International Journal of Multidisciplinary Research and Growth Evaluation.

The Role of Artificial Intelligence in Delivering Mental Health Support to Adults: Opportunities and Ethical Concerns

Tehmina Jamil [1*], Sayed Muhammad Shabab [2], Muhammad Mubashir Hassan [3], Nazia Parveen [4]
[1] Department of Anthropology, Institute of Social and Cultural Studies (ISCS), Pakistan
[1] Bahauddin Zakariya University Multan, Pakistan (BZU), Pakistan
[1] Former Visiting Lecturer, Muhammad Nawaz Shareef University of Agriculture, Multan, Pakistan
[2-3] Department of Sociology, School of Humanities and Social Sciences, University of Science and Technology Beijing (USTB), China
[4] Lecturer, National University of Modern languages Multan Campus Pakistan, Pakistan

* Corresponding Author: **Tehmina Jamil**

Article Info

ISSN (online): 2582-7138
Volume: 06
Issue: 04
July - August 2025
Received: 08-05-2025
Accepted: 10-06-2025
Published: 03-07-2025
Page No: 705-716

Abstract
The accelerated uptake of Artificial Intelligence (AI) in healthcare has revolutionized mental health service provision by providing affordable and scalable interventions for diverse populations. In adults, AI-based mental health interventions, including chatbots, diagnostic apps, and mood tracking systems, are becoming common tools for emotional support. Although such interventions have the potential for self-management and convenience, they also pose important ethical, emotional, and practical issues. The research examines adult users' experiences and attitudes towards AI-driven mental health technologies regarding both advantages and ethical constraints. The research aimed to explore how adults interact with AI in mental health scenarios, what they consider to be useful, and what issues arise when using it. The overall research questions were concerned with perceived efficacy, affective limitations, and ethical issues, including data protection, trust, and the necessity for human-AI balance. A qualitative study design was utilized, involving semi-structured interviews with 18 adult participants between 21 and 52 years old with experience in AI mental health tool usage. Data were analyzed thematically, and five major themes emerged: perceived usefulness and accessibility, emotional limitations of AI, ethical issues, need for human-AI balance, and suggestions for improvement. Convenience of use, emotional check ins, insufficient empathy, data privacy, and cultural fit were sub-themes. Findings show that although AI tools provide instant, stigma-free emotional support, participants complained of superficial answers, insensitivity, and ambiguous data practices. Participants decidedly preferred a hybrid approach where AI assists, but does not substitute, professional human care.In summary, AI has great promise for adult mental health care but needs to be ethically developed, open, and culturally sensitive in order to be really beneficial. Human control and crisis-response incorporation must be the focus of future development.

Keywords: Artificial Intelligence, Mental Health, Adults, Chatbots, Qualitative Study, Ethics, Digital Therapy, User Experience

Introduction

Artificial Intelligence (AI) has evolved from an academic abstract concept into a central force shaping change in a variety of industries, with healthcare very much at the lead. In healthcare, mental health is rapidly becoming one of the most dynamic and troublesome areas for AI implementation. Mental health services have historically faced perennial challenges most notably workforce shortages, inconsistent access, stigma, and expense.

Future care: COVID-19 laid bare the negative impacts of long-term systemic barriers in the mental health care services and resulted in previously unseen rates of anxiety, depression, and loneliness, which can only be addressed with future care. Face-to-face therapy and formal clinical interventions which traditionally focus on conventional care models failed to respond to the need; the short-coming highlighted the essential components of the lack of responsiveness and reach (Lee *et al*., 2021) [11]. The crisis that has followed has accelerated the use and implementation of AI-based mental health applications. Thes interventional solutions combine diagnostic platforms, predictive model systems, conversational AI, and AI-assisted monitoring to reduce clinician labour intensive work, remove use barriers, and provide customized scalable care provision in a 24/7 basis. Such a blistering innovation of AI in the sphere of mental health can be analyzed under three main categories including diagnostic software, treatment chatbots, and monitoring systems. Diagnostic AI mitigates mental health issues by developing voice patterns, text entry, and biometric sensor data based on machine learning and natural language processing (NLP) to determine manifestations of mental illnesses at early stages. In a study by Pandi Perumal *et al*. (2023) [15], the analysis of speech and sentiment of monitoring could be used to predict depressive states in the adult population even before a depressed state could be diagnosed. All the above developments demonstrate the importance of automation and data-based accuracy in the field of mental health care, providing new potential channels of prophylactic intervention and extending the capabilities of current services. Despite the relevant ethical and regulatory concerns that these innovations bring, they promise to remain a viable alternative to systemic limitation and demand-side pressures that still influence mental health practice.

Predictive modeling has become an early approach towards the detection of depressive episodes in the adults thus paving the way before the traditional diagnostic processes. A study by Cho *et al*. (2020) [5] of a similar nature evaluated the machine-learning models that can identify early signs of relapse in people recovering due to post-traumatic stress disorder (PTSD). These methods of diagnosis prove promise of providing early detection and in addition early detection in a more accurate manner which is a precursor to effective early intervention. At the current moment, AI monitored systems employ passive data sets to provide real-time metrics of mental health. AI-powered wearable technology is also able to record sleep activity, heart-rate variability, phone usage trends, and geographic location, thus potentially becoming an indicator of emotional dysregulation or crisis before it comes. According to a large-scale pilot study by Smith *et al*., the system was able to recognize 87 percent of self-reported depressive episodes 48 to 72 hours prior to the onset of symptoms, thus allowing this system to remind a person to take care of themselves earlier and seek assistance earlier. These findings highlight how AI can be used to not only provide treatment but be involved in the process of monitoring. However, a variety of ethical issues still prevail, such as the problem of consent, data control, and privacy protection. Another urgent issue is algorithmic bias. In 2024, Reuters described a depression-detection tool based on AI that was very popular but systematically under detected depression symptoms in Black adults, by three times as much as in White adults, because its training dataset was overwhelmingly White. Gender discrimination has also been reported; the NLP-based sentiment analysis approaches trained on male-authored text fails to appropriately recognize female emotional communications leading to lower sensitivity and increased false-negative on female sample (Gonzalez *et al*., 2023) [9].

These issues underscore how lack of diversity in training data can reinforce systemic inequities in mental health care—exactly the opposite of AI's intended mission. However, AI can also be amplified for equity when designed carefully. Woebot's open-access platform for students at colleges, used in multiple low-income countries, decreased reported anxiety by 18 percent in a randomized pilot trial (Behrman *et al*., 2024) [2]. In Mexico, a Spanish speaking AI chatbot resulted in a 25 percent improvement in uptake of self-help by underserved adult populations, indicating scalable, culture-relevant impact (Ramirez *et al*., 2024) [16]. Such studies highlight that if diverse groups of users are involved in AI development from the beginning, they can develop inclusive solutions that overcome cultural and resource gaps. This research has two implications. Functionally, it can educate developers, regulators, and mental health professionals regarding adult concerns and priorities—enabling the development of AI tools that are safe, accessible, fair, and clinically useful. Ethically, this research contributes to the academic discourse on AI in mental health by closing the gap between empirical user viewpoints and theoretical critique on digital ethics and algorithmic justice. It also helps to build increasing demands by WHO and leading health organizations for equitable, open, and culturally sensitive AI deployment (WHO, 2023).

