

# **Limits of Artificial Intelligence Treatment for Depression and Anxiety**

A Research Paper submitted to the Department of Engineering and Society

Presented to the Faculty of the School of Engineering and Applied Science  
University of Virginia • Charlottesville, Virginia

In Partial Fulfillment of the Requirements for the Degree  
Bachelor of Science, School of Engineering

**Binh An Dang Nguyen**

Spring, 2022

On my honor as a University Student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments

Advisor

Bryn E. Seabrook, Department of Engineering and Society

## STS Research Paper

### Current Treatment for Depression and Anxiety Disorders

Amid the COVID-19 pandemic, “more than 42% of adults reported symptom of either an anxiety or depressive disorder” (Stephenson, 2021). The necessary social restrictions enforced during the pandemic led to isolation and unemployment, furthering the risk of mental health disorders. Demand for psychiatric care escalated during the quarantine period. However, America’s current system for mental health services is unable to keep up with the growing patient population. In fact, adults who “needed but hadn’t received counseling in the past 4 weeks also rose, from 9% (during August 19-31) to 12% (during December 9-21)” (Stephenson, 2021). The pandemic intensified pre-existing barriers with mental health care and increased “the number of individuals in need of treatment” (Carlo, 2021). Artificial Intelligence (AI) has the potential to narrow “unmet needs in psychiatry” (Allen). Current psychiatric treatments strongly rely on face-to-face interactions between the patient and clinicians” (Ducharme, 2019). Unfortunately, “clinicians actually get very little time to interact with patients” (Ducharme, 2019). Additionally, many patients find it difficult to schedule needed appointments. Not only does AI expand access to treatment options and allow more effective monitoring of patient health, but also aids in diagnosis and treatment planning.

Current diagnosis for mental disorders revolves around “physician-patient questionnaires that are most of the time inaccurate and ineffective in providing a reliable assessment of symptoms” (M; F), but AI could be utilized to make “diagnoses more quickly and accurately”(Ducharme, 2019). AI has immense potential in psychiatric treatment for depression and anxiety disorders if harnessed properly. However, it is important to ensure that undertaking this new technology will not lead to any consequences in the current quality of treatment for

patients. Concerns revolving the use of AI treatment focus on misdiagnosis of disorders and the inherent bias in machine learning models. To analyze the how AI can best be implemented in psychiatric care, Technological Fix theory is applied to understand the benefits and consequences of AI. The perspective of the Technological Fix gives professionals more insight into the “unforeseen and deleterious side effects” (Weinberg, 1978) of AI in America’s mental health care system. By highlighting these consequences, psychiatrists and other experts are aware of the limitations of technology and can fully utilize AI to its full potential, revolutionizing America’s current psychiatric system. In order to reach these conclusions and best integrate AI technology into our mental health care system, the following question needs to be addressed: What are the sociotechnical considerations for implementing AI in psychiatric treatment?

### **Discourse Analysis for Studying AI Influence on American Psychiatric Treatment**

The research question being discussed is: What are the sociotechnical considerations for implementing AI in psychiatric treatment? The question is answered through an analysis of scientific journals and research papers on the effectiveness of certain AI applications made for patients with depression. The published papers indicate benefits of AI treatment for depression and anxiety, offering an insight on how mental health professionals can approach this new technology. Prior research on AI in health care and other fields indicate AI’s shortcomings in mental health treatments. Since AI technology aims to reduce the gaps within America’s current mental health care system, studies focusing on the system’s weaknesses are explored to identify how AI bridges these disparities. This research paper is organized thematically based on benefits and potential complications of AI solutions in America’s mental healthcare system. Key terms in the study include: “machine learning,” “artificial intelligence,” “mental health care.”

### **History of Psychiatric Treatment in America**

The coronavirus uprooted American society and dragged the economy into a recession, piling additional stressors into people's daily lives. Many people suffering from mental health disabilities "face disproportionately high rates of poverty, housing and employment discrimination, and criminalization" (Altiraifi & Rapfogel 2021). The coronavirus widened these disparities that this oppressed community already face. As the pandemic wears on, people's mental health deteriorated as well, highlighting the urgent need for better mental health services in America.

Long before the pandemic, "the U.S. mental health care system was already failing to meet people's needs" (Altiraifi & Rapfogel 2021). The psychiatric system in America is riddled with biases and discrimination. Mental health providers have "denied communities facing various forms of oppression and control over their mental care" (Altiraifi & Rapfogel 2021). Treatment is not equal throughout the US and can be "cost-prohibitive, scarce, and coercive" (Altiraifi & Rapfogel 2021). Mental health services are unaffordable and inaccessible in America. A survey in 2016 indicated that "1.8 million American adults believed that they had a need for mental health services during the past year that went unmet" (Christidis 2018). The American mental health system provides inadequate care due to high cost, inconsistent treatment plans, and inaccessibility.

