Students’ attitudes towards AI in Psychiatry

Anne-Kathrin Kleine

# Seed for random number generation  
set.seed(42)

# Introduction

The application of AI technology in mental healthcare is on the rise Jiang et al. (2017). Although many applications are still in development, several AI products that aid in diagnosing and treating mental illnesses, selecting appropriate treatment approaches, predicting treatment success, and improving psycho-therapeutic interventions are already available to mental health professionals Cummins et al. (2019). Given the large number of registered patents related to identifying and diagnosing psychiatric illnesses, the number of AI applications in the mental health sector may be expected to increase further over the following years (Zheng et al., 2022).

# The application of AI technology in mental health care

The application of AI technology in mental healthcare ranges from selecting appropriate treatment regimens to supporting therapist skill development. In their scoping review of machine learning in psychotherapy research, Aafjes-van Doorn et al. (2021) identified 51 studies that developed and tested a machine learning algorithm focused on selecting appropriate treatment regimes, predicting treatment adherence and response, and therapist skill development. Shatte et al. (2019) identified 190 mental health tools focused on detecting and diagnosing mental health conditions, 67 on prognosis, treatment, and support, 26 on public health applications, and 17 on research and clinical administration. In the following, we provide a brief overview of how AI technology is used in the areas of a) assistance with detection, diagnosis, and prognosis, b) assistance with treatment and treatment selection, and c) improvement of psychotherapy quality and therapist skill development. Detailed descriptions of AI technologies relevant for clinical practice in the different areas of mental health care have been provided elsewhere (e.g., Lee et al., 2021).

## Assistance with detection, diagnosis, and prognosis

## Assistance with treatment and treatment selection

## Improvement of psychotherapy quality

Supervision and receiving performance-feedback on their therapy sessions supports psychotherapy trainees’ acquisition of skills and increases retention Helge Rønnestad & Ladany (2006). However, the process of providing ongoing feedback is labor intensive and thus rarely used in training. Feedback is often based on trainees’ self-reports and usually only available long after the actual performance (Tanana et al., 2019). Using artificial intelligence technology for training purposes in mental health care may leverage the problem by providing immediate and performance-specific feedback to psychotherapists and trainees.

Most of the tools developed to improve psychotherapy quality rely on natural language processing-based feedback Tanana et al. (2019). For example, *TIM* (Therapy Insights Model) uses real-time chat messages exchanged between therapists and patients to provide therapists with feedback regarding the topics that were sufficiently covered in the session and the topics that should be addressed in the next sessions (Cummins et al., 2019). *CORE-MI* (Counselor Observer Ratings Expert for Motivational Interviewing uses audio recordings of motivational interviewing (MI) sessions to generate feedback on psychotherapists’ adherence with MI principles. The user receives feedback on six summary measures of MI fidelity: empathy, MI spirit, reflection-to-question ratio, percent open questions, percent complex reflections, and percent MI adherence. *CORE-MI* includes a visual summary of counseling sessions based on the fidelity assessment that may be used by the therapist to improve their MI performance (Hirsch et al., 2018). Similar tools include the *ClientBot*, a training tool that mimics typical patient responses to therapist questions and provides real-time feedback on therapists’ use of open questions and reflections (Tanana et al., 2019); or *Partner*, a reinforcement learning agent that may increase the quality of mental health support conversations by suggesting sentence-level edits to posts that enhance the level of empathy while maintaining conversation quality (Sharma et al., 2021).

# Concerns regarding the application of AI technology in mental health practice

Thus far, insight into health practitioners’ attitudes towards the application of AI technology is limited to the physical health domain, with inconclusive results Pinto dos Santos et al. (2019). One study suggests that despite the strong belief in the efficiency of AI tools, only a minority (a quarter) of radiologists in the sample believed that AI tools can be trusted (Jungmann et al., 2020). Other research findings suggest generally positive attitudes towards the application of AI in radiology (e.g., Codari et al., 2019).

Despite a large amount of research in the area of mental health AI technology, many of the algorithms developed and tested in research have not yet been integrated into clinical care (Sendak et al., 2020). Reasons for this research-practice gap include technical difficulties and the black box problem of AI-based recommendations. Technical difficulties primarily complicate the use of AI tools in diagnosis and prognosis and selecting treatment approaches. They encompass, for example, the inaccuracy of predictions, accuracy-interpretability tradeoffs, data privacy issues, and the difficulty of deriving clear-cut diagnoses in the light of ambiguous symptomatology or symptoms overlapping with other (mental) illnesses Aafjes-van Doorn et al. (2021).