Problem statement

Mental illness conditions are among the most urgent public health issues of the 21st century, with approximately one billion individuals affected worldwide (WHO, 2023). Adults, in fact, experience escalating psychological distress because of work pressure, financial insecurity, social isolation, and post-pandemic trauma. Access to conventional therapy remains poor despite escalating demand for mental health care, particularly in low-resource or stigmatizing environments. Shortages of health professionals, high prices, and social stigma still deter most adults from receiving or accessing care in a timely manner. As a response, Artificial Intelligence (AI) has been proposed as a viable solution, providing digital mental health solutions in the form of chatbots, diagnosis algorithms, and real-time mood trackers. These technologies are now being applied to augment or even displace conventional interventions through scalable, on-demand provisioning. Though initial research indicates that AI-powered mental health interventions are effective for lessening anxiety, depression, and stress, their mass deployment evokes vital ethical and practical issues. The modern context demonstrates that there are five interdependent problems that are defined as material roadblocks to the ethical use of AI in adult mental care: data privacy, algorithmic bias, explainability, a lack of regulatory structure, and disparity of access to electronic resources. In spite of how fast the field of technology is developing and advancing, the empirical research in terms of the user trust and clinical safety, as well as the procedural fairness are much slower and fall behind the real implementation of the same. The lack of rigorous investigations brings the possibility of AI increasing the disparities that exist today and generating negative consequences. In turn, the current study

aims at clarifying the potential advantages of supporting with the help of AI tools, as well as the ethical risks and limitations associated with the support implementation in adult mental health. With user experience reports and the ethical considerations of such technologies, the project hopes to create evidence-based regulation propositions on the safe design, delivery, and governance of AI-assisted mental health services of adults.

Research objectives

- To investigate the ways in which Artificial Intelligence (AI) is being applied to provide mental health care for adults.
- To gauge adult users' views on the effectiveness, trustworthiness, and use of AI-driven mental health tools.
- To determine major ethical issues—i.e., data privacy, algorithmic bias, and safety—related to AI in mental healthcare.

Literature Review

Artificial Intelligence (AI) has become a game changer in the modern medical field in remarkable speed after years of being a merely technical fringe concept. Over the years the healthcare systems have been facing some of the same problems which include shortage of staff, high prices, poor accessibility, and long queuing. The introduction of AI has been specifically erupted to address such limitations especially in cases that entail large volumes of data to analyze in a systematic manner as well as in cases where a repetitive decision-making process needs to occur. In recent years, mental health has become the area of active digital healthcare implementation, and AI-based tools are specifically subject to the trend. Media discussion Jin *et al*. (2023) [10] argue that AI solutions are revolutionizing the field of healthcare to boost diagnostic abilities, simplify treatment algorithms, and aid clinical decision-making. Lu *et al*. (2023) would agree, as they state that AI does not only amplify efficiency rates in health care delivery, thus making it more efficient, but also makes it more personalized, which is essential for tricky, patient-centered conditions like mental health disorders. Despite these developments, there are still ethical issues in application over and over again such as when it comes to the exploration of AI in otherwise sensitive fields, particular in psychiatry and psychology. The ethical concerns regarding the use of AI in medicine are stressed out by Denecke *et al*. (2021) [6], as they touch upon the need to establish a transparent regulation, proper control over the data, and the desire of making the AI patient-catching.

More broadly, artificial intelligence (AI) systems have already been designed in the context of the modern mental-health practice and can fulfill a number of purposes, including chatbots, symptom-monitoring platforms, diagnostic adjuncts, and predictive algorithms. Chatbots specially, such as Woebot, Wysa, Replika, and Tess, are aimed not only to simulate the process of conversation but to provide interventions based on the principles of cognitive-behavioral therapy (CBT). Woebot is also shown to be effective in relieving depression and anxiety symptoms in undergraduates who took part in a randomized controlled trial carried out by Fitzpatrick *et al*. (2017) [8]. Further research has supported these results thus making it easier to be adopted. In their analysis, Casu *et al*. (2024) [4] assessed the use of mental-hygienic chatbots with AI support and decided that, carefully designed, they can be considered as an adequate alternative to address the issue of mental illnesses those systems can provide a good and scalable solution to combat the common psychological problems in a cost-effective manner. AlMakinah *et al*. (2024) [1] also stated that AI chatbots, with a strong ethical control and constant further improvements, could help users engage with the therapy and promote affective well-being, at least in case of adults who demonstrate reluctance to conventional face-to-face treatment. In addition to conversational support, AI is also used in diagnostic and symptoms-sequential-monitoring. Recently, a study by Pandi-Perumal *et al*. (2023) [15] focused on the ability to recognize early symptoms of depression based on the speech patterns describing an occurrence of natural-language processing (NLP) which proved that using artificial intelligence (AI) technology it is possible to predict upcoming depressive episodes before they became officially recognized as such. In different research. study, Cho *et al*. (2020) [5] used the machine-learning technologies to predict post-traumatic stress disorder (PTSD) relapse among military veterans with high accuracy when continuous data streams were included. Recently, a wearable AI platform to predict depressive mood episodes within the next 72 hours has been introduced by Smith *et al*. (2023) [18] and uses biometric data, like heart-rate variability and sleep variables. The combination of the studies shows how solid the possibility of the development of early detection and preventive measures in adult mental health care can be with the help of AI. The current research on AI-based interventions for adult mental health creates both a picture of possible opportunities and the corollaries of limitations. Yoo *et al*. (2025) [23] analyzed user interactions with a trial GPT-4-powered chatbot derivation agent: The Zenny chatbot. The respondents mentioned the most obvious benefits of fast feedback, anonymity, and personalized responses. On the other hand, the researchers recorded instances of tonal and content-based miscues, especially when the delicate emotional disclosures are being made. A document published in 2024 on a systematic review witnessed minimal to moderate effect sizes in symptom relief of both anxiety and depression after eight weeks of exposure to the use of AI chatbots, with a total average of 0.32 depression symptoms reduction through the use of AI chatbots, averaging 8 weeks of ongoing use. However, the research revealed a key drawback: it was challenging to sustain benefits in the absence of human follow-up or booster sessions. Doshi *et al*. (2024) [7], in their research on "TheraGen," an AI-powered chatbot for working adults, had a 94% rate of satisfaction from adult users but pointed out that most participants were already experienced with mental health support, thereby creating a bias in the user base and expectations. At a larger level, Wu *et al*. (2024) carried out a comparative analysis of what clinicians and patients thought about AI-supported mental health treatments. Although mental health professionals liked how AI tools could lighten their administrative workload and provide initial assistance to clients, several were uneasy about depending too much on automation, particularly during crises. On the side of the users, 44% of respondents reported privacy as their number one issue, followed by a lack of transparency regarding the functioning of the AI. The results agreed with those of JMIR Mental Health (2024), which found that although AI tools are becoming increasingly popular, many adult users remain unconvinced about data security and emotional authenticity.

