



Key Considerations for Incorporating Conversational AI in Psychotherapy

Adam S. Miner^{1,2,3*}, Nigam Shah⁴, Kim D. Bullock¹, Bruce A. Arnow¹, Jeremy Bailenson³ and Jeff Hancock³

¹ Department of Psychiatry and Behavioral Sciences, Stanford University School of Medicine, Stanford, CA, United States, ² Department of Epidemiology and Population Health, Stanford University School of Medicine, Stanford, CA, United States, ³ Department of Communication, Stanford University, Stanford, CA, United States, ⁴ Stanford Center for Biomedical Informatics Research, Stanford University School of Medicine, Stanford, CA, United States

OPEN ACCESS

Edited by:

Michelle Burke Parish,
University of California, Davis,
United States

Reviewed by:

Stefanie Kristiane Gairing,
University Psychiatric Clinic Basel,
Switzerland
Donald M. Hilty,
UC Davis Health,
United States
Peter Yellowlees,
University of California, Davis,
United States

*Correspondence:

Adam S. Miner
aminers@stanford.edu

Specialty section:

This article was submitted to
Public Mental Health,
a section of the journal
Frontiers in Psychiatry

Received: 09 December 2018

Accepted: 17 September 2019

Published: 18 October 2019

Citation:

Miner AS, Shah N, Bullock KD,
Arnow BA, Bailenson J and
Hancock J (2019) Key Considerations
for Incorporating Conversational
AI in Psychotherapy.
Front. Psychiatry 10:746.
doi: 10.3389/fpsy.2019.00746

Conversational artificial intelligence (AI) is changing the way mental health care is delivered. By gathering diagnostic information, facilitating treatment, and reviewing clinician behavior, conversational AI is poised to impact traditional approaches to delivering psychotherapy. While this transition is not disconnected from existing professional services, specific formulations of clinician-AI collaboration and migration paths between forms remain vague. In this viewpoint, we introduce four approaches to AI-human integration in mental health service delivery. To inform future research and policy, these four approaches are addressed through four dimensions of impact: access to care, quality, clinician-patient relationship, and patient self-disclosure and sharing. Although many research questions are yet to be investigated, we view safety, trust, and oversight as crucial first steps. If conversational AI isn't safe it should not be used, and if it isn't trusted, it won't be. In order to assess safety, trust, interfaces, procedures, and system level workflows, oversight and collaboration is needed between AI systems, patients, clinicians, and administrators.

Keywords: natural language processing, artificial intelligence, expert systems, psychotherapy, conversational AI, chatbot, digital assistant, human-computer interaction

INTRODUCTION

Clinicians engage in conversations with patients to establish a patient-therapist relationship (i.e., alliance), make diagnoses, and provide treatment. In traditional psychotherapy, this conversation typically involves a single patient and a single clinician (1). This model of psychotherapy is being modified because software programs that talk like people (i.e., conversational artificial intelligence, chatbots, digital assistants) are now beginning to provide mental health care (2). Conversational artificial intelligence (AI) is gathering diagnostic information (3, 4) and delivering evidence-based psychological interventions (5–7). Additionally, conversational AI is providing clinicians with feedback on their psychotherapy (8) and talking to young people about suicide, sex, and drug use (9, 10).

Conversational AI appears unlikely to achieve enough technical sophistication to replace human therapists anytime soon. However, it does not need to pass the Turing Test (i.e., able to hold human seeming conversations) to have a significant impact on mental health care (2). A more proximal challenge is to plan and execute collaborative tasks between relatively simple AI systems and human practitioners (11–13). Although AI in mental health has been discussed broadly (for a review

see 14), specific formulations of clinician-AI collaboration and migration paths between forms remain vague.

Articulating different forms of collaboration is important, because the deployment of conversational AI into mental health diagnosis and treatment will be embedded within existing professional services. Conversational AI will likely interact with traditional workers (i.e., clinicians), but how these roles and responsibilities will be allocated between them has not been defined. To guide future research, we outline four approaches and dimensions of care that AI will affect.

Within the four approaches of AI-human integration in mental health service delivery, one extreme is a view that any involvement by conversational AI is unreasonable, putting both patients and providers at risk of harmful unintended consequences. At the other extreme, we explore how conversational AI might uniquely serve a patient's needs and surpass the capacity of even the most experienced and caring clinician by overcoming entrenched barriers to access. Although embodiment (e.g., virtual avatars or robots) can have a significant impact on interactions with virtual systems, we focus exclusively on the potential benefits and challenges of verbal and written language-based conversation and ignore the implications of embodiment or presence (15). **Table 1** summarizes the four approaches and our related assumptions.

CARE DELIVERY APPROACHES

It is unclear whether the path forward will involve simultaneous experimentation with all four degrees of digitization, or progression through these approaches. We first briefly describe how these compare to the way individual psychotherapy is most often delivered today. Perhaps surprisingly, laws, norms and the ethics of data sharing represent a nonobvious but critical factor in how these alternative approaches can operate now or develop in the future.

