola S, Alonso J, et al. Use of mental health services for anxiety, mood, and substance disorders in 17 countries in the WHO world mental health surveys. *Lancet*. 2007;370(9590):841-850. doi:10.1016/S0140-6736(07)61414-7

28. Doyle R, Turner N, Fanning F, et al. First-Episode Psychosis and Disengagement From Treatment: A Systematic Review. *Psychiatr Serv*. 2014;65(5):603-611. doi:10.1176/appi.ps.201200570

29. Penttilä M, Jääskeläinen E, Hirvonen N, Isohanni M, Miettunen J. Duration of

untreated psychosis as predictor of long-term outcome in schizophrenia: systematic review and meta-analysis. *Br J Psychiatry*. 2014;205(2):88-94. doi:10.1192/bjp.bp.113.127753

30. Luxton DD. Recommendations for the ethical use and design of artificial intelligent care providers. *Artif Intell Med*. 2014;62(1):1-10. doi:10.1016/j.artmed.2014.06.004

31. Wampold BE. How important are the common factors in psychotherapy? An update. *World Psychiatry*. 2015;14(3):270-277. doi:10.1002/wps.20238

32. McKay KM, Imel ZE, Wampold BE. Psychiatrist effects in the psychopharmacological treatment of depression. *J Affect Disord*. 2006;92(2-3):287-290. doi:10.1016/j.jad.2006.01.020

33. Lane GW, Noronha D, Rivera A, et al. Effectiveness of a social robot, “Paro,” in a VA long-term care setting. *Psychol Serv*. 2016;13(3):292-299. doi:10.1037/ser0000080

34. Luxton DD, June JD, Sano A, Bickmore T. Intelligent Mobile, Wearable, and Ambient Technologies for Behavioral Health Care. In: Luxton DD, ed. *Artificial Intelligence in Behavioral and Mental Health Care*. Elsevier; 2016:137-162. doi:10.1016/B978-0-12-420248-1.00006-4

35. Parsons TD, Rizzo AA. Affective outcomes of virtual reality exposure therapy for anxiety and specific phobias: A meta-analysis. *J Behav Ther Exp Psychiatry*. 2008;39(3):250-261. doi:10.1016/j.jbtep.2007.07.007

36. Opriş D, Pintea S, García-Palacios A, Botella C, Szamosközi Ş, David D. Virtual reality exposure therapy in anxiety disorders: a quantitative meta-analysis. *Depress Anxiety*. 2012;29(2):85-93. doi:10.1002/da.20910

37. Powers MB, Emmelkamp PMG. Virtual reality exposure therapy for anxiety disorders: A meta-analysis. *J Anxiety Disord*. 2008;22(3):561-569. doi:10.1016/j.janxdis.2007.04.006

38. Scozzari S, Gamberini L. Virtual Reality as a Tool for Cognitive Behavioral Therapy: A Review. In: *Studies in Computational Intelligence*. Vol 337. ; 2011:63-108. doi:10.1007/978-3-642-17824-5_5

39. Riva G. The Key to Unlocking the Virtual Body: Virtual Reality in the Treatment of Obesity and Eating Disorders. *J Diabetes Sci Technol*. 2011;5(2):283-292. doi:10.1177/193229681100500213

40. Botella C, Serrano B, Baños R, García-Palacios A. Virtual reality exposure-based therapy for the treatment of post-traumatic stress disorder: a review of its efficacy, the adequacy of the treatment protocol, and its acceptability. *Neuropsychiatr Dis Treat*. 2015;11:2533-2545. doi:10.2147/NDT.S89542

41. Rizzo A, Cukor J, Gerardi M, et al. Virtual Reality Exposure for PTSD Due to Military Combat and Terrorist Attacks. *J Contemp Psychother*. 2015;45(4):255-264. doi:10.1007/s10879-015-9306-3

42. Rizzo A, Roy MJ, Hartholt A, et al. Virtual Reality Applications for the Assessment and Treatment of PTSD. In: Bowles S V., Bartone PT, eds. *Handbook of Military Psychology*. Cham: Springer International Publishing; 2017:453-471. doi:10.1007/978-

3-319-66192-6_27

43. Lakoff A. *Pharmaceutical Reason: Knowledge and Value in Global Psychiatry*. Cambridge: Cambridge University Press; 2006.

44. Cutter GR, Liu Y. Personalized medicine: The return of the house call? *Neurol Clin Pract*. 2012;2(4):343-351. doi:10.1212/CPJ.0b013e318278c328

45. Ashrafian H. Artificial Intelligence and Robot Responsibilities: Innovating Beyond Rights. *Sci Eng Ethics*. 2015;21(2):317-326. doi:10.1007/s11948-014-9541-0