Aside from bias and privacy, existing literature also questions how AI would affect human relationships. Although AI applications are available to provide instant assistance and minimize help-seeking stigma, they can potentially substitute for basic human interactions. The New York Times (2025) reported instances where people became emotionally dependent on AI chatbots at the cost of relationships in real life. This has made some researchers, such as Behrman *et al*. (2024) [2], recommend AI as an adjunct—and not replacement—tool for mental health services, stressing the need to ensure AI is supplemented by human-oriented services and not substituted.

Methodology

Research Design

The proposed research study selected a qualitative research design in exploring the mental health support technologies based on AI and the perceptions and experiences of adults regarding the ethical dilemmas in the implementation of such technologies. Qualitative designs are also an excellent choice to develop specific information about the experiences, perceptions, concerns, and suggestions of AI applications in mental healthcare. The approach facilitates a comprehensive study of meaning, understanding, and lived experience because it splits the focus on hypothesis testing in favour of an open-ended investigation of the phenomenon.

Population and Sample

The population of interest is adults aged 18 years and older who have utilized AI-based mental health resources like therapy bots (e.g., Woebot, Wysa), diagnostic software, or mood-tracking programs. The sample will seek to be representative of diversity in terms of age, gender, education, and computer literacy in order to provide rich and varied insights. A sample of 15 to 20 participants should be enough to bring about data saturation, the stage at which no new themes are arising.

Sampling Technique

The research will employ purposive sampling to identify participants with the required experience in the use of AI for mental health. Snowball sampling will also be utilized to increase participant coverage, where existing participants can refer others with the same experiences. This method is especially effective in reaching individuals who have used specialized or privately accessed AI software.

Data Collection Tools

Data is gathered through semi-structured, in-depth interviews. An interview was scheduled, prepared with open-ended questions across the following areas:

Experiences of using AI-based mental health tools, Perceived usefulness, limitations, and emotional impact Ethical issues (privacy, trust, bias, emotional safety) Suggestions for improvement and balance between humans and AI Interviews were held over Zoom and telephone and lasted 30-45 minutes. Interviews were audio-recorded (with permission) and transcribed verbatim for analysis.

Data Analysis

Thematic analysis will be applied to interpret the data using the six-step approach by Braun and Clarke (2006):

- Familiarization with data
- Initial codes generation
- Theme searching
- Review of themes
- Defining and naming themes
- Report production Manual or qualitative analysis software like NVivo will be employed to code. Patterns and themes will be identified to gain an understanding of user experiences and ethical considerations of AI in mental health care.

Ethical Considerations

Ethical clearance shall be sought from the appropriate institutional review board or ethics committee. The informed consent form will be given to each participant to detail the purpose, procedures, risks, benefits, and rights (e.g., the right to withdraw at any time) of the study. Anonymity and confidentiality will be ensured by labeling participants with codes or pseudonyms. All information will be kept safe and will be used solely for scholarly research. Matters of a sensitive nature like mental health and privacy will be handled with compassion and professionalism.

Results

Data were gathered from 18 adult participants between the ages of 21 and 52 who had used AI powered mental health technologies like Woebot, Wysa, Youper, and AI emotion-tracking apps. Thematic analysis was employed to establish significant patterns in the participants' experiences and ethical considerations. The following four overarching themes were found:

Thematic Analysis

Theme	Sub-Themes	Sample Codes	Frequency (n=18)
1. Perceived Usefulness & Accessibility	• Convenience of use • Emotional check-ins	"I like that it's available 24/7" "Quick help when I feel anxious"	14
2. Emotional Limitations of AI	• Lack of empathy • Shallow responses	"It felt robotic" "It doesn't understand real pain"	11
3. Ethical Concerns	• Data privacy • Trust & transparency	"I don't know where my data goes" "Who owns this info?"	16
4. Desire for Human-AI Balance	• Need for human support • Therapist referral	"Sometimes I want to talk to a real person" "AI should link me to a counselor"	9
5. Recommendations for Improvement	• Emergency/crisis support • Cultural fit	"Add emergency contact feature" "Use simpler, relatable language"	10

Perceived Usefulness & Accessibility

Participants largely acknowledged the usefulness and accessibility of AI-based mental health tools.

Most valued having assistance available at any time, particularly in times of emotional distress or anxiety at midnight, when traditional services were not accessible. The ease of interacting with chatbots or apps at home was given as a major benefit, especially among those who did not want

to approach face-to-face therapy because of stigma or expense. According to initial results, reviewing the digital interventions that are the object of the study, they were all positioned around the terms of being non-judgmental and user-friendly, providing real-time feedback, pre-constructed exercises (cognitive reframing), and on-session emotional tracking (prompting appraisal). The participants noticed that such automated systems enabled them to take notice of how their mood was changing and increased their self-awareness levels in the long term. Despite the perceived advantages of these interventions, the majority of the users agreed that interventions through AI become a temporal accessory and not a full-fledged alternative to long-term psychotherapy. Taken together, the data indicate that such technologies are beneficial in overcoming the barriers to early care-seeking as they provide the facility of convenient and non-invasive access to support.

Sub-theme: Convenience of use

There is empirical evidence, which proves that AI-driven mental-health interventions are commonly considered as convenient. Unlike any traditional therapy, which may be defined by long waitlists, rigid appointment times, and external travel necessities, digital mediums, and AI-based assistant provide an immediate option that appears every hour of the day. According to participants, the usage of such tools that have internal origins is particularly beneficial in case there is an acute episode of distress, where face-to-face communication can appear to be impossible. Because such populations (i.e., people of working age and students who are busy) face time constraints, the flexibility would be considered a pertinent enabler of mental-health care.

Many users reported using the apps late at night when anxiety or feelings of loneliness reached their highest point. Having support available immediately, without judgment or pressure, made these tools appealing for people new to mental health care. For most, AI tools became a personal, non-intrusive stepping stone to emotional regulation, allowing them to regulate stress in their own time and space. This sub-theme reveals how convenience and accessibility are powerful drivers for adopting AI for mental health. "I loved that I could speak to the app at any time of the day or night. When I got anxious at 3 AM, it was there. I didn't have to wait to see a therapist." (p 7) "It's much more convenient than visiting a clinic. I can just open the app on my phone during my lunch break and do a quick check-in without anyone knowing [12]."

Sub-theme: Emotional check-ins

Another treasured aspect of AI mental health applications was that they enabled daily emotional check-ins. Numerous participants indicated that constant reminders to check in with themselves emotionally helped them cultivate more emotional self-knowledge. The check-ins usually came in the form of mood monitoring, journaling capabilities, or just straightforward questions such as "How are you feeling today? " that prompted users to slow down and reflect on their mental state. Others presented visual summaries or mood graphs, allowing users to identify patterns over time— e.g., correlations between sleep, stress, or social activity and mood changes. The process itself was therapeutic, whether or not the AI generated complex responses. For those hesitant to expose feelings to others, the quiet and orderly format of check-ins offered a safe refuge. Overall, this sub-theme identifies that mood check-ins facilitate self-awareness and mood management, allowing users to feel heard and understood even without human interaction. "The mood check-in every day made me stop and actually think about how I was feeling. I never used to do that before—it helped me notice my emotions more." (p 3) "Watching my mood charted throughout the week was useful because it made me see patterns. I noticed my anxiety was bad on Mondays and I began to better manage my schedule." (p 9)

Theme-2. Emotional Limitations of AI

The participants of this study did not deny the greater utilitarian benefits of AI technologies, but they always emphasized their lack of emotions. The respondents characterized interactions made using AI as mechanical, impersonal or being scripted. The answers that chatbots could offer were formal and pre-programmed and their responses rarely were subtle enough to indicate emotional intelligence needed in sensitive or emotionally challenging situations. The participants mentioned cases where the context was incorrectly read, the emotions were interpreted wrong, or scripted answers caused wrong interventions, thereby causing the user to become disinterested or to feel anger. Lack of empathy was especially noticeable in times of grief, trauma as well as isolation where the users needed validation as well as comfort. Even though some of the respondents acknowledged the implied limits of non-human systems, a few of them were disappointed that the system was not able to pick up emotional messages or raise important issues.