The pandemic emphasized the correlation between mental illnesses and systemic racism. Even though racial and ethnic minorities "overall have similar mental disorder than whites, consequences of mental illness in minorities [are longer lasting]" (*Mental health disparities: Diverse populations*). Many racial and ethnic groups who "have historically been discriminated against such as African Americans, American Indians, and Alaska Natives" (Altiraifi & Rapfogel 2021)" suffer a "disproportionately high burden of disability resulting from mental disorders"

*(Mental health disparities: Diverse populations)*. Despite lower rates of depression in the Hispanic and Black community compared to white Americans, depression is more likely to be persistent within people of color. This correlation can be due to a “lack of cultural understanding by health care providers” such as “language differences between patient and provider, stigma of mental illness among minority groups, and cultural presentation of symptoms” (*Mental health disparities: Diverse populations*). Minority groups and women also “use mental health services at substantially lower rates than white Americans” (Altiraifi & Rapfogel 2021). Within America’s mental health care system, there is a large disparity between white Americans and minority groups that needs to be addressed.

### **Current Diagnosis Methods of Depression and Anxiety**

Doctors determine a diagnosis of depression and anxiety based on the following: physical exams, lab tests, and psychiatric evaluation. Physical exams and lab tests are not discussed within the scope of the paper because AI development focuses on optimizing psychiatric evaluation. Both physical exams and lab tests are better understood than psychiatric evaluations because they have “specific biomarkers to make a diagnosis” and have less variability. Psychiatric evaluations collect information about symptoms, thoughts, feelings, and behavioral patterns through questionnaires (Mayo Foundation for Medical Education and Research 2018). Based on the evaluation and criteria listed in the Diagnostic and Statistical Manual of Mental disorders (DSM-5), psychiatrists determine the diagnosis. It is important to note that the DSM manual has “been updated seven times since it was first published in 1952 [to keep up] with scientific developments in psychiatry” (First 2016). One of the challenges with diagnosis is that “mental health disorders are ambiguously defined” (Newson 2022). Often, mental health assessment tools are inconsistent due to overlap between disorders. Additionally, “symptoms are

often assessed in a binary manner, asking simply about their presence or absence, rather than on a spectrum or scale.” There are many symptoms that “fall within the spectrum of normal mental functioning. Understanding this distinction prevents false positives in diagnosis” and “ensure that people receive appropriate treatment” (Newson 2022).

### **Current State of AI in Psychiatry**

Most psychiatric AI technologies exist in a research environment and not used professionally yet. These AI technologies focus on helping “researchers assess the heterogeneity of psychiatric conditions more fully” (Lindsey, 2021). For example, machine learning was used “to identify two forms of post-traumatic stress disorder” (Lindsey, 2021).

AI technology has made its way into digital mental health applications, “tallying over 10,000 mental health apps available to consumers” on the Apple and Google Play stores. Mental-health apps and programs such as Woebot incorporate AI and cognitive behavioral therapy to treat patients (Ducharme, 2019). Although there are a wide-range of options available for digital mental health treatments, “there is little research on the acceptability of available app-based treatments and whether they can actually reduce symptoms of psychopathology” (Mehta et al, 2021).

### **Technological Fix**

The STS research question is analyzed under the perspective of the Technological Fix. A technological fix is “a means for resolving a societal problem by adroit use of technology and with little or no alteration of social behavior. (Weinberg, 1978)” Although technology “has proved powerfully effective for solving any number of problems” (Newberry, 2005), there are “some types of problems that cannot or should not be fixed by technology” (Newberry, 2005).

This STS framework is an ideal method for discussing the research question because it considers “deleterious and unforeseen side effects” of AI use in America’s mental health system (Weinberg, 1978). Although mental health services in America is flawed, it is important that the introduction of this new technology does not create or worsen new problems while temporarily solving others.

The criticism of technological fix “asserts that attempting to solve problems by technological means is counterproductive; that, in the long run, it will make things worse” (Scott, 2011). People who support the technological fix stance are often “wary of technological fixes because they do not get to root causes,” believing them to be “a band-aid solution.” (Weinberg, 1978). A problem with technological fix theory is the heavy-handed blame towards technology. Many social critics fail to realize that social fixes also have unintended consequences as well (Weinberg, 1978). However, technological fix recognizes the “social bias in favor of technology” (Scott, 2011). It accepts technology as a solution to solve problems, but heeds “cautions about defects in this way of solving problems based on historical experiences” (Scott, 2011). The technological fix falls short due to the “clear distinction between technological problems and social ones” (Scott, 2011). In modern times, the line between technology and social problems is blurred. Society is intertwined with technology and cannot simply be addressed with just a social or technological fix. The technological fix falsely assumes that the “single root cause” to a problem that is purely social and behavioral in nature, and that technological fixes only address “symptoms” (Scott, 2011). The technological fix misleading believes that these problems can be resolved without any “side-effect problems” if the root of the problem is solved (Scott, 2011). However, each decision made will have unavoidable trade-offs.

## **Results and Discussion**

America's current mental health care system lacks the resources to support the influx of patients. In order to keep up with the fast-growing demand for mental health care, AI solutions can be implemented to facilitate more accurate and effective treatments and expand treatment access to American people.