Currently, psychotherapy sessions are rarely recorded except in training institutions for supervision. When they are, for example during training or to assess clinician fidelity during clinical trials, trained human clinicians with prescribed roles and responsibilities are the listeners and provide oversight. With few exceptions, such as immediate risk of serious harm to the patient or others, clinicians need explicit permission to share identifiable patient information. When one of these exceptions is invoked, there is an obligation

to limit the sharing strictly to the extent needed to provide effective treatment and ensure safety (16, 17). Against this backdrop, having conversational AI listen to psychotherapy sessions or talk directly with patients represents a departure from established practice.

In the “humans only” approach, psychotherapy remains unchanged. Most psychotherapy sessions are heard only by the patient and clinician who are in the room. If a session were recorded, the labor intensiveness of human review would ensure most sessions would never be analyzed (8). The second approach, “human delivered, AI informed,” introduces into the room a listening device connected to software that detects clinically relevant information (18) such as symptoms or interventions (19), and relays this information back to the patient or clinician. Quantitative analysis of recorded psychotherapy is in its early stages, but it shifts to software programs the burden of extracting relevant information from audio or text. In the third approach, “AI delivered, human supervised,” patients speak directly to a conversational AI with the goal of establishing diagnoses or providing treatment (20). A human clinician would either screen patients and hand off specific tasks to conversational AI or supervise conversations between front-line conversational AI and patients. The fourth approach, “AI only,” would have patients talk to a conversational AI with no expectation of supervision by a human clinician.

One of the less developed but more alluring ideas of AI psychotherapy is “AI delivered, human supervised.” Even the most ardent supporters of AI will acknowledge that there are certain things humans do better than computers. Combining people and algorithms may potentially build on the best of both approaches, and AI-human collaboration has been suggested as a way to address limitations in planning treatment in other medical areas such as oncology (21). Indeed, the prevailing opinion of expert systems researchers in the 1980s argued that computer-human collaboration would outperform either people or computers alone (for a review see 22).

In assessing any system to augment the practice of psychotherapy the first consideration of its impact should be that it will ensure patients and clinicians are helped and not harmed (23, 24). In the discussion below, we consider salient issues that impact the potential value and harm of different delivery mechanisms by focusing on four dimensions of impact: access to care, quality, clinician-patient relationship, and patient self-disclosure.

TABLE 1 | Delivery approaches and dimensions of impact for conversational AI.

Care delivery approach	Dimensions of impact			
	Access to care	Quality	Clinician-patient relationship	Patient self-disclosure and sharing
Humans only	Unchanged	Established	No disruption	Unchanged
Human delivered, AI informed	Unchanged	Potentially improved	Potentially disrupted ^a	Unknown
AI delivered, human supervised	Improved, but limited scalability	Unknown	Likely disrupted	Unknown
AI only	Improved, not restrained by human attention	Unknown	Nonexistent	Unknown

^aBy “disrupt” we do not mean to signal that the result will be necessarily good or bad.

DIMENSIONS OF IMPACT

Access to Care

Limited access to mental health treatment creates a demand for scalable and non-consumable interventions (25, 26). Despite the high costs and disease burden associated with mental illness (27), we have a decreasing number of clinicians per capita available to provide treatment in the US (28). Increasing the number of human clinicians is not currently feasible, in part because of the decline from 2008 to 2013 per capita for both psychologists (from a ratio of 1:3,642 to 1:3,802) and psychiatrists (from a ratio of 1:7,825 to 1:8,476) (28). Conversational AI has the potential to help address insufficient clinician availability because it is not inherently limited by human clinician time or attention. Conversational AI could also bridge one of the current tensions in care delivery: although clinicians value patient conversations, they have no financial incentive to engage in meaningful but lengthy conversations (29).

The decreasing amount of time spent in meaningful conversations exacerbates the shortage of psychiatrists and psychologists. Psychiatrists' use of talk therapy has been consistently and steadily declining, meaning fewer patients are receiving talk therapy during psychiatric visits (30). In contrast to a human clinician's time and attention, conversational AI is relatively non-consumable, making it an attractive alternative to delivery of care by a human. If conversational AI is effective and acceptable to both patients and clinicians, it may address longstanding challenges to mental health access. These include the ability to accommodate rural populations and to facilitate increased engagement from people who may experience traditional talk therapy as stigmatizing (31).