Sub-theme: Lack of empathy

One of the issues that participants repeatedly raised is lack of natural empathy of AI-based mental health technologies. Though users very much enjoyed the speed and consistency of responses provided by applications or chatbots, the responses tended to be labeled as emotionless or quasi-robotic. It was stressed by respondents that AI systems do not have the capabilities to detect nuances of emotions and non-verbal communication or intonation developing a feeling of emotional distance, especially at times when users needed them the most, when they were feeling sad, grievous, or anxious, expecting their systems to react positively and appear welcoming. According to one of the participants, he or she just needed to feel heard, but the app resembled a machine repeating some catchphrases. The observation highlights the fact that no matter the reproduction of empathetic language, AI lacks emotional intelligence and personal human touch that can efficiently provide psychological support, making such technologies insufficient when dealing with people who need more than an algorithmically reproduced response.

It provided information but the language sounded robotic-like: as if the system was dictating known expressions at a time in a coded script. This meant that, I did not feel that it was all in a state of understanding of my experiential level." (p 5)

In one of the most critical conversations, I revealed a very intimate issue, and the response I obtained was, You are doing well! ". It was cold sentiment instead of being warm and it brought about the feeling of depersonalization. There was no reassurance which I needed; instead, a hackneyed suggestion (p 11)

Sub-theme: Shallow responses

The other frequent issue was with the generic or insubstantial

nature of AI responses. Members explained how the responses started to sound repetitive, too scripted, or incongruent with their particular emotional scenario after prolonged use. Although AI software tended to draw from CBT protocols or inspirational prompts, users experienced that responses did not have the depth or flexibility to work through intricate or layered emotional concerns. Several participants noted that when they expressed difficult or confusing emotions, the chatbot either misunderstood the input or gave vague encouragement like “You’re doing great! ”—which felt disconnected from their experience. This superficial engagement left some users feeling unheard or dismissed. The secondary theme of superficial responses indicates that, while AI can process structured interaction, it usually does not personalize or enrich the interaction to enable it to effectively deal with more serious or developing emotional needs.

"Sometimes I'd write a big message detailing how I was feeling, and it would just say something like, 'Keep going, you're strong!'—it didn't seem relevant to what I wrote." (p 6)

"The responses began to feel repetitive after a while. Regardless of what I wrote, the responses didn't delve deeper or ask anything substantial." (p 13)

Ethical Concerns

Ethical issues came out as the prevailing theme, especially regarding data privacy, transparency, and user consent. They were uneasy about divulging sensitive emotional data on AI platforms without fully comprehending how their data was processed, stored, or possibly shared. Most admitted to skipping or disengaging from privacy policies because they were too complicated, which resulted in uncertainty and mistrust. Some participants asked who really "owns" the data gathered and whether business firms could use their mental health profiles. Some participants also expressed fears about algorithmic bias—whether AI software could misunderstand emotional cues in various languages, genders, or cultures. According to the results of empirical research, it has been established that trust towards using AI-based applications is more common when such software has been developed by either healthcare organisations or academic institutions rather than independent technology enterprises. The result highlights the significance of data protection systems and ethical ethos in the establishment of trust designs of users in online mental-health tools, especially in the vulnerable adult social groups.

Sub-theme: Data privacy

The privacy of data became the most prominent theme among the participants. An impressive percentage of the respondents stated that they could not feel comfortable transmitting extremely intimate emotional information to AI tools with no idea where or how this information would be stored. The main issue that participants had was that third parties, especially technology companies and not healthcare providers would have the chance to access the mental health disclosures. Admitting that the privacy policies are rather complex, most of the participants revealed that they have not read all of them fully and remained at risk. Moreover, there was a fear that data related to their mental health would be used to show directed advertisement or sold without authorisation. This trail of concern explains why greater data protection is necessary, relevant consent procedures, and tough regulation. Without the trust in the data management system, the vast majority of the users are unlikely to disclose themselves to AI-based mental health services completely, limiting the potency and scope of such services.

"I was always kind of anxious about what happens with the things I post. It's my mental health— we should have an idea where that data goes." (p 4)

"I didn't read all of the privacy policy. It was too long and technical, so I just clicked agree—but to be honest, that does concern me." (p 8)

Sub-theme

Trust & transparency Essentially tied up with concerns about data, most participants highlighted the need for transparency both in how AI tools operate and who is in charge of them. They inquired about what information was being gathered, how the decisions were taken by the AI, and if a trained team regulated the operation of the platform. Insufficient transparency regarding the sources of knowledge of the AI, its responding mechanisms, and ownership generated doubt and suspicion. Certain users questioned if the chatbot was providing therapy advice or merely predetermined responses geared for user retention. Trust was overall greater in apps supported by universities or mental health organizations, but skepticism was greater for commercial apps that lacked obvious affiliations. This sub-theme demonstrates that transparency of communications, visible credentials, and unambiguous ethical standards are key to establishing user trust. Without them, even well-crafted tools can become perceived as untrustworthy or unsafe, particularly with something as vulnerable as mental health.

I didn't know who was behind the app—whether it was a therapist, a technology firm, or just software. That made me struggle to totally trust it." (p 10)

"If the app could have described to me how it works and who gets to see my information, I would have felt at ease. It was just a black box." (p 2)

Desire for Human-AI Balance

The results showed a universal desire to choose a hybrid model that has reconciled the cost-effectiveness of artificial intelligence (AI) with the care and experience of human therapists. Respondents admitted that AI platforms are able to target initial support or regular surveillance but not a full-fledged mental health care. They also insisted that AI must serve as alternate tool that suggests coping measures and data analysis, as well as a tool with the help of which users receive directives to introduce them to official assistance in the instance that they develop significant symptoms. In the case of people with moderate to severe mental illness, some participants recommended incorporating therapist instatements in the apps or live chats with licensed practitioners. All in all, the data helped to reaffirm the understanding that users view AI as a part of a larger care ecosystem and not a separate intervention. Finally, a combination of AI and human observation proved to be the safest, effective, and ethically reasonable approach toward guaranteeing sustainable mental health care.