Although AI has great potential in assisting clinicians and individuals, many professionals have concerns with this new technology. Since AI “[challenges] current norms and conceptual systems” in psychiatric care (Müller, 2020), there are many ethical concerns regarding this rising technology. An ethical discussion is needed to understand the “technology in its context” and “shape our societal response, including regulation and law” (Müller, 2020). To properly implement AI into psychiatric care and treatments, specialists need to understand the risks AI imposes. The main debate regarding use of AI in the complex field of mental health includes biases in AI systems, accessibility, effectiveness and quality of AI solutions, how AI can affect human behavior and relationships, and privacy and security. AI “will always cause some uses to be easier, and thus more frequent, and hinder other uses” (Müller, 2020). The focus of this analysis is to highlight risks AI technology poses in mental health care and approaches professionals take to tackle these problems.

### ***How Human Bias Affects AI***

One of the biggest challenges in the mental health care system is the “racial and ethnic bias” that “adversely affects mental health assessment and intervention” (Merino et al, 2018). This “biased provision of health services is a well-documented barrier to health for marginalized populations” (Merino et al, 2018). AI can reduce bias and discrimination in psychiatry, but



professionals need to consider how AI could affect the pre-existing biases for marginalized groups.

The bias found in AI is a result of human biases. Implicit bias causes unequal mental health care services in America's current psychiatric system because it can "negatively influence a provider's willingness to engage in patient-centered care, provide referrals to specialized treatment, or even adhere to evidence-based guidelines when serving diverse populations" (Merino et al, 2018). The bias manifests itself through physician's condescending language, lack of cultural awareness, and stigmatizing views towards certain populations. The subtle form of discrimination adversely affects patient's trust, diagnosis accuracy, and treatment quality for marginalized populations.

Implicit bias "refers to attitudes and beliefs that occur outside of our conscious awareness and control" (Ruhl, 2020). These biases can manifest through actions without the person being aware. To reduce this implicit bias, AI can be implemented. Since humans are unaware of how perspectives hurt minority groups, they might not understand the factors that influence their thinking. Although human behavior is hard to understand and document, machine decision is easier to examine and evaluate. Thus, AI minimizes "humans' subjective interpretation of data," resulting in fairer decision making. AI only considers factors that improve their predictive accuracy which means that they can remove variables such as sex, race, age, etc while humans struggle with eliminating these biases. (Silberg & Manyika, 2020)

Additionally, machine learning excels at "digesting large amounts of data very quickly and identifying patterns or finding anomalies or outliers in that data" (Walch, 2020). AI recognizes patterns of behavior and actions in data "that humans may not have even thought to look for" (Walch, 2020). AI can detect subtle discrimination and bias that people cannot see.

Although AI cannot completely solve the problem of biases and discrimination within the psychiatric field, it allows professionals to be more aware of the implicit biases that exist within their field of work and how to work against these unconscious behaviors.

Since mental health care services often use the one-on-one approach for treatment, a provider can act like a “gatekeeper to accessing mental care” (Merino et al, 2018). In fact, a study which uses “audio recordings of potential psychotherapy clients...show that middle-class white women are far more likely than working-class Black men to get a call back when requesting an appointment” (Merino et al, 2018). Psychiatrist set the standard for selecting and evaluating patients. If providers chose to disregard minority groups over others, their harmful values are reflected in the data that is used to create AI. Thus, the biases present in the psychiatric field translate into the algorithms used for creating machine learning models. Although AI can reduce the implicit bias in psychiatric care, AI can also create bias on a larger scale.

AI is developed by “combining large amounts of data with fast, iterative processing and intelligent algorithms” (“Artificial Intelligence (AI) – what it is and why it matters”, 2018). The AI learns by detecting patterns within the data that engineers feed it. To understand the requirements needed for psychiatric care for depression and anxiety, AI needs data that is representative of the true patient population for depression and anxiety. However, this data is interlaced with “preferences or exclusions” of certain groups. This bias “can also be introduced by how data is obtained, how algorithms are designed, and how AI outputs are interpreted” (Siwicki, 2021). If the training data cannot be generalized and does not represent the population, the created AI will reflect those same biases. Disproportionate data that favors certain groups over others further perpetuate the pre-existing bias in the American mental health care, pushing

away groups who are already shunned by the system. AI technology can unknowingly exacerbate medical discrimination.

The biases within medical AI are “a result of the social and material systems in which they are built and embedded” (Shaw & Donia, 2021). AI is unable to solve these problems of implicit bias within our society and can even usher bias on a massive scale. There is a deep history of medical discrimination of people of color in America that cannot be solved through technology. Even though overtly discriminatory practices are no longer present, this prejudiced way of thought still resonates in psychiatry today. Forms of discrimination manifest themselves in psychiatric treatment and diagnoses. In some cases, “psychiatrists added the word ‘aggressive’ to the definition of schizophrenia, and marketed pharmaceuticals directly at Black patients who they felt were more ‘out of control’” (Smith, 2020). The long-term effect of this mindset causes a “disparity in diagnostic rates” (Smith, 2020). For example, “black men are more likely to diagnosed with schizophrenia than white, with no scientific basis” (Smith, 2020). Implicit bias and lack of cultural awareness led to unintentional and harmful judgments about people of color.

AI alone is not enough to reduce the bias within psychiatry. Society needs to recognize how these implicit biases impact America’s mental health care system. In order to reduce the biases and discrimination, practitioners need to understand what factors shape the way “people approach mental health” (Smith, 2020) and what actions they unconsciously take that causes these biases.