Quality

Technology has been highlighted as a way to better understand and disseminate high quality psychotherapy (32, 33). Clinicians are already using texting services to deliver mental health interventions (34), which demonstrates a willingness by patients and clinicians to test new approaches to patient-clinician interaction. These new approaches facilitate novel measures of intervention quality. For example, innovations in computer science (e.g., natural language processing and machine learning) are being used to assess language patterns of successful crisis interventions in text-based mental health conversations (18, 35). Computational analysis of psychotherapy is encouraging researchers and companies to identify patterns of patient symptomology and therapist intervention (36, 37). This approach may improve psychotherapy quality by better understanding what effective clinicians actually do. This assessment has historically occurred through clinicians' self-reports or time intensive human audits (e.g., 38).

Although its efficacy is not definitively established, there are reasons to expect that conversational AI could constructively enhance mental health diagnosis and treatment delivery (39, 40). A diagnostic interview aids the patient and clinician in understanding the patient's presenting problem and provides a working model of how problems are being maintained. Approaches vary from highly structured diagnostic interviews

[e.g., Structured Clinical Interview for DSM-5 (41)] to unstructured interviews in which the conversation develops based on the clinician's expertise, training, and the patient's features. Conversational AIs have interviewed patients about symptoms for PTSD with a high level of patient acceptance (20). Conversational AI has been piloted across numerous clinically relevant groups such as clinical depression (6) and adolescent stress (42). In a study in which students believed they were speaking with a conversational AI, the students reported feeling better after talking about their problems following the encounter (43). Although these early findings point to potential benefits, there is a lack of rigorous clinical trial data and uncertainty about regulatory oversight (2).

Yet while there is reason for optimism, inflated or unsubstantiated expectations may frustrate patients and weaken their trust in psychotherapeutic interventions (44, 45). Many current computation methods can be used to search for specific dialogue acts, but additional work is needed to map theoretically important constructs (e.g., therapeutic alliance) to causal relationships between language patterns and clinically relevant outcomes. Psychotherapy quality will be difficult to assess without disentangling causal inferences and confounding factors. Beyond computation, patients' attitudes matter in psychotherapy because those who have a negative experience compared with their expectations have worse clinical outcomes (46). If a patient loses trust in a conversational AI, they may be less likely to trust human clinicians as well. As conversational AI becomes more sophisticated and expectations of benefit increase, there are growing concerns that users will transition from feeling let down to feeling betrayed (47). These factors suggest that careful experimentation about sub-processes in AI-mediated communication merits research attention.

Clinician–Patient Relationship

Modern medicine views the patient–clinician relationship as critical to patient health (48), and provider wellness (49). Indeed, appreciation of the importance of the patient–clinician relationship in modern medicine can be traced back to the influence of clinical psychology (50). Therapeutic alliance develops from clinicians' collaborative engagement with patients and reflects agreement on treatment goals, the tasks necessary to achieve such goals, and the affective bond between patient and provider (51). Therapeutic alliance is consistently associated with symptom improvement in psychotherapy (52–54). Numerous approaches exist to create alliance during psychotherapy, including the use of supportive language, mirroring emotions, and projecting warmth. Although originally conceptualized for human-to-human conversations, users have reported experiencing a sense of therapeutic alliance when speaking directly with conversational AI, suggesting this bond may not necessarily be restricted to human-human relationships (3). If conversational AI can create and maintain a therapeutic alliance, the provision of psychotherapy will not be necessarily limited by human clinicians' time and attention.

Establishing therapeutic alliance with conversational AI may benefit both patients and providers. By allowing conversational AI to take over repetitive, time-consuming tasks, clinicians' attention

and skill could be deployed more judiciously (55). Allowing clinicians to do less of the work that contributes to burnout, such as repetitive tasks performed with little autonomy, may improve clinicians' job satisfaction (56). Clinician burnout is associated with worse patient outcomes and is increasingly recognized as a problem which must be more adequately addressed (57, 58).

At the same time, software that augments clinical duties has been criticized for distancing clinicians from patient care (59). In mental health, this risk is especially salient because the content of therapy is often quite intimate. Some of the repetitive, time-consuming tasks clinicians engage in with patients, such as reviewing symptoms or taking their history, are precisely the vehicles by which clinicians connect with and understand their patients' experiences and develop rapport. It is unknown whether having a conversational AI listen in on psychotherapy will significantly impact patients' and clinicians' sense of therapeutic alliance. This area merits further research.

Patient Self-Disclosure and Sharing

Patient self-disclosure of personal information is crucial for successful therapy, including sensitive topics such as trauma, substance use, sexual history, forensic history, and thoughts of self-harm. Patient self-disclosures during psychotherapy are legally and ethically protected (24) and professional norms and laws have been established to set boundaries for what a clinician can share (60). Unauthorized sharing of identifiable patient information can result in fines, loss of license, or even incarceration. Moreover, because of the natural limitations of human memory, patients are unlikely to expect a human clinician to remember entire conversations perfectly in perpetuity. This capacity is in stark contrast to conversational AI, which has near-limitless capacity to hear, remember, share, and analyze conversations as long as desired. Because humans and machines have such different capacities, patient expectations of AI capabilities may impact treatment decisions and consent to data sharing (23).