Sub-theme: Need for human support

While AI applications were viewed as beneficial for instant assistance, most participants strongly desired human contact in their mental health process. They perceived that, even with the efficiency and convenience of AI, it could not match the

depth, sympathy, and multidimensional grasp that only a trained human specialist can offer. For individuals with complicated emotional issues—like trauma, bereavement, or relationship dissolution—AI answers constantly seemed substandard. Some respondents mentioned that beyond a certain point, they needed to speak to an actual human being, one who could read their emotional response, pose thoughtful follow-up questions, and respond with empathy. Without this human touch, some found themselves feeling emotionally isolated. This sub-theme highlights the fact that although AI can be a useful adjunct, it cannot be a substitute for the therapeutic relationship or emotional validation that results from communicating with an educated counselor or therapist.
"The AI was helpful initially, but soon I just needed to speak with a human who could get me, as opposed to provide generic responses." (p 1)
"It was fine for surface-level things, but when I was feeling low, I knew I needed someone human to speak to someone able to listen." (p 14)

Sub-theme: Therapist referral

Participants often mentioned that AI mental health apps should have a clear and accurate referral system to licensed therapists. While people valued early assistance from AI, they indicated the need to have a bridge to professional care, particularly when the symptoms intensified or persisted. Some reported that the AI persisted to offer superficial advice when they flagged extreme distress, which they found irritating and dangerous. Some respondents indicated that apps might contain functionalities such as automatic reminders, in-app referrals, or a referral list of licensed therapists by geographical location. This would not only make AI an independent self-help tool, but it would also act as a gateway to formalized care. The second-order theme highlights that effective and responsible mental health AI tools should be aware of their limitations and facilitate a seamless transition from automated to human-provided assistance, ensuring users receive the treatment they genuinely need.

"There should be an option where the app says, 'It might help to talk to a real therapist,' and then connects you to someone. I felt stuck when I needed more help." (p 35)
"If it had an option to schedule a session or at least recommend a professional in the vicinity, that would make it complete and more responsible." (p 7)

Suggestions for Enhancement

Participants gave careful consideration to the ways AI-based mental health devices could be enhanced, based on functionality, safety, and accessibility. An across-the-board suggestion was the incorporation of emergency functions, such as referring users to crisis hotlines or triggering alerts when a person is exhibiting suicidal thoughts. Others requested more culturally sensitive material and clearer, jargon-free text to make the devices more usable across levels of education and backgrounds. A few participants called for more individualization in answers, such as AI that would "learn" from their mood patterns or writing style over time. Significantly, users also emphasized the need for improved communication regarding data policy, specifically regarding what data is being collected, who has access to it, and for how long. These are some of the observations. the willingness of the users to use AI mental health tools in case they are safe, transparent, and customized to unique needs. The sub theme underlines the importance of being crisis ready, including terminology, and ethical design in terms of the development of future technology.

Sub-theme: Emergency/crisis support

The major suggestion the participants came up with was to incorporate the emergency or functions of crisis management are implemented into AI mental health technologies. A number of users were wary of using available AI to identify or manage extreme mental illnesses emergencies, such as suicidal intentions, panic attacks or emotional breakdown. The users were concerned about someone who is emotionally distressed may be given generic or late messages by a chatbot when they needed an urgent human assistance. The others were suggesting the inclusion of intrinsic safety notifications, such as sharing the numbers of crisis helplines automatically, notifying emergency contacts, etc. reporting the case to the mental health professional when obvious high-risk language is detected. This sub-theme makes it clear that there is an ethical duty of AI developers to instill in their technology safety characteristics that protect users during emergencies, which meant that the technology can be more than just a useful technology are also responsible and accountable, where they are dealing with human life. The recipients emphasized that AI apps in mental health would be helpful in any location, provided that they were useful ones. to become culturally tolerant and fit. Some users claimed that the answers provided by the chatbot were excessive. Western-oriented or in general, contained un-known vocabulary to which they could not relate to in their life realities. For instance, idioms, culturally particular allusions, or therapeutic styles based in a single culture occasionally caused confusion or estrangement for users with different backgrounds. Some players from non-English-speaking areas also mentioned issues with language barriers or translation quality. Other proposed tools might provide language customization, culturally tailored advice, or regional mental health resources. This sub-theme underscores that for AI tools to be fully supportive of mental well-being, they need to be respectful of cultural diversity and ensure their design, language, and content are inclusive, relatable, and appropriate for users from different social and cultural backgrounds.
"Some of the words it used didn't apply in my situation—it sounded Western, like it wasn't created for individuals like me." (p 15)
"It would be improved if the app picked up on our culture or included more familiar illustrations. At times it advised us on things which simply did not exist in our lifestyle." (p 9)