### ***How AI Resolves Accessibility Issues***

Each year, “Americans with mental health illnesses struggle to find mental health care” (“The doctor is out,” 2017). Statistics show that patients “who looked for a new mental health

provider in the last year contacted psychiatrists who were not accepting new patients (55%) or who did not accept their insurance (56%)” (“The doctor is out,” 2017). There are not enough providers to meet with the needs of American people. Financial and availability barriers prevent Americans from receiving mental health treatment and support. AI can overcome these barriers “by obviating the need for a trained clinician” (Mehta et al, 2021). Digital mental health interventions reduce the cost and increase accessibility to those with internet. Since AI can provide fully automated treatments, it will also be a favorable way for patients “who are uncomfortable seeking help from another person” (Mehta et al, 2021). These AI treatment options such as Youper and Woebot are “available through mobile apps that “can be installed on a person’s mobile device” (Mehta et al, 2021).

This medium of treatment allows users to access their interventions “anytime and anywhere” (Mehta et al, 2021) which allows patients to access healthcare treatment whenever is convenient for them. It reduces the barriers of time and cost that are obstacles for patients today. Although AI can give more options for patients who need treatment options to manage their emotional disorders, this technology will not fix any structural problems within the psychiatric system.

While AI in health care aims to increase equality of access to health care services, it can potentially hurt “healthcare services among economically disadvantaged and vulnerable populations” (Jumelle & Ispas, 2014). Digital health care relies on “different kinds of infrastructure” (Shaw & Donia, 2021). These infrastructures include “the cables and wires that enable digital signals to travel over distance that make digital communication available” (Shaw & Donia, 2021). Since AI technologies rely heavily on high-speed internet and phones or similar kinds of technology, AI healthcare can further disadvantage groups of people who already do not

“have the financial means, technical knowledge or desire to use the Internet, communication devices and digital health technologies and applications” (Jumelle & Ispas, 2014).

Even though AI aims to expand care to all patients, it can also intimidate and exclude certain populations from needed care. To combat this unintended effect, AI technologies and “applications should be included as part of the universal healthcare services to improve the health of the population” rather than serve as a replacement for treatment (Jumelle & Ispas, 2014). Adding more options to the healthcare pool will give more options for patients who need health care. A wider range of treatment planes will improve overall mental healthcare and reduce the barriers for these mental health care services.

Due to the increased accessibility of psychiatric care through AI on mainstream apps, AI aims to reduce attitude barriers to mental health care. Stigma surrounding mental health “comes from lack of understanding or fear. Inaccurate or misleading media representations of mental illness contribute to both those factors. (“Stigma, Prejudice and Discrimination Against People with Mental Illness”). A negative view on mental illness, manifests into public stigma, self-stigma and institutional stigma. The stigma around mental illnesses “limit opportunities for people with mental illness” through lower funding and services for mental healthcare (“Stigma, Prejudice and Discrimination Against People with Mental Illness”). Many patients who need mental health treatment do not seek it because of their “internalized shame, that about their own condition” (“Stigma, Prejudice and Discrimination Against People with Mental Illness”). Although AI will not directly solve the negative view of mental illnesses, it can still help destigmatize mental health issues by “making mental health diagnosis and treatment more quantifiable and less subjective” (Asar, 2020). Mainstreaming healthcare encourages patients to seek treatment because they recognize that many people suffer from the same conditions, and it

normal and not shameful. Thus, the public, individuals, and big institutions will recognize the need for mental healthcare and revolutionize the way people view mental illnesses, bettering mental health services and care in America.

### ***Effectiveness of Patient Empowerment***

Applications that use AI therapy for anxiety and depression treatment such as Youper and Woebot use self-guided intervention techniques to “target emotional disorders” (Mehta et al, 2021). These digital health technologies like Youper are “infused with self-tracking mechanisms” that “[encourage] people to self-police their own actions and habits, meaning that they have heightened awareness about whether and how their action align with expected social norms” (Shaw & Donia, 2021). Although patient empowerment gives more independent knowledge to patients, it can control human behavior and restrict them. The resulting behavior can generate new problems such as heightened patient self-awareness in behaviors leading to excessive inspection of actions.

Although “empowerment may be a major benefit of digital health, an overemphasis may be harmful to patients” (Jumelle & Ispas, 2014). Self-policing can lead to increased dissatisfaction, insecurity, and depression if the patient feels that he/she is underperforming. Shifting the responsibilities to patients “can become an added responsibilities and burden in daily life,” causing patients to “resist their new role” (Jumelle & Ispas, 2014). Furthermore, some patients may “not have the resources or may lack support in developing empowerment” (Jumelle & Ispas, 2014). Patient empowerment is best used for patients with a supportive community or group. However, not all patients have access to these resources which means that self-intervention techniques could worsen symptoms for patients or prevent patients from getting effective treatment.

A large shift of responsibility can overwhelm patients and increase the “risk of non-adherence and likelihood of patients modifying their treatment plans” (Jumelle & Ispas, 2014). However, there are also benefits of relying on patient empowerment. Through patient empowerment, patients are encouraged to “take more active responsibility in the self-management of their health with support from digital health technologies” (Jumelle & Ispas, 2014). Thus, patients can make more informed choices about what they need. However, an abundance of knowledge can overwhelm patients and worsen their symptoms. Overall, patients should be able to decide what they want to do with the information given to them and participate using the information they have. By giving patients knowledge and access to these self-intervention tools and information, they can better cultivate their own sense of well-being if they choose too. Withholding this information and power from patients that “could be detrimental to a few, is overly restrictive” (diagnostic imaging). Although these risks should not be ignored, digital health technologies should be utilized to increase patient active responsibility and self-management of their mental health.