In mental health, conversational AI has been shown to both facilitate and impede disclosure in different contexts. For example, users were more open with a conversational AI than with a human listener in reporting mental health symptoms (20), and have been successfully used to treat persecutory delusions for people with psychosis (61). Conversely, users were more reluctant to disclose sensitive information such as binge drinking behavior to a conversational AI compared to a non-responsive questionnaire (62). Because personal disclosures are central to diagnosis and treatment in psychotherapy, users' expectations and behavior towards technology-mediated conversations merit further assessment (63, 64, 65).

Certain disclosures in a psychotherapy context carry specific ethical and legal mandates, such as reporting suicidal or homicidal ideation. In 1969, a therapist at the University of California did not share the homicidal ideation of a patient with the intended victim. The patient subsequently killed the named victim, and the victim's family sued. This case (*Tarasoff v. Regents of the University of California, 1974*) established clinicians' duty not only to protect the confidentiality of their patients but also to notify individuals their patient might harm. A failure to warn leaves a clinician liable to civil judgment (66). Most case law and norms have

been established on the premise of a dyadic relationship between patient and clinician. The extent to which conversational AI inherits liability for harm is untested. As conversational AI takes on clinical duties and informs clinical judgment, expectations must be clarified about how and when these systems will respond to issues related to confidentiality, safety, and liability.

DISCUSSION

Experts in AI, clinicians, administrators, and other stakeholders recognize a need to more fully consider safety and trust in the design and deployment of new AI-based technologies (67, 68). A recent Lancet commission on global mental health states that "technology-based approaches might improve the reach of mental health services but could lose key human ingredients and, possibly, lower effectiveness of mental health care" (33). To inform future research directions, we have presented four approaches to integrating conversational AI into mental health delivery and discussed the dimensions of their impact.

Because conversational AI may augment the work of psychotherapy, we seek to encourage product designers, clinicians, and researchers to assess the impact of new practices on both patients and clinicians. Other areas of medicine have seen success with AI, such as lung cancer imaging and building diagnostic or prognostic models (69–73), and conversational AI for health is an emerging field with limited research on efficacy and safety (40, 63, 74).

Before we deploy AI-mediated treatment, workflow changes must be considered in the context of other demands on clinician time and training. Clinicians are already being asked to be familiar with telehealth (75) social media (76), and mobile health (77), while simultaneously being reminded of the need for self-care in light of clinician burnout (58). Before we insert new devices into clinical care, it will be crucial to engage clinicians and design evaluation strategies that appreciate the skills, attitudes, and knowledge of affected workers. Just as we can't expect technology companies to easily understand healthcare, we can't expect medical professionals to intuit or work in harmony with new technology without thoughtful design and training.

A limitation of this work is that we do not set out a specific research agenda, and some important considerations are beyond the scope of this work (e.g., the cost and feasibility of each approach). We propose instead that initiatives using conversational AI anticipate challenges and leverage lessons learned from existing approaches to deploying new technology in clinical settings that involve clinician training and patient protections from the start (32, 77). We instead encourage those proposing to put AI into care settings to directly consider and measure impact on access, quality, relationships, and data sharing.

The potential benefits are clear for mental health. If diagnosis or treatment can be done by conversational AI, the societal burden of treating mental health could be diminished. Additionally, conversational AI could have a more long-term relationship with a patient than clinicians who rotate out of training centers. Despite these potential benefits, technology carries risks related to privacy,

bias, coercion, liability, and data sharing that could harm patients in expected (e.g., denial of health insurance) and unintended ways (33, 44, 74, 78, 79–81). Conversations are valuable for patients and clinicians, and it is crucial to make sure they are delivered safely and effectively, regardless of who or what does the talking.

AUTHOR CONTRIBUTIONS

ASM and JH contributed to the initial conceptualization and design of the manuscript. AM wrote the first draft. NS, KDB, BAA, and JB contributed to manuscript revision, read and approved the submitted version.