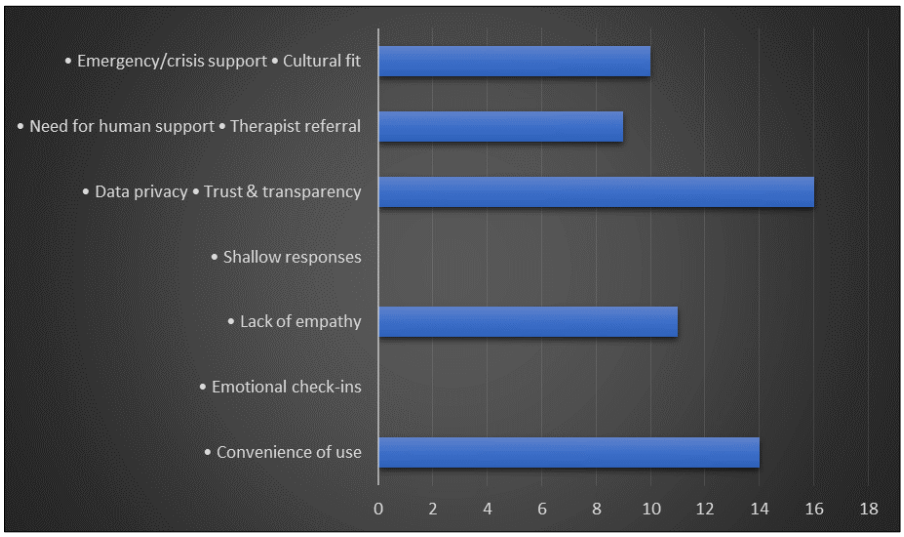

Fig 1: Frequency of User Feedback Themes on AI in Mental Health Context

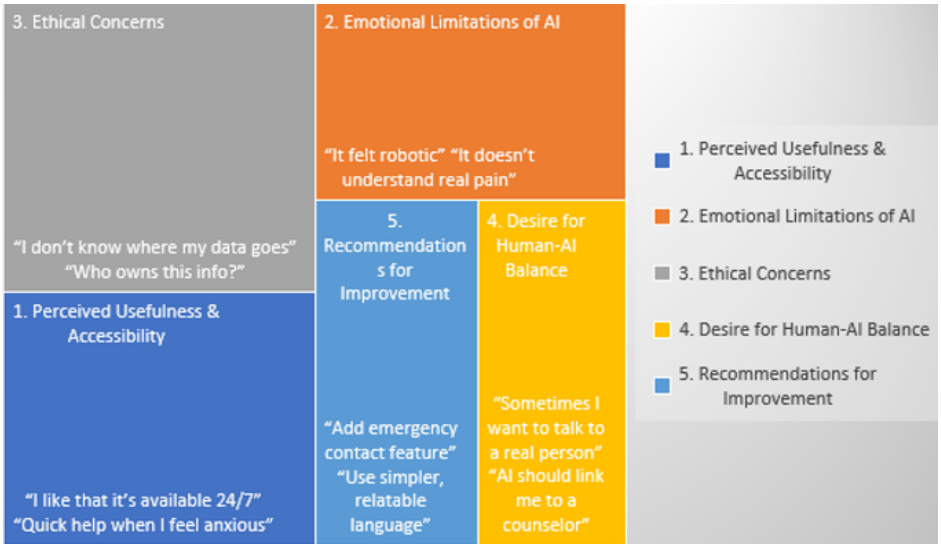

Fig 2: Thematic Categorization of User Feedback on AI Interaction

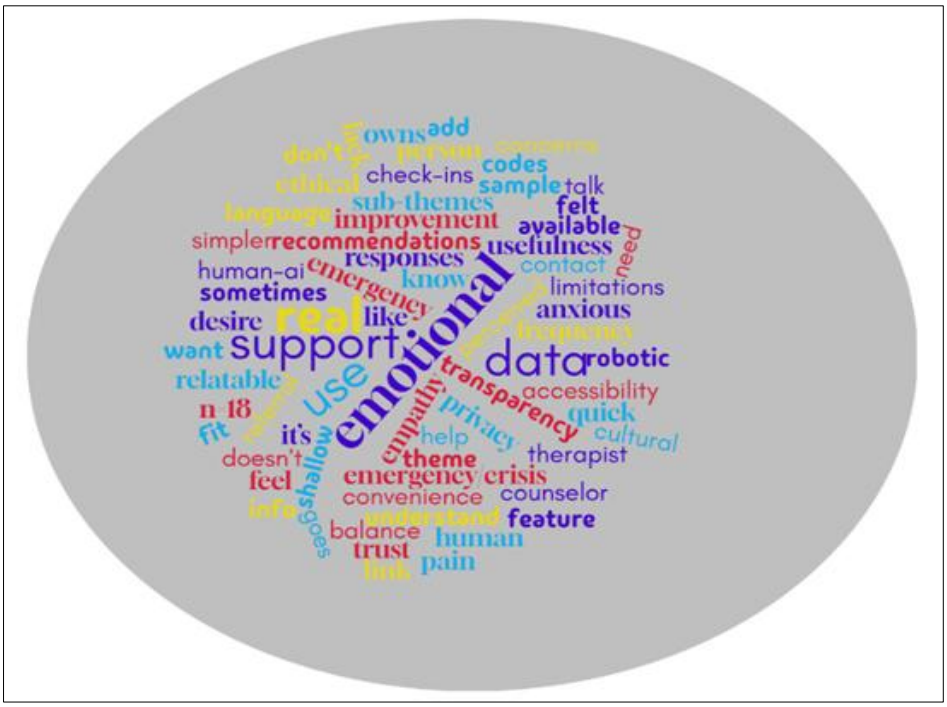

Fig 3: Word cloud depicting key themes and sub-themes

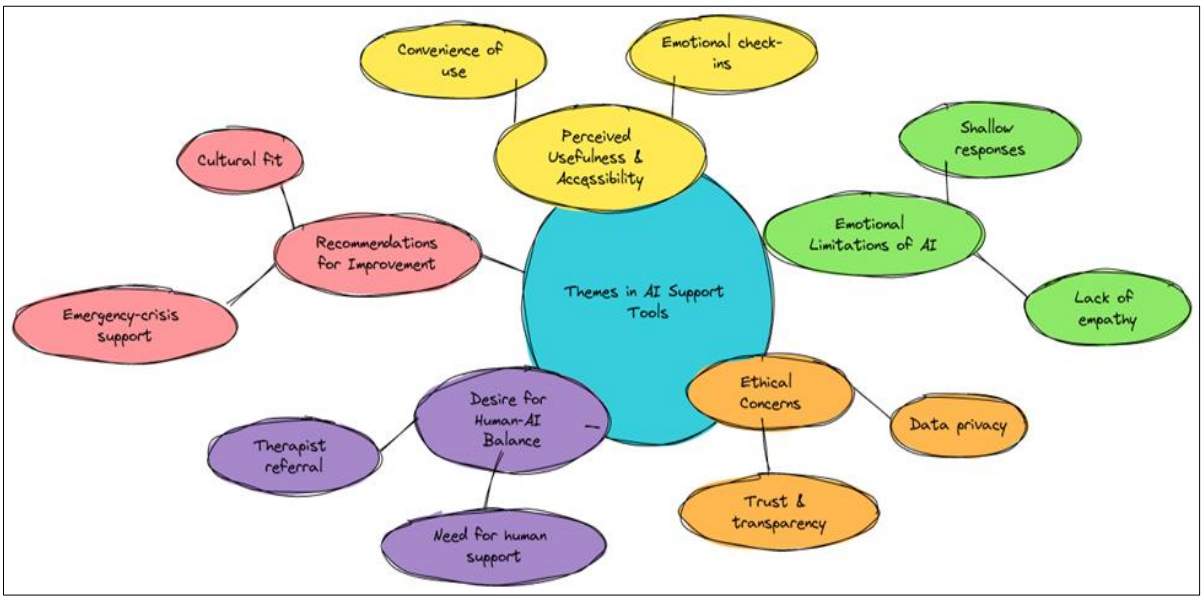

Fig 4: Themes in AI Support Tools

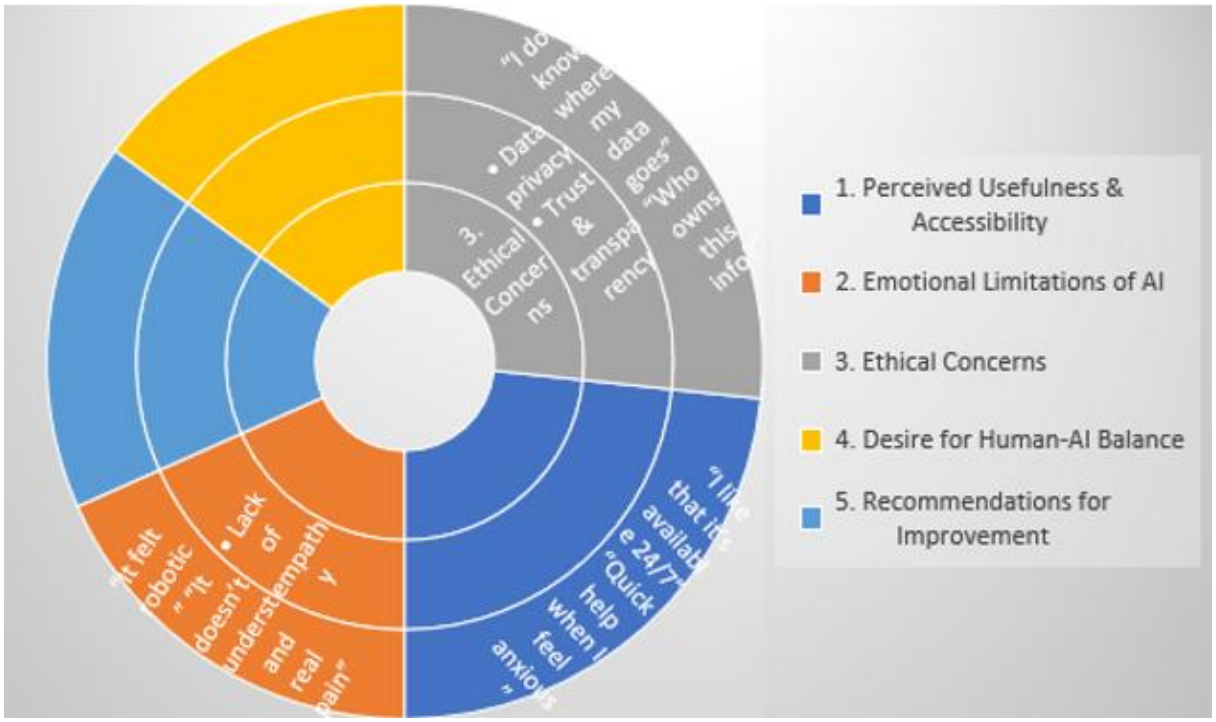

Fig 5: Thematic Analysis of User Feedback on AI

Discussion Submission

This research sought to identify adult users' experiences and ethical issues regarding the use of AI based mental health applications. Five main themes were derived from thematic analysis of qualitative data from 18 adult users: perceived usefulness and accessibility, emotional limitations of AI, ethical concerns, desire for human-AI balance, and suggestions for improvement. These findings offer important insights into both the potential and limitations of AI in supporting adult mental health, especially in an era where digital health interventions are rapidly expanding but still lack regulatory oversight and emotional depth.

The first theme was Perceived Usefulness and Accessibility, which demonstrates the position of participants as they valued AI. the availability, easy access, and speed of emotional provision of mental health tools, which can be taken out, one per day and put back together at any time of the day or night support. These findings are aligned with past studies that emphasize on the convenience and when needed nature of the service aspects of AI applications. AlMakinah *et al*. (2024) [1] found out that adult users got the enjoyment the anonymity, privacy and easy access ability offered by artificial intelligence (AI) powered chatbots such as Woebot and Wysa. Similarly, Pandey (2024) [14] reported the fact that the majority of the users of the AI chatbots were adults who were are not ready to receive their treatment face to face because of the stigma and logistic reasons. Emotional check-ins and mood monitoring, in particular, turned out to be especially appreciated by the participants of this research results of Smith *et al*. (2023) [18], which stated that constant utilization of AI tools resulted in improved the emotional

awareness of the users. The convenience of AI tools therefore seems to act as a door to mental health care for users who may otherwise eschew official care environments. Nevertheless, as noted in the present research, such utility is mainly restricted to mild or moderate emotional distress; for more serious mental health requirements, AI utilities were perceived as insufficient. The second topic, Emotional Limitations of AI, captures prevailing concerns among participants about the absence of real empathy and shallow responses by chatbots. This is in accordance with the literature, which often reports emotional shallowness as one of the main limitations of AI therapy. Yoo *et al*. (2025) [23], for example, tested an experimental GPT-4-powered mental health chatbot and determined that users enjoyed the organized guidance but frequently felt disconnected on an emotional level with the system. Casu *et al*. (2024) [4] also mentioned that users reported answers given by AI tools to be repetitive or too scripted.

In the present study, subjects reported a feeling of frustration or misunderstanding when the chatbots could not comprehend emotional depth context depth. As one of the subjects said, even though the chatbot wrote generic positive somethings, it still helped since I could feel that I was in contact with a real person. rather than an innuendo of severe distress, it has failed to reply to one, which should be reinforced--a finding which is seconded by according to the report in Stanford (2025) [19], AI therapy bots misjudge 20 percent of the prompts in suicidal ideation. Such issues highlight a critical dilemma that, on the one hand, AI technologies are able to imitate aspects of therapeutic discussion, they have not reached the emotional intelligence and the human intuition that characterise human therapists (Gonzalez *et al*., 2023) [9]. The main concern of this study was the ethical issue since it is the superordinate issue in the research about data privacy and transparent. It was quite bothering to participants how sensitive