Furthermore, blame cannot be fully placed on AI technology for increased anxiety and non-compliance as these are problems found in traditional psychiatric treatments as well. In fact, “one of the main factors contributing to poor response to pharmacological treatment is adherence” (Carvajal, 2004). Reasons for “poor adherence can be due to drug-related factors (tolerance, complexity of prescription, side effects, or cost), patient-related variables (illness symptoms, comorbidity, insight capacity, belief system, or sociocultural environment), and physician-related factors (communication or psychoeducational style)” (Carvajal, 2004). Treatments are difficult for patients to maintain due to variable of reasons, AI’s reliance on self-policing and an increased shift to patient responsibility can’t be to blame for lack of patient

involvement. Overall, AI technology can serve as a powerful assistive technology that increase patient independence and knowledge. However, the “appropriateness of digital health technologies depends on whether the technology replaces more restrictive measures, becomes a restriction on autonomy” (Jumelle & Ispas, 2014).

### ***Misdiagnosis***

Specialists question if AI can effectively diagnose mental illnesses when professionals still struggle with understanding mental illnesses. The ambiguity and lack of understanding of mental illnesses is reflected in the numerous changes to the DSM manual. Depression and anxiety are “still not well understood in research and clinical settings” (Canales, 2019).

If the quality of AI technology used in clinical settings is subpar, it can further perpetuate problems within patients who are seeking treatment. An inaccurate diagnosis “keeps people from receiving treatment for their affliction, allowing distress to grow unchecked as hope diminishes,” leading “to increasingly dangerous psychological conditions that threaten both emotional and behavioral function” (Kvarnstorm, 2020). Misdiagnosis leads to treatments that can potentially “increase dysfunction by exposing [patients] to pharmacological therapies that augment existing symptomatology or cause the emergence of new symptoms” (Kvarnstorm, 2020). For instance, “people whose bipolar disorder is treated with antidepressants are at heightened risk of mania, increased frequency of mood switching, and more severe depressive episodes than if they received no treatment at all.” (Kvarnstorm, 2020). Misdiagnosis and mistreatment cause immediate and long-term damage to the patient and can further “complicate treatment and compromise treatment outcomes” (Kvarnstorm, 2020). Inaccurate information given by AI technology can potentially impede on patient treatment for depression and anxiety instead of solving them.



Even without the addition of AI technology, misdiagnosis of mental illnesses such as depression and anxiety “is an alarmingly prevalent phenomenon that keeps [patients] from being able to achieve psychological wellness. (Kvarnstorm, 2020).” Misdiagnosis is already concerning problem within the mental health system. Factors that contribute to the high rate of misdiagnosis include “clinicians [who] are inclined to diagnose disorders that they feel more comfortable treating,” failing “to recognize emotions as behaviors as symptomatic of mental illness and thus fail to report them,” and cultural and ethical ignorance. (Kvarnstorm, 2020). Although inaccurate AI can potentially cause misdiagnosis, professionals also fail to accurately diagnoses patients due to lack of understanding and inability to recognize behaviors and symptoms.

### ***How AI Affects Doctor-Patient Relationships***

Using AI technology for treatment will “increase [patient] reliance on digital health technologies,” altering traditional doctor-patient relationship (Jumelle & Ispas, 2014). AI technology changes how patients and doctors interact with each other “in terms of distance and time of communication and how medical care is delivered” (Jumelle & Ispas, 2014).

Patient-doctor relationship is built on “a patient’s and clinician’s mutual trust, respect, and commitment” (Nagy & Sisk, 2020). Their relationship needs to be maintained and strengthened through mutual understanding and effective communication. Building and improving these relationships “is also essential to improving the overall health care system” (Nagy & Sisk, 2020). When this relationship is weakened, patients “experience anxiety, frustration, and second-guessing.” Integrating AI into therapy for depression and anxiety should strive to foster the relationship between clinicians and patients rather than hurt it.

Specialists fear that patients will value the “increased diagnostic and treatment accuracy offered by AI” (Nagy & Sisk, 2020) over the care provided by clinicians. The information and treatment plans given by AI and clinicians may differ. In cases where patients value AI knowledge more, it casts doubt towards the clinician’s viewpoint, straining the strength of the doctor-patient relationship. Professionals also fear that the “humanistic aspects of medical care” (Nagy & Sisk, 2020) can also be met by AI through conversational agent systems. Introducing AI systems in health care could cause a bias towards technology where patients favor AI solutions over human viewpoints. The bias towards technology diminishes doctor-patient relationship and reduces trust towards professionals in the field, inhibiting the growth of mental health treatments.