REFERENCES

- Goldfried MR, Greenberg LS, Marmar C. Individual psychotherapy: process and outcome. *Annu Rev Psychol* (1990) 41(1):659–88. doi: 10.1146/annurev.ps.41.020190.003303
- Miner AS, Milstein A, Hancock JT. Talking to machines about personal mental health problems. *JAMA* (2017) 318(13):1217–8. doi: 10.1001/jama.2017.14151
- Bickmore T, Gruber A, Picard R. Establishing the computer–patient working alliance in automated health behavior change interventions. *Patient Educ Couns* (2005) 59(1):21–30. doi: 10.1016/j.pec.2004.09.008
- Rizzo A, Scherer S, DeVault D, Gratch J, Artstein R, Hartholt A, et al. Detection and computational analysis of psychological signals using a virtual human interviewing agent. *10th Intl Conf Disability, Virtual Reality & Associated Technologies*; Gothenburg, Sweden. (2014) Available at: http://ict.usc.edu/bibtexbrowser.php?key=rizzo_detection_2014&bib=ICT.bib (Accessed 15 Oct. 2018).
- Bickmore TW, Puskas K, Schlenk EA, Pfeifer LM, Sereika SM. Maintaining reality: relational agents for antipsychotic medication adherence. *Interact Comput* (2010) 22(4):276–88. doi: 10.1016/j.intcom.2010.02.001
- Fitzpatrick KK, 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 Ment Health* (2017) 4(2):e19. doi: 10.2196/mental.7785
- Oh KJ, Lee D, Ko B, Choi HJ. A chatbot for psychiatric counseling in mental healthcare service based on emotional dialogue analysis and sentence generation. *Mobile Data Management (MDM), 2017 18th IEEE International Conference*; 29.IEEE. (2017) pp. 371–375. doi: 10.1109/MDM.2017.64
- Imel ZE, Steyvers M, Atkins DC. Computational psychotherapy research: scaling up the evaluation of patient–provider interactions. *Psychotherapy* (2015) 52(1):19. doi: 10.1037/a0036841
- Crutzen R, Peters GJY, Portugal SD, Fisser EM, Grolleman JJ. An artificially intelligent chat agent that answers adolescents' questions related to sex, drugs, and alcohol: an exploratory study. *J Adolesc Health* (2011) 48(5):514–9. doi: 10.1016/j.jadohealth.2010.09.002
- Martínez-Miranda J. Embodied conversational agents for the detection and prevention of suicidal behaviour: current applications and open challenges. *J Med Syst* (2017) 41(9):135. doi: 10.1007/s10916-017-0784-6
- Bailenson JN, Beall AC, Loomis J, Blascovich J, Turk M. Transformed social interaction: decoupling representation from behavior and form in collaborative virtual environments. *Presence: Teleop Virt Environ* (2004) 13(4):428–41. doi: 10.1162/1054746041944803
- Luxton DD. Recommendations for the ethical use and design of artificial intelligent care providers. *Artif Intell Med* (2014) 62(1):1–0. doi: 10.1016/j.artmed.2014.06.004
- Hyman L. *Temp: how American work, American business, and the American dream became temporary*. New York, New York: Penguin Random House (2018). ISBN: 9780735224070.
- Luxton DD. *Artificial intelligence in behavioral and mental health care*. Elsevier/Academic Press (2016). doi: 10.1016/B978-0-12-420248-1.00001-5
- Rehm IC, Foenander E, Wallace K, Abbott JA, Kyrios M, Thomas N. What role can avatars play in e-mental health interventions? *Front Psychiatry* (2016) 7:186. doi: 10.3389/fpsy.2016.00186
- American Psychiatric Association. *The principles of medical ethics with annotations especially applicable to psychiatry*. (2001) Washington, DC: Author
- American Psychological Association. Ethical principles of psychologists and code of conduct. *Am Psychol* (2002) 57(12):1060–73. doi: 10.1037/0003-066X.57.12.1060
- Althoff T, Clark K, Leskovec J. Large-scale analysis of counseling conversations: an application of natural language processing to mental health. *Trans Assoc Comput Lingu* (2016) 4:463. doi: 10.1162/tacl_a_00111
- Xiao B, Imel ZE, Georgiou PG, Atkins DC, Narayanan SS. Rate my therapist: automated detection of empathy in drug and alcohol counseling via speech and language processing. *PLoS One* (2015) 10(12):e0143055. doi: 10.1371/journal.pone.0143055
- Lucas GM, Gratch J, King A, Morency LP. It's only a computer: virtual humans increase willingness to disclose. *Comput Hum Behav* (2014) 37:94–100. doi: 10.1016/j.chb.2014.04.043
- Goldstein IM, Lawrence J, Miner AS. Human–machine collaboration in cancer and beyond: the centaur care model. *JAMA Oncol* (2017) 3(10):1303–4. doi: 10.1001/jamaoncol.2016.6413
- Metaxiotis KS, Samouilidis JE. Expert systems in medicine: academic illusion or real power? *Inform Manage Comput Secur* (2000) 8(2):75–9. doi: 10.1108/09685220010694017
- Martínez-Martin N, Dunn LB, Roberts LW. Is it ethical to use prognostic estimates from machine learning to treat psychosis? *AMA J Ethics* (2018) 20(9):804–11. doi: 10.1001/amajethics.2018.804
- Roberts LW. *A clinical guide to psychiatric ethics*. Arlington, VA: American Psychiatric Pub (2016). ISBN: 978-1-61537-049-8.
- Kazdin AE, Rabbitt SM. Novel models for delivering mental health services and reducing the burdens of mental illness. *Clin Psychol Sci* (2013) 1(2):170–91. doi: 10.1177/2167702612463566
- Kazdin AE, Blase SL. Rebooting psychotherapy research and practice to reduce the burden of mental illness. *Perspect Psychol Sci* (2011) 6(1):21–37. doi: 10.1177/1745691610393527
- Dieleman JL, Baral R, Birger M, Bui AL, Bulchis A, Chapin A, et al. US spending on personal health care and public health, 1996–2013. *JAMA* (2016) 316(24):2627–46. doi: 10.1001/jama.2016.16885
- Olfson M. Building the mental health workforce capacity needed to treat adults with serious mental illnesses. *Health Affairs* (2016) 35(6):983–90. doi: 10.1377/hlthaff.2015.1619
- Kaplan RS, Haas DA, Warsh J. Adding value by talking more. *N Engl J Med* (2016) 375(20):1918–20. doi: 10.1056/NEJMp1607079
- Mojtabai R, Olfson M. National trends in psychotherapy by office-based psychiatrists. *Arch Gen Psychiatry* (2008) 65(8):962–70. doi: 10.1001/archpsyc.65.8.962
- Perle JG, Langsam LC, Nierenberg B. Controversy clarified: an updated review of clinical psychology and tele-health. *Clin Psychol Rev* (2011) 31(8):1247–58. doi: 10.1016/j.cpr.2011.08.003

