Reading Response

Key Considerations for Incorporating Conversational AI into Psychotherapy Miner et al (2019)

This reading response will center on the paper *Key Considerations for Incorporating Conversational AI into Psychotherapy* (Miner et al. 2019). This paper introduces four approaches to AI-human integration in psychotherapy with four dimensions of impact on each approach. These four approaches being: Humans only, Humans delivered with AI informed, AI delivered with human supervised, and AI only. The four dimensions being: access to care, quality, clinician-patient relationship, and patient self-disclosure and sharing. This paper analyzes the potential benefits and harm of the four approaches with these four dimensions.

Beginning with access to care, the paper explains how using a human only and human delivered with AI approach is no longer feasible due to a declining number of practitioners in the psychotherapy industry. For the third approach, AI delivered with human supervision would still require a reliance on human clinicians but with a human supervising potentially multiple clients in a short period of time. This would not be scalable as the number of clients increases, but is still an improvement to the previous two approaches. Lastly, the AI only approach would improve access to care generally as machines would not be limited by human attention.

In terms of the quality dimension, the paper states that a human only approach would have no change as that is already the most commonly practiced. A human delivered with AI informed approach would see potential improvements as AI can detect symptoms that are potentially missed by the human. The last two approaches lack rigorous clinical data to conclude potential improvements or harm.

Similarly, a clinical-patient relationship for the human only approaches would bear no changes. There would be disruption for human delivered with AI informed and AI delivered with human supervision approaches as the clinical-patient relationship would have an external AI intruding in the relationship between the clinician and the patient. For AI only, there would be a non-existing clinician which would cease any clinician-patient relationship.

Finally, patient self-disclosure and sharing remains the same for the human only approach with not enough clinical research for the remaining three approaches.

A critical point

One critical point I would like to develop more about the reading is the access to care for the AI only approach. This reading makes the argument that because the AI only approach is not limited by human attention, it improves access to care for patients. This argument has a fair claim, an AI only approach is not limited by human attention, the declining rate of human clinicians would not affect this approach.

The other side of how an AI only approach does not improve access to care has a stronger argument than what the paper claims. Assuming that it's possible that an AI model is just as competent as a human clinician, there would be no guarantee that rural populations will be able to get this type of treatment without exerting vast amounts of money. Less financially fortunate geographic locations would have the same access to care, almost none at all, due to the lack of money that would be needed to afford a fully automated AI only approach.

This side would maintain a viewpoint on how an approach that improves access to care regardless of financial status would be a true approach that improves access to care.