In addition, is often not apparent to clinicians how AI-based recommendations are generated. Despite the efforts made to enhance the explainability of AI recommendations, such as the Explainable Artificial Intelligence (XAI) Initiative, the complexity of deep learning approaches necessarily limits the extent to which they can be made accessible to a broader user group (Feldman et al., 2019). Especially in the area of mental health care, where transparency and the explainability of clinical decision-making are highly valued, the black box problem of AI-based recommendations creates a significant obstacle against its adoption into clinical practice Kelly et al. (2019).

## Concerns among medical and psychology students and early practitioners

*Research Question 1*: Are there any differences in the attitude towards the application of AI technology in mental health care between medical students and psychology students?

# The relationship between AI knowledge and attitudes towards the application of AI technology in mental health care

Technical difficulties concerning the generation of accurate predictions and a seamless adoption of AI tools into clinical practice may be and will be addressed by data scientists and mental health experts over the coming years. It has been suggested that reluctance against the adoption of new technology remains high as long as there is a lack of understanding on the practitioner side (Feldman et al., 2019).

A basic understanding how AI recommendations are generated makes it easier for clinicians to a) explain their decisions to patients and stakeholders, b) make fewer errors in using the generated information in their practice, c) provide feedback to help improving AI-based diagnostic and treatment tools, d) improve their professional skills (Adadi & Berrada, 2018).

*Hypothesis 1:*: There is a positive relationship between the knowledge about AI tools used in a) assistance with detection, diagnosis, and prognosis, b) assistance with treatment and treatment selection, and c) improvement of psychotherapy quality and positive attitudes towards the application of AI technology in the respective area.

# The moderating role of need satisfaction

|  |
| --- |
| Need satisfaction scale |

* Idea: Adapt scale items to need satisfaction threat in future job; for example:
  + “I feel that AI tools used to [area of mental health] will make it more difficult to express my ideas and opinions in this job.”
  + “AI tools used to [area of mental health] will make it more difficult to be myself at my job.”

*Hypothesis 2:*: The positive relationship between the knowledge about AI tools used in different mental health areas and positive attitudes towards the application of AI technology in these areas is attenuated by perceived threats to a) the need for autonomy, b) the need for competence, c) the need for relatedness posed by AI technology in the respective mental health area.

# Experimental setup

The ability to interpret AI algorithms’ mistakes and formulate the best strategies to correct these applications requires specialized training. Consequently, medical and health informatics education must emphasize algorithm-based platforms, and include relevant data analytics and AI topics in their curricula. Moreover, computer science and health informatics programs should consist of health care–focused digital skills training.

Medical students need to become familiar with clinical AI applications and predictive modeling techniques to assess biased data and evaluate innovative AI technologies.

# Potential control variables (Other predictors of AI attitudes)

* personality (**park\_woo22?**)

## AI Attitudes

* affectice, cognitive, behavioral attitudes (**park\_woo22?**)
  + affective: Positive emotional attitude (or positive emotion) was measured with four emotions: excitement, amusement, happiness, and surprise. Negative emotional attitude (or negative emotion) was measured with four emotions: unease, confusion, fear, and disappointment.
  + cognitive: Needs satisfaction was assessed with six items measuring each of the following three needs with two items: autonomy, competence, and relatedness. Consistent with the literature where need satisfaction is measured within a specific domain such as work (Broeck et al., 2010), we modified items to make it directly applicable to AI. We used a general basic needs satisfaction measure validated in Korean (Lee & Kim, 2008) and among original 18 items, we used six items due to survey length constraints. For example, “AI would allow me to decide things for myself and think more independently (autonomy), “Using AI would make me feel more capable and effective in performing tasks (competence)” and “Using AI would help me to build more reliable relationships with others (relatedness).” The items were rated on a 7-point Likert scale (1=Strongly Disagree, 7=Strongly Agree). Our data showed adequate levels of internal consistency reliability for all three needs (α = .78 for autonomy need satisfaction, α = .66 for competence need satisfaction, and α = .89 for relatedness need satisfaction).
  + emotional and competence: The warmth perception was measured with two items (α = .72) and competence perception was measured with two items (α = .81) developed by Fiske and colleagues (2002). The items have shown evidence of being a valid and reliable scale among Korean participants (Kim et al., 2019). Example items include “I think that AI is warm (warmth perception)” and “I think that AI is efficient (competence perception).” These items were rated on a 5-point Likert scale (1 = Strongly Disagree, 5 = Strongly Agree). In addition, behavioral intent to use AI, which has been widely used in prior studies (e.g. Venkatesh & Davis, 2000), was measured with the following item on a 7-point Likert scale (1 = Strongly Disagree, 7 = Strongly Agree). The item is “I intend to use products and service using AI.”