ACKNOWLEDGMENTS

This work was supported by grants from the National Institutes of Health, National Center for Advancing Translational Science, Clinical and Translational Science Award (KL2TR001083 and UL1TR001085), the Stanford Department of Psychiatry Innovator Grant Program, and the Stanford Institute for Human-Centered Artificial Intelligence. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH. We thank Nicole Martínez-Martin JD PhD, Victor Henderson MD MS, and Stan Fisher for their valuable feedback. Reference formatting assisted by Charlesworth Author Services.

32. Mohr DC, Schueller SM, Montague E, Burns MN, Rashidi P. The behavioral intervention technology model: an integrated conceptual and technological framework for eHealth and mHealth interventions. *J Med Internet Res* (2014) 16(6):e146. doi: 10.2196/jmir.3077
33. Patel V, Saxena S, Lund C, Thornicroft G, Baingana F, Bolton P, et al. The Lancet Commission on global mental health and sustainable development. *Lancet* (2018) 392(10157):1553–98. doi: 10.1016/S0140-6736(18)31612-X
34. Schaub MP, Wenger A, Berg O, Beck T, Stark L, Buehler E, et al. A web-based self-help intervention with and without chat counseling to reduce cannabis use in problematic cannabis users: three-arm randomized controlled trial. *J Med Internet Res* (2015) 17(10):e232. doi: 10.2196/jmir.4860
35. Dinakar K, Chen J, Lieberman H, Picard R, Filbin R. Mixed-initiative real-time topic modeling & visualization for crisis counseling. *Proceedings of the 20th International Conference on Intelligent User Interfaces*; Atlanta GA: ACM. (2015) pp. 417–426. doi: 10.1145/2678025.2701395
36. Owen J, Imel ZE. Introduction to the special section “Big ‘er’Data”: scaling up psychotherapy research in counseling psychology. *J Couns Psych* (2016) 63(3):247. doi: 10.1037/cou0000149
37. Iter D, Yoon J, Jurafsky D. Automatic detection of Incoherent speech for diagnosing schizophrenia. *Proceedings of the Fifth Workshop on Computational Linguistics and Clinical Psychology: From Keyboard to Clinic*; New Orleans, LA. (2018) pp.136–146. doi: 10.18653/v1/W18-0615
38. Cook JM, Biyanova T, Elhai J, Schnurr PP, Coyne JC. What do psychotherapists really do in practice? An Internet study of over 2,000 practitioners. *Psychotherapy* (2010) 47(2):260. doi: 10.1037/a0019788
39. Haque A, Guo M, Miner AS, Fei-Fei L. Measuring depression symptom severity from spoken language and 3D facial expressions. Paper presented at NeurIPS 2018 Workshop on Machine Learning for Health. Montreal, Canada (2018).
40. Laranjo L, Dunn AG, Tong HL, Kocaballi AB, Chen J, Bashir R, et al. Conversational agents in healthcare: a systematic review. *J Am Med Inform Assoc* (2018) 25(9):1248–58. doi: 10.1093/jamia/ocy072
41. First MB, Williams JBW, Karg RS, Spitzer RL. *Structured clinical interview for DSM-5 disorders, clinician version (SCID-5-CV)*. Arlington, VA: American Psychiatric Association (2016).
42. Huang J, Li Q, Xue Y, Cheng T, Xu S, Jia J, et al. Teenchat: a chatterbot system for sensing and releasing adolescents’ stress. *International Conference on Health Information Science*; Queensland, Australia. Springer, Cham (2015) pp. 133–145. doi: 10.1007/978-3-319-19156-0_14
43. Ho A, Hancock J, Miner AS. Psychological, relational, and emotional effects of self-disclosure after conversations with a Chatbot. *J Commun* (2018) 68(4):712–33. doi: 10.1093/joc/jqy026
44. Aboujaoude E. Telemental health: Why the revolution has not arrived. *World Psychiatry* (2018) 17(3):277. doi: 10.1002/wps.20551
45. Miner AS, Milstein A, Schueller S, Hegde R, Mangurian C, Linos E. Smartphone-based conversational agents and responses to questions about mental health, interpersonal violence, and physical health. *JAMA Int Med* (2016) 176(5):619–25. doi: 10.1001/jamainternmed.2016.0400
46. Watsford C, Rickwood D. Disconfirmed expectations of therapy and young people’s clinical outcome, help-seeking intentions, and mental health service use. *Adv MentHealth* (2013) 12(1):75–86. doi: 10.5172/jamh.2013.12.1.75
47. Brooker N. “We should be nicer to Alexa.” *Financial Times* (2013). <https://www.ft.com/content/4399371e-bcbd-11e8-8274-55b72926558f> (Accessed October 15, 2018).
48. Martin DJ, Garske JP, Davis MK. Relation of the therapeutic alliance with outcome and other variables: a meta-analytic review. *J Consult Clin Psychol* (2000) 68(3):438. doi: 10.1037//0022-006X.68.3.438
49. Rosenthal DI, Verghese A. Meaning and the nature of physicians’ work. *N Engl J Med* (2016) 375(19):1813–5. doi: 10.1056/NEJMp1609055