AI strives to off-load tedious work for psychiatrist and assist them in diagnosing and treating patients. Clinicians spend a considerable amount of time “analyzing patient data, developing a differential diagnosis, and evaluating potential treatment options,” (Nagy & Sisk, 2020) making it difficult for doctors to allot more time for patients. AI helps physicians reduce the time needed to do “tedious charting responsibilities both before, during, and after the patient encounter” (Nagy & Sisk, 2020), allowing doctors to shift their attention towards patients. More time spent interacting with patients improves communication and understanding between patient and doctor about the nature of the patient’s illness and symptoms. However, the integration of AI technology could have the opposite effect, potentially reducing the time doctors spend with their patients.

With the addition of AI technologies, there will be “structural and personal barriers might hinder clinicians from using this time to further develop their relationship with their patients” (Nagy & Sisk, 2020). AI assistance is able to support more patient care by decreasing the time it

takes to treat each patient. However, due to the nature of America's health care system and business model of medicine, AI can result in impersonal relationships between doctor and patients. If this new technology "decreases the time required for a patient visit, the health care system might respond by increasing the volume of patients seen per day rather than allowing time for relationship development and shared decision making" (Nagy & Sisk, 2020) Physicians can end up with a more tightly packed schedules with less time for each individual patient.

Even if the patient volume remains the same, "personal barriers might prevent some clinicians from engaging with their patients to develop trust and elicit their values regarding goals of care" (Nagy & Sisk, 2020). Some clinicians might feel uncomfortable with "highly personal and emotional communication" (Nagy & Sisk, 2020). With more time allotted for each session, patients may begin to share more sensitive information to their clinicians. Patients may want their clinician to "provide sensitive, acknowledging, and supportive statements" (Nagy & Sisk, 2020). However, clinicians struggle with engaging and understanding their patients. In fact, "many clinicians report low confidence in their ability to engage in difficult or emotionally charged conversations as a reason for not engaging in shared decision making" (journal). Their inability to respond confidently or interact with their patients hinders trust between their patients. A longer patient session will only highlight the clinicians lacking communication skills. However, this is not the fault of the AI but due to lack of training or experience for clinicians. Clinicians' primary job is to be able to communicate with their patients, AI only highlights the need for "improving clinicians' communication and social skills" (Nagy & Sisk, 2020).

Although AI provides psychiatrists with effective tools for diagnosing patients, "persistent personal barriers can impede the development of healing relationships" (Nagy & Sisk, 2020). Implementing AI can bring structural changes to psychiatric care that affect the

patient-doctor relationship. AI can assist in building deeper connections with patients but can also hurt these developing relationships if there aren't rules that "aim to address personal and professional barriers that can hinder the development of a trusting and open patient-clinician relationship" (Nagy & Sisk, 2020). Overall, the issues within mental health cannot fully be addressed by technology or social solutions, but they can be used together to make up for their downfalls.

### ***Privacy Concerns with AI***

AI can be used to create diagnosis and treatment plans. However, in order to create an accurate, inclusive model, the machine learning algorithms need personal data from individuals. Although AI creates more efficient care, including AI in mental health care raises concerns about privacy due to the "the high volume of personal healthcare data flowing between different devices and different healthcare information systems" (Jumelle & Ispas, 2014). AI trains and delivers effective treatment plans and diagnostic evaluations using patient data. This data "may include diagnostic images, video, audio and email" (Jumelle & Ispas, 2014). Although the data will be anonymized to give more protection to patients, there is still a "security and the potential of re-identifying patients through unauthorized methods" (Jumelle & Ispas, 2014) Due to this security and privacy risk, many patients may feel uncomfortable with utilizing AI and sharing their healthcare information. AI technology is a greater risk to patient data and health care information.

AI could endanger patient privacy and security of their healthcare information. The information that gets leaked "impact of organizations such as insurance companies" (Jumelle & Ispas, 2014). These companies utilize patient information and "collect information about individual behaviors and shape their product offerings accordingly." AI poses security and

privacy concerns regarding sensitive information about patients. However, it's important to note that AI is not the only technology that poses these concerns. Many organizations have been transitioning “paper healthcare records to electronic records” (Jumelle & Ispas, 2014). These organizations “gather data meticulously and to keep personal health records with appropriate levels of privacy protections and to seek patients’ consent regarding what data will be collected, who has access to which information in the EHR, and releasing information to a third party” (Jumelle & Ispas, 2014). Implementing AI requires heavy considerations for privacy protections and stricter policies.

### **Next Steps for Studying AI’s Role and Influence in American Psychiatric Care**

Since AI is relatively new to America’s mental health system, the effect of AI is not well recognized yet. Most of the AI made for depression patients are newly created which means that research on the effectiveness of AI technology is limited. Most of the studies used for the research is observational and uses self-reported measures and cannot show “a complete picture of the impact of [psychiatric AI] on clinical symptoms and overall functioning” (Mehta et al, 20221). Future work should focus on researching the use of AI used by clinicians and the effectiveness of long-term treatments for patients with depression and anxiety. When AI begins to form a stronger foothold in psychiatric care, more research can be done on the specific impacts AI has on psychiatry compared to other fields. By seeing how AI performs in mental health treatments for depression and anxiety first-hand, professionals gain better insight on the true dangers and benefits of the technology.