### Attitudes towards AI in radiology

* The first section was aimed at evaluating whether the students had already heard of deep learning and AI in the context of radiology and whether they felt they had a basic grasp of the underlying technologies
* The second section was introduced with a statement that AI is already being employed in relatively common software such as speech- and text-recognition, spam-filters and recommendation algorithms. The respondents were then asked to state whether or not they had already heard this in the media, on social media, during lectures or from friends/family
* In the third and fourth sections, the students were presented with various statements and asked to indicate their level of agreement on a four-point Likert scale (disagree entirely, rather disagree, rather agree, or agree entirely). While the third section aimed at specific possible applications of AI in radiology, the fourth section had a broader scope and tried to assess the students’ general fear of algorithms replacing human radiologists and other physicians.
* The last section consisted of questions regarding respondent demographics as well as one question regarding whether the respondents considered themselves tech-savvy or not

### Recruitment

* (**pintodossantosMedicalStudentsAttitude2019?**):
  + The questionnaire was sent out via email and advertised on social media to undergraduate medical students at three major German universities (University of Cologne, University of Bonn, University of Mainz). Participation was voluntary and had no relation to the students’ curricular activities. The students were informed that the results of the survey would be used for further statistical evaluation and scientific publication. Respondent anonymity was guaranteed by design.

# References

Aafjes-van Doorn, K., Kamsteeg, C., Bate, J., & Aafjes, M. (2021). A scoping review of machine learning in psychotherapy research. *Psychotherapy Research*, *31*(1), 92–116. <https://doi.org/10.1080/10503307.2020.1808729>

Adadi, A., & Berrada, M. (2018). Peeking Inside the Black-Box: A Survey on Explainable Artificial Intelligence (XAI). *IEEE Access*, *6*, 52138–52160. <https://doi.org/10.1109/ACCESS.2018.2870052>

Atkins, D. C., Steyvers, M., Imel, Z. E., & Smyth, P. (2014). Scaling up the evaluation of psychotherapy: Evaluating motivational interviewing fidelity via statistical text classification. *Implementation Science: IS*, *9*, 49. <https://doi.org/10.1186/1748-5908-9-49>

Can, D., Marín, R. A., Georgiou, P. G., Imel, Z. E., Atkins, D. C., & Narayanan, S. S. (2016). "It sounds like...": A natural language processing approach to detecting counselor reflections in motivational interviewing. *Journal of Counseling Psychology*, *63*(3), 343–350. <https://doi.org/10.1037/cou0000111>

Chekroud, A. M., Bondar, J., Delgadillo, J., Doherty, G., Wasil, A., Fokkema, M., Cohen, Z., Belgrave, D., DeRubeis, R., Iniesta, R., Dwyer, D., & Choi, K. (2021). The promise of machine learning in predicting treatment outcomes in psychiatry. *World Psychiatry*, *20*(2), 154–170. <https://doi.org/10.1002/wps.20882>

Chen, Z. S., Prathamesh, Kulkarni, Galatzer-Levy, I. R., Bigio, B., Nasca, C., & Zhang, Y. (2022). *Modern Views of Machine Learning for Precision Psychiatry*. arXiv. <http://arxiv.org/abs/2204.01607>

Codari, M., Melazzini, L., Morozov, S. P., Kuijk, C. C. van, Sconfienza, L. M., Sardanelli, F., & European Society of Radiology (ESR). (2019). Impact of artificial intelligence on radiology: A EuroAIM survey among members of the European Society of Radiology. *Insights into Imaging*, *10*(1), 105. <https://doi.org/10.1186/s13244-019-0798-3>

Cummins, R., Ewbank, M., Martin, A., Tablan, V., Catarino, A., & Blackwell, A. (2019). *TIM: A Tool for Gaining Insights into Psychotherapy*. <https://doi.org/10.1145/3308558.3314128>

Feldman, R., Aldana, E., & Stein, K. (2019). Artificial Intelligence in the Health care Space: How We Can Trust What We Cannot Know. *Policy Review*, *30*, 23. <https://repository.uchastings.edu/faculty_scholarship/1753/>

Gong, B., Nugent, J. P., Guest, W., Parker, W., Chang, P. J., Khosa, F., & Nicolaou, S. (2019). Influence of Artificial Intelligence on Canadian Medical Students’ Preference for Radiology Specialty: ANational Survey Study. *Academic Radiology*, *26*(4), 566–577. <https://doi.org/10.1016/j.acra.2018.10.007>

Helge Rønnestad, M., & Ladany, N. (2006). The impact of psychotherapy training: Introduction to the special section. *Psychotherapy Research*, *16*(3), 261–267. <https://doi.org/10.1080/10503300600612241>

Hirsch, T., Soma, C., Merced, K., Kuo, P., Dembe, A., Caperton, D. D., Atkins, D. C., & Imel, Z. E. (2018). “It’s hard to argue with a computer:” Investigating Psychotherapists’ Attitudes towards Automated Evaluation. *DIS. Designing Interactive Systems (Conference)*, *2018*, 559–571. <https://doi.org/10.1145/3196709.3196776>