50. Szasz TS, Hollender MH. A contribution to the philosophy of medicine: the basic models of the doctor–patient relationship. *AMA Arch Int Med* (1956) 97(5):585–92. doi: 10.1001/archinte.1956.00250230079008
51. Horvath AO, Greenberg LS. Development and validation of the working alliance inventory. *J Couns Psych* (1989) 36(2):223. doi: 10.1037//0022-0167.36.2.223
52. Flückiger C, Del Re AC, Wampold BE, Symonds D, Horvath AO. How central is the alliance in psychotherapy? A multilevel longitudinal meta-analysis. *J Couns Psychol* (2012) 59(1):p.10. doi: 10.1037/a0025749
53. Horvath AO, Del Re AC, Flückiger C, Symonds D. Alliance in individual psychotherapy. *Psychotherapy* (2011) 48(1):9. doi: 10.1037/a0022186
54. Norcross JC ed. *Psychotherapy relationships that work: therapist contributions and responsiveness to patient needs*. New York: Oxford University Press (2002).
55. Jha S, Topol EJ. Adapting to artificial intelligence: radiologists and pathologists as information specialists. *JAMA* (2016) 316(22):2353–4. doi: 10.1001/jama.2016.17438
56. Harvey SB, Modini M, Joyce S, Milligan-Saville JS, Tan L, Mykletun A, et al. Can work make you mentally ill? A systematic meta-review of work-related risk factors for common mental health problems. *Occup Environ Med* (2017) 74(4):301–10. doi: 10.1136/oemed-2016-104015
57. Delgadillo J, Saxon D, Barkham M. Associations between therapists’ occupational burnout and their patients’ depression and anxiety treatment outcomes. *Depress Anxiety* (2018) 35:844–50. doi: 10.1002/da.22766
58. Panagioti M, Panagopoulou E, Bower P, Lewith G, Kontopantelis E, Chew-Graham C, et al. Controlled interventions to reduce burnout in physicians: a systematic review and meta-analysis. *JAMA Int Med* (2017) 177(2):195–205. doi: 10.1001/jamainternmed.2016.7674
59. Verghese A. Culture shock-patient as icon, icon as patient. *N Engl J Med* (2008) 359(26):2748–51. doi: 10.1056/NEJMp0807461
60. Edwards G. Doing their duty: an empirical analysis of the unintended effect of Tarasoff v. Regents on homicidal activity. *J Law and Econ* (2014) 57(2):321–48. doi: 10.1086/675668
61. Craig TK, Rus-Calafell M, Ward T, Leff JP, Huckvale M, Howarth E, et al. AVATAR therapy for auditory verbal hallucinations in people with psychosis: a single-blind, randomised controlled trial. *Lancet Psychiatry* (2018) 5(1):31–40. doi: 10.1016/S2215-0366(17)30427-3
62. Schuetzler RM, Giboney JS, Grimes GM, Nunamaker JF The influence of conversational agents on socially desirable responding. *Proceedings of the 51st Hawaii International Conference on System Sciences*; Waikoloa Village, Hawaii. (2018) pp. 283–292. ISBN: 978-0-9981331-1-9. doi: 10.24251/HICSS.2018.038
63. Bickmore T, Trinh H, Asadi R, Olafsson S. (2018a) Safety first: Conversational agents for health care. In: Moore, R, Szymanski, M, Arar, R, Ren, GJ, editors. *Studies in Conversational UX Design. Human–Computer Interaction Series*. Springer, Cham. doi: 10.1007/978-3-319-95579-7_3
64. French M, Bazarova NN. Is anybody out there?: understanding masspersonal communication through expectations for response across social media platforms. *J Comput Mediat Commun* (2017) 22(6):303–19. doi: 10.1111/jcc4.12197
65. Liu B, Sundar SS. Should machines express sympathy and empathy? Experiments with a health advice chatbot. *Cyberpsychol Behav Soc Netw* (2018) 21(10):625–36. doi: 10.1089/cyber.2018.0110
66. Swerdlow BA. Tracing the evolution of the Tarasoff Duty in California. *J Sociol Soc Welfare* (2018) 45:25.
67. Bhugra D, Tasman A, Pathare S, Priebe S, Smith S, Torous J, et al. The WPA-lancet psychiatry commission on the future of psychiatry. *Lancet Psychiatry* (2017) 4(10):775–818. doi: 10.1016/S2215-0366(17)30333-4
68. Stone P, Brooks R, Brynjolfsson E, Calo R, Etzioni, O, et al. “Artificial Intelligence and Life in 2030.” One Hundred Year Study on Artificial Intelligence: Report of the 2015-2016 Study Panel, Stanford University, Stanford, CA. (2016) Doc: <http://ai100.stanford.edu/2016-report>. (accessed October 15, 2018).
69. Avati A, Jung K, Harman S, Downing L, Ng A, Shah NH. Improving palliative care with deep learning. *IEEE International Conference on Bioinformatics and Biomedicine*; Kansas City:IEEE, MO. (2017). pp. 311–316. doi: 10.1109/BIBM.2017.8217669
70. Jung K, Covington S, Sen CK, Januszky M, Kirsner RS, Gurtner GC, et al. Rapid identification of slow healing wounds. *Wound Repair Regen* (2016) 24(1):181–8. doi: 10.1111/wrr.12384

71. Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA* (2016) 316(22):2402–10. doi: 10.1001/jama.2016.17216
72. Pusiol G, Esteva A, Hall SS, Frank M, Milstein A, Fei-Fei L. Vision-based classification of developmental disorders using eye-movements. *International Conference on Medical Image Computing and Computer-Assisted Intervention*; Athens, Greece: Springer, Cham (2016) pp. 317–325. doi: 10.1007/978-3-319-46723-8_37
73. Yu KH, Zhang C, Berry GJ, Altman RB, Ré C, Rubin DL, et al. Predicting non-small cell lung cancer prognosis by fully automated microscopic pathology image features. *Nat Commun* (2016) 7:12474. doi: 10.1038/ncomms12474
74. Bai G, Jiang JX, Flasher R. Hospital risk of data breaches. *JAMA Int Med* (2017) 177(6):878–80. doi: 10.1001/jamainternmed.2017.0336
75. Maheu MM, Drude KP, Hertlein KM, Lipschutz R, Wall K, Hilty DM. An interprofessional framework for telebehavioral health competencies. *J Technol Behav Sci* (2017) 2(3–4):190–210. doi: 10.1007/s41347-017-0038-y
76. Zalpuri I, Liu HY, Stubbe D, Wrzosek M, Sadhu J, Hilty D. Social media and networking competencies for psychiatric education: skills, teaching methods, and implications. *Acad Psychiatry* (2018) 42(6):808–17. doi: 10.1007/s40596-018-0983-6
77. Hilty DM, Chan S, Torous J, Luo J, Boland RJ. A Telehealth framework for mobile health, smartphones, and apps: competencies, training, and faculty development. *J Technol Behav Sci* (2019) 1–18. doi: 10.1007/s41347-019-00091-0
78. Bickmore TW, Trinh H, Olafsson S, O’Leary TK, Asadi R, Rickles NM, et al. (2018b) Patient and consumer safety risks when using conversational assistants for medical information: an observational study of Siri, Alexa, and google assistant. *J Med Int Res* 20(9):e11510. doi: 10.2196/11510
79. Caliskan A, Bryson JJ, Narayanan A. Semantics derived automatically from language corpora contain human-like biases. *Science* (2017) 356(6334):183–6. doi: 10.1126/science.aal4230
80. De Choudhury M, Sharma SS, Logar T, Eekhout W, Nielsen RC. Gender and cross-cultural differences in social media disclosures of mental illness. *Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*; Portland, OR. (2017) pp. 353–369. doi: 10.1145/2998181.2998220
81. Martinez-Martin N, Kreitmair K. Ethical issues for direct-to-consumer digital psychotherapy apps: addressing accountability, data protection, and consent. *JMIR Ment Health* (2018) 5(2):e32. doi: 10.2196/mental.9423

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2019 Miner, Shah, Bullock, Arnow, Bailenson and Hancock. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.