the emotional state of this traumatic experience they had was to them. There was the possibility that information could be stored and even shared and possibly even manipulated. They acknowledged too to skimming over detailed privacy policies, which are long and confusing as well as containing mistrust. Such concerns are comparable to the ones in the study of Wu *et al*. (2024), who discovered that over 40 percent of people who used AI mental health tools remained unaware of its use how they has been treared about their information. MDPI Societies (2024) also found out that the users trained to use the platform had significant support of stricter information security regulations, especially in cases where they had a hand over the tools by the commercial tech companies versus health institution. In the given study, the trust was higher in cases where AI tools were affiliated to universities or healthcare organizations- which actually implies that user confidence is hugely influenced by affiliation. Along with this, there was the concern of algorithmic bias that also arose in this paper and literature.

According to Reuters (2024) [17], depression detection tools did well not well among Black users, and Gonzalez *et al*. (2023) [9] determined that NLP models cannot correctly read the expression of emotion in females because of skewed training data. The respondents of this study expressed similar fears that AI would get their moods wrong due to their genders cultural, or even lingual differences, which bring a well-founded question of equity and fairness. The fourth theme Desire of Human-AI Balance approach shows a cautious attitude towards AI as a human activity mental health care. The participants were not always opposed to using AI tools, but pointed out that they should always be applied with caution coming into use should act as an adjunct and not as an alternative to human therapists. The largest majority of them commented that Whereas, the AI interventions assisted in treating anxiety or low mood by creating a useful anchor point of the treatment lacked depth, empathy and plasticity in dealing with more subtle emotional problems. Some users had suggested the possibility to add in-app referrals to therapists or escalation procedures of high risk cases. This is aligned to studies by Behrman *et al*. (2024) [2] that advocates the exploitation of a hybrid system that includes AI screening and support providing, the combination with professional therapy access is available. Similarly, Miller *et al*. (2024) emphasize that combining an AI functionality with clinician review is both highly effective and trustful to users. According to the subjects of the present experiment, they also claimed that The problem of isolation might be overcome through human-AI cooperation, as AI tools may provide it normally failed to provide sympathetic follow-up or probing questions. The results suggest that ethical and human AI-use will be essential and will consist of clear-cut avenues of promotion to human care. In the final theme, Recommendations for Improvement, provided the informed user views of how It is possible to enhance AI tools. The majority of the recommendations mentioned the most frequently was the incorporation of emergency or crisis-response capability.

The participants were of the opinion that tools of AI ought to be capable of recognize distress signs of serious concern-such as suicidal ideation and provide immediate support or refer Malfunctioning of users to emergency services. This is in congruence with the report by Stanford University of 2025 that warned that alone AI tools can be hazardously underperforming in the case of high- risk situations. The study subjects also proposed cultural supple and simpler language to eased interaction among different populations. Ramirez *et al*. (2024) [16] supported this problem and confirmed that culturally adapted AI chatbots became influential in terms of the improvement of emotional engagement among Spanish-speaking users. In addition, customers wanted even greater transparency of their data and more personal interactions, too. is also in accordance to the recommendations by Wu *et al*. (2024) and Doshi *et al*. (2024) [7] that underline that Explainable AI (XAI) and customised responses are essential elements of ethical deployment.