Jiang, F., Jiang, Y., Zhi, H., Dong, Y., Li, H., Ma, S., Wang, Y., Dong, Q., Shen, H., & Wang, Y. (2017). Artificial intelligence in healthcare: Past, present and future. *Stroke and Vascular Neurology*, *2*(4), 230–243. <https://doi.org/10.1136/svn-2017-000101>

Jungmann, F., Jorg, T., Hahn, F., Pinto dos Santos, D., Jungmann (geb. Görgen), S., Düber, C., Mildenberger, P., & Kloeckner, R. (2020). Attitudes Toward Artificial Intelligence Among Radiologists, IT Specialists, and Industry. *Academic Radiology*, *28*. <https://doi.org/10.1016/j.acra.2020.04.011>

Kelly, C. J., Karthikesalingam, A., Suleyman, M., Corrado, G., & King, D. (2019). Key challenges for delivering clinical impact with artificial intelligence. *BMC Medicine*, *17*(1), 195. <https://doi.org/10.1186/s12916-019-1426-2>

Lee, E. E., Torous, J., De Choudhury, M., Depp, C. A., Graham, S. A., Kim, H.-C., Paulus, M. P., Krystal, J. H., & Jeste, D. V. (2021). Artificial Intelligence for Mental Health Care: Clinical Applications, Barriers, Facilitators, and Artificial Wisdom. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, *6*(9), 856–864. <https://doi.org/10.1016/j.bpsc.2021.02.001>

Megerian, J. T., Dey, S., Melmed, R. D., Coury, D. L., Lerner, M., Nicholls, C. J., Sohl, K., Rouhbakhsh, R., Narasimhan, A., Romain, J., Golla, S., Shareef, S., Ostrovsky, A., Shannon, J., Kraft, C., Liu-Mayo, S., Abbas, H., Gal-Szabo, D. E., Wall, D. P., & Taraman, S. (2022). Evaluation of an artificial intelligence-based medical device for diagnosis of autism spectrum disorder. *Npj Digital Medicine*, *5*(1), 57. <https://doi.org/10.1038/s41746-022-00598-6>

Miller, D. D., & Brown, E. W. (2018). Artificial Intelligence in Medical Practice: The Question to the Answer? *The American Journal of Medicine*, *131*(2), 129–133. <https://doi.org/10.1016/j.amjmed.2017.10.035>

Moyers, T. B., Martin, T., Manuel, J. K., Hendrickson, S. M. L., & Miller, W. R. (2005). Assessing competence in the use of motivational interviewing. *Journal of Substance Abuse Treatment*, *28*(1), 19–26. <https://doi.org/10.1016/j.jsat.2004.11.001>

Pinto dos Santos, D., Giese, D., Brodehl, S., Chon, S. H., Staab, W., Kleinert, R., Maintz, D., & Baeßler, B. (2019). Medical students’ attitude towards artificial intelligence: A multicentre survey. *European Radiology*, *29*(4), 1640–1646. <https://doi.org/10.1007/s00330-018-5601-1>

Roth, C. B., Papassotiropoulos, A., Brühl, A. B., Lang, U. E., & Huber, C. G. (2021). Psychiatry in the Digital Age: A Blessing or a Curse? *International Journal of Environmental Research and Public Health*, *18*(16), 8302. <https://doi.org/10.3390/ijerph18168302>

Sendak, M. P., D’Arcy, J., Kashyap, S., Gao, M., Nichols, M., Corey, K., Ratliff, W., & Balu, S. (2020). A Path for Translation of Machine Learning Products into Healthcare Delivery. *EMJ Innovations*. <https://doi.org/10.33590/emjinnov/19-00172>

Sharma, A., Lin, I. W., Miner, A. S., Atkins, D. C., & Althoff, T. (2021). Towards Facilitating Empathic Conversations in Online Mental Health Support: A Reinforcement Learning Approach. *Proceedings of the Web Conference 2021*, 194–205. <https://doi.org/10.1145/3442381.3450097>

Shatte, A. B. R., Hutchinson, D. M., & Teague, S. J. (2019). Machine learning in mental health: A scoping review of methods and applications. *Psychological Medicine*, *49*(9), 1426–1448. <https://doi.org/10.1017/S0033291719000151>

Tanana, M. J., Soma, C. S., Srikumar, V., Atkins, D. C., & Imel, Z. E. (2019). Development and Evaluation of ClientBot: Patient-Like Conversational Agent to Train Basic Counseling Skills. *Journal of Medical Internet Research*, *21*(7), e12529. <https://doi.org/10.2196/12529>

Zheng, F., Wang, L., Zeng, Z., & Wu, S. (2022). International Technologies on Prevention and Treatment of Neurological and Psychiatric Diseases: Bibliometric Analysis of Patents. *JMIR Mental Health*, *9*(2), e25238. <https://doi.org/10.2196/25238>