### **Importance of Understanding AI’s Influence in Psychiatric Care**

Depression and anxiety are often addressed through medication and psychotherapy. The mental health care system can be “immensely improved” through technology and “[drive] more informed decisions, accessible treatment and operational efficiencies” (McAllister, 2021). However, technology alone will not overcome the “historic, persistent and endemic structural inequities,” and deep-rooted stigma (McAllister, 2021). Utilizing AI technology without careful consideration can even worsen existing issues including widespread bias and further disabling particular ethnic groups. A sociotechnical approach to AI technology in psychiatric care broadens the issues, infrastructures, and communities that can be implicated due to this fix. With a focus on people, the concern is directed towards ethics and bettering communities long-term rather than important stakeholders. The paper outlines the importance of ethics and socio-technical considerations and strategies needed for merging technology and psychiatric care.

## References

- Allen, S. (n.d.). *Artificial Intelligence and the future of Psychiatry*. EMBS. Retrieved December 8, 2021, from <https://www.embs.org/pulse/articles/artificial-intelligence-and-the-future-of-psychiatry/>.
- Altiraifi, A., & Rapfogel, N. (2021, November 7). *Mental health care was severely inequitable, then came the coronavirus crisis*. Center for American Progress. Retrieved February 3, 2022, from <https://www.americanprogress.org/article/mental-health-care-severely-inequitable-came-coronavirus-crisis/>
- Artificial Intelligence (AI) – what it is and why it matters*. SAS. (n.d.). Retrieved March 20, 2022, from [https://www.sas.com/en\\_in/insights/analytics/what-is-artificial-intelligence.html#howitworks](https://www.sas.com/en_in/insights/analytics/what-is-artificial-intelligence.html#howitworks)
- Asar, A. (2020, August 18). *Council post: Five ways AI can help revolutionize mental healthcare*. Forbes. Retrieved March 20, 2022, from <https://www.forbes.com/sites/forbestechcouncil/2020/08/19/five-ways-ai-can-help-revolutionize-mental-healthcare/?sh=7f815a3013ab>
- Bateman, K. (n.d.). *4 ways artificial intelligence is improving mental health therapy*. World Economic Forum. Retrieved April 2, 2022, from <https://www.weforum.org/agenda/2021/12/ai-mental-health-cbt-therapy/>

- Carlo, Andrew D. M. D. (2021, April 1). *Harnessing collaborative care to meet mental health demands in the era of covid-19*. JAMA Psychiatry. Retrieved December 8, 2021, from <https://jamanetwork.com/journals/jamapsychiatry/fullarticle/2771405>.
- Carvajal, C. (2004). Poor response to treatment: Beyond medication. *Predictors of Response to Treatment in Neuropsychiatry*, 6(1), 93–103.  
<https://doi.org/10.31887/dcms.2004.6.1/ccarvajal>
- Chen, J., Mullins, C. D., Novak, P., & Thomas, S. B. (2015). Personalized strategies to activate and empower patients in health care and reduce health disparities. *Health Education & Behavior*, 43(1), 25–34. <https://doi.org/10.1177/1090198115579415>
- Choi, S. W., Ramos, C., Kim, K., & Azim, S. F. (2019, April 15). *The Association of Racial and ethnic social networks with mental health service utilization across minority groups in the USA - Journal of racial and ethnic health disparities*. SpringerLink. Retrieved February 3, 2022, from <https://link.springer.com/article/10.1007/s40615-019-00583-y>
- Christidis, P. (2018, April). *An unmet need for mental health services*. Monitor on Psychology. Retrieved February 3, 2022, from <https://www.apa.org/monitor/2018/04/datapoint>
- Ducharme, J. (2019, November 20). *How artificial intelligence could save psychiatry*. Time. Retrieved December 8, 2021, from <https://time.com/5727535/artificial-intelligence-psychiatry/>.



First, M. B. (2016, December 16). *Psychiatry's list of disorders needs real-time updates*. STAT. Retrieved February 3, 2022, from <https://www.statnews.com/2016/12/12/psychiatry-dsm-update/>

Garg, P., & Glick, S. (2018, October 24). *Ai's potential to diagnose and treat mental illness*. Harvard Business Review. Retrieved March 20, 2022, from <https://hbr.org/2018/10/ai-potential-to-diagnose-and-treat-mental-illness>

Jumelle, A. K., & Ispas, I. (2014). Ethical issues in Digital Health. *Requirements Engineering for Digital Health*, 75–93. [https://doi.org/10.1007/978-3-319-09798-5\\_4](https://doi.org/10.1007/978-3-319-09798-5_4)

Kvarnstrom, E. (2020, October 25). *The dangers of mental health misdiagnosis: Why accuracy matters*. Bridges to Recovery. Retrieved March 20, 2022, from <https://www.bridgestorecovery.com/blog/the-dangers-of-mental-health-misdiagnosis-why-accuracy-matters/>

Lakdawala, P. D. (2015). Doctor-patient relationship in psychiatry. *Mens Sana Monographs*, 13(1), 82. <https://doi.org/10.4103/0973-1229.153308>

Lindsey, H. (2021, September 30). *Ai's push to understand psychiatry research has the potential to tackle mental illness*. Business Insider. Retrieved April 2, 2022, from <https://www.businessinsider.com/how-ai-is-advancing-psychiatry-medicine-to-tackle-mental-illness-2021-9>

M;, F. (n.d.). *Artificial Intelligence in psychiatry*. Advances in experimental medicine and biology. Retrieved December 8, 2021, from <https://pubmed.ncbi.nlm.nih.gov/31705492/>.