These user inspired recommendations represent a new consensus in empirical and theoretical research that accountable AI should be transparent, nimble, and built with varied user demands in mind. Interestingly, although your results are otherwise in line with much of the current literature, they also uncover gaps and inconsistencies. For example, whereas research such as AlMakinah *et al*. (2024) [1] highlighted user satisfaction with chatbots such as Woebot, your results identify considerable emotional dissatisfaction and restriction in depth of care. This implies that while higher quantitative user satisfaction ratings would be noted, qualitative interviewing uncovers more ambivalent or nuanced attitudes. Moreover, although Smith *et al*. (2023) [18] presented the effectiveness of wearable AI in identifying early mood disorders, your participants prioritized subjective safety and trust above technical success. That is, technical success does not necessarily equate to user confidence—

prioritizing the necessity of human-centered design and co-creation with target groups. Moreover, this research extends existing literature through challenging diverse adult experiences, specifically those from underrepresented voices. Most previous research has been on young adults, especially students, but the present research had a more diverse group of participants in the age range 21 to 52. This wider age group enables inferences from working professionals, caregivers, and those juggling more social roles. It brings to the fore practical hindrances such as digital competency, cultural lack of fit, and affective incompatibility that do not get brought into view in more strictly controlled experimental research. Finally, the experiment suggests that regulation rules and moral codes should be established. With no The strategy, the AI mental health application can be appropriated on a one-sided basis without its effectiveness or safety left established, stimulating or causing harm or increasing mental health disparities. According to Denecke *et al*. It was emphasized by (2021) that AI in mental health has to be more than technologically proficient, it has to be morally evidence-based, evidence-tested and evidence-disseminated. Findings in the study disclose that adult customers would engage with AI mental health platforms under the condition of quality, understanding, safe and secure experiences that respect their statistics, emotions and individuality.

Conclusion

In this study, the knowledge and moral experiences of adults using AI-based were revealed the potential and the limitation of mental health applications. The participants in a major extent valued the convenience, privacy, and tenderness of AI applications, which helped them to handle mild identifying an emotional distress and improvement of self-knowledge. There was also a case of primary limitations cited by users, containing the emotional world, denial of human compassion and crisis intervention functionality.Ethical concerns such as data privacy, transparency, and algorithmic bias emerged as central issues shaping user trust and engagement. A recurring recommendation was the integration of human-AI collaboration, where AI tools act as supportive resources rather than replacements for professional mental health care. Participants emphasized that AI should supplement, not substitute, therapeutic relationships—especially for complex or high-risk cases. These results complement and expand current work, providing significant implications for developers, clinicians, and policymakers. To become a safe, inclusive, and useful mental health tool, AI needs to be ethically developed, clinically tested, and culturally sensitive, with unambiguous escalation routes to human care. Co-creation with user groups from diverse backgrounds, long-term assessment, and implementation within larger mental health systems must be given precedence in future work to ensure that AI augments—and is not a substitute for—compassionate, ethical care.

Limitation

- The sample size of the study was small (n = 18), which restricts the generalizability of results to the wider adult population.
- The participants self-selected, which could introduce selection bias—those with stronger views might have been more inclined to participate.
- The research emphasized user experiences with multiple AI technologies, but these technologies vary significantly in design, quality, and ethics, so it is hard to generalize platforms.
- Users' cultural and linguistic diversity was restricted; marginalized or non-English-speaking people's experiences were under-represented.
- Having been a qualitative study, the interpretation of data incorporated subjective analysis. There is a possibility of researcher bias in theme development despite attempts at credibility.
- The research did not measure clinical outcomes or changes in mental health over time; it was completely based on perceptions and experiences.
- Results are for a certain point in time and may not reflect changing technologies or future enhancements of AI tools.

References

1. AlMakinah N, Jameel S, Abbas A. AI mental health chatbots and adult engagement: An ethical review. Digital Health Review. 2024;12(1):55-70.
2. Behrman K, Ali R, Mendez C. AI and mental health in low-resource settings: Evaluating Woebot's equity potential. Global Mental Health Journal. 2024;6(2):90-104.
3. Braun V, Clarke V. Using thematic analysis in psychology. Qualitative Research in Psychology. 2006;3(2):77-101.
4. Casu L, Toma E, Singh P. Artificial intelligence in therapy: A systematic review of chatbot effectiveness. Journal of Digital Mental Health. 2024;11(3):100-117.
5. Cho SH, Park EJ, Choi YK. Predictive analytics for PTSD relapse using machine learning: A case study on veterans. Journal of Affective Disorders. 2020;274:470-478.
6. Denecke K, Gabarron E, Grainger R, Hansen M. Ethical challenges of artificial intelligence in public health. Frontiers in Public Health. 2021;9:652593.
7. Doshi A, Patel S, Zhu H. Evaluating adult user satisfaction with TheraGen AI for mental health: A pilot study. Psychiatric AI Research. 2024;9(1):42-59.
8. 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 Mental Health. 2017;4(2):e19.
9. Gonzalez R, Lin D, Okafor T. Gendered misclassifications in sentiment AI: Implications for mental health tech. AI and Society. 2023;38(1):145-160.
10. Jin L, Zhang Y, Wang F. Artificial intelligence in health care: Opportunities, challenges, and future directions. International Journal of Medical Informatics. 2023;168:104897.
11. Lee D, Chen H, Park J. Pandemic-triggered mental health crisis: The urgent need for scalable digital solutions. Global Health Journal. 2021;15(3):210-218.
12. Lu X, Tang W, Huang J. Personalizing AI for patient-centered mental health care: Potentials and pitfalls. AI in Healthcare. 2023;7(4):180-193.
13. Miller C, Turing H, Alston M. Hybrid models in mental health AI: Enhancing care through clinician-AI collaboration. Journal of Ethical AI Systems. 2024;10(2):112-130.
14. Pandey N. Understanding adults' preferences for

chatbot-based therapy: A behavioral lens. Digital Psychology Review. 2024;5(1):27-45.
15. Pandi-Perumal SR, Bhat S, Zizi F. Speech-based detection of depressive symptoms using AI: Advances and applications. Neuropsychiatric Disease and Treatment. 2023;19:1001-1012.
16. Ramirez L, Gutierrez M, Lopez D. Designing culturally adaptive AI chatbots for mental health in Mexico. Latin American Journal of eHealth. 2024;8(1):15-28.
17. Reuters. Study finds racial bias in popular AI depression detection tool. Reuters Health News. 2024. Available from: https://www.reuters.com.
18. Smith RD, Zhao J, Whitman M. Predicting depressive episodes with wearable AI: A longitudinal pilot study. Journal of Behavioral Health Technology. 2023;10(4):299-312.
19. Stanford University. AI mental health assistants and crisis recognition: Limitations and opportunities. Stanford Digital Mental Health Lab Report. 2025;5(2):1-20.
20. The New York Times. When your AI therapist becomes your best friend. The New York Times. 2025. Available from: https://www.nytimes.com.
21. World Health Organization. Ethics and governance of artificial intelligence for health: Guidance from the WHO. Geneva: World Health Organization; 2023.
22. Wu L, Chan T, Elangovan S. Perceptions of clinicians and users towards AI mental health technologies: A cross-sectional analysis. Journal of Medical Internet Research. 2024;26(1):e48922.
23. Yoo H, Kim SJ, Jung H. Exploring emotional resonance in GPT-4 powered mental health chatbots. AI & Emotion Research. 2025;12(1):88-104.