Mayo Foundation for Medical Education and Research. (2018, February 3). *Depression (major depressive disorder)*. Mayo Clinic. Retrieved February 3, 2022, from <https://www.mayoclinic.org/diseases-conditions/depression/diagnosis-treatment/drc-20356013>

McAllister, K. (2021, September 2). *Stigmatized issues and cross-system collaboration have still evaded tech solutions to date, members of the Braintrust say*. Protocol. Retrieved February 6, 2022, from <https://www.protocol.com/braintrust/healthcare-problems-tech-cant-fix?rebelltitem=1#rebelltitem1>

Mehta, A., Niles, A. N., Vargas, J. H., Marafon, T., Couto, D. D., & Gross, J. J. (2021). Acceptability and effectiveness of artificial intelligence therapy for anxiety and depression (Youper): Longitudinal observational study. *Journal of Medical Internet Research*, 23(6). <https://doi.org/10.2196/26771>

*Mental health disparities: Diverse populations*. American Psychiatric Association. (n.d.). Retrieved February 3, 2022, from <https://www.psychiatry.org/psychiatrists/cultural-competency/education/mental-health-facts>

Merino, Y., Adams, L., & Hall, W. J. (2018). Implicit bias and mental health professionals: Priorities and directions for research. *Psychiatric Services*, 69(6), 723–725. <https://doi.org/10.1176/appi.ps.201700294>

Müller, V. C. (2020, April 30). *Ethics of Artificial Intelligence and Robotics*. Stanford Encyclopedia of Philosophy. Retrieved March 19, 2022, from <https://plato.stanford.edu/entries/ethics-ai/>

Nagy, M., & Sisk, B. (2020). How will artificial intelligence affect patient-clinician relationships? *AMA Journal of Ethics*, 22(5). <https://doi.org/10.1001/amajethics.2020.395>

NAMI. (2017, November). *The doctor is out*. NAMI. Retrieved March 20, 2022, from <https://www.nami.org/Support-Education/Publications-Reports/Public-Policy-Reports/The-Doctor-is-Out>

Newberry, B. (2005). Technological Fix. In C. Mitcham (Ed.) *Encyclopedia of Science, Technology and Ethics*. (Volume 4., pp. 1901-1903). New York, New York. Macmillan Reference USA.

Newson, J. J. (2022, January 29). *3 challenges in Mental Health Assessment - Sapien Labs: Neuroscience: Human Brain Diversity Project*. Sapien Labs | Neuroscience | Human Brain Diversity Project. Retrieved February 3, 2022, from <https://sapienlabs.org/mentalog/3-challenges-in-mental-health-assessment/>

Ruhl, C. (2020, July 1). *Implicit or unconscious bias*. Implicit Bias is a type of Unconscious Bias | Simply Psychology. Retrieved March 20, 2022, from <https://www.simplypsychology.org/implicit-bias.html>

Scott, D. The Technological Fix Criticisms and the Agricultural Biotechnology Debate. *J Agric Environ Ethics* 24, 207–226 (2011). <https://doi.org/10.1007/s10806-010-9253-7>

Shaw, J. A., & Donia, J. (2021). The sociotechnical ethics of Digital Health: A Critique and extension of approaches from Bioethics. *Frontiers in Digital Health*, 3. <https://doi.org/10.3389/fdgth.2021.725088>

- Silberg, J., & Manyika, J. (2020, July 22). *Tackling bias in artificial intelligence (and in humans)*. McKinsey & Company. <https://www.mckinsey.com/featured-insights/artificial-intelligence/tackling-bias-in-artificial-intelligence-and-in-humans>
- Siwicki, B. (2021, November 30). *How ai bias happens – and how to eliminate it*. Healthcare IT News. Retrieved March 20, 2022, from <https://www.healthcareitnews.com/news/how-ai-bias-happens-and-how-eliminate-it>
- Smith, K. M. (2020, July 6). *Discrimination and racism in the history of Mental Health Care*. NAMI. Retrieved March 20, 2022, from <https://www.nami.org/Blogs/NAMI-Blog/July-2020/Discrimination-and-Racism-in-the-History-of-Mental-Health-Care>
- Stephenson, Joan P. D. (2021, April 6). *CDC study finds worsening anxiety and depression, especially in young adults, during COVID-19 pandemic*. JAMA Health Forum. Retrieved December 8, 2021, from <https://jamanetwork.com/journals/jama-health-forum/fullarticle/2778458>.
- Stigma, Prejudice and Discrimination Against People with Mental Illness*. American Psychiatric Association . (n.d.). Retrieved March 20, 2022, from <https://www.psychiatry.org/patients-families/stigma-and-discrimination>
- Walch, K. (2020, May 9). *How AI is finding patterns and anomalies in your data*. Forbes. Retrieved March 20, 2022, from <https://www.forbes.com/sites/cognitiveworld/2020/05/10/finding-patterns-and-anomalies-in-your-data/>

Weinberg, A. (1978). *Beyond the Technological Fix*. Internal Institute for Energy Analysis report. Oak Ridge Associated Universities.