Pakistan Journal of Positive Psychology

Print ISSN : 3078-6428 | Online ISSN : 3078-6436
Pakistan Journal of Positive Psychology

Original Article

How Artificial Intelligence is Changing Psychological Practice: A Qualitative Study of Clinicians' Perceptions and Ethical Concerns

Fazle Khaliq[a*] Haider Ali [b], & Mukrameen Khan [c]

[a] Lecturer, National Cyber Security Academy, Air University, Islamabad, Pakistan
[b] Clinical Psychologist, Swat Psychiatry and Rehab Hospital, Swat, Pakistan
[c] Lecturer, Department of Psychology, University of Veterinary and Animal Sciences, Swat, Pakistan

*Correspondence to: Fazle Khaliq, Lecturer, National Cyber Security Academy, Air University Islamabad, Pakistan. Email: fazlekhaliqpsychology@gmail.com

KEYWORDS:

Artificial Intelligence, Clinicians' Perceptions, Ethical Concerns, Mental Health Technology, Thematic Analysis

CITE THIS ARTICLE:

Khaliq, F., Ali, H., & Khan, M. (2026). (2026). How artificial intelligence is changing psychological practice: a qualitative study of clinicians' perceptions and ethical concerns. Pakistan Journal of Positive Psychology, 3(10), 35-45.

ABSTRACT

Artificial intelligence (AI) is increasingly embedded in psychological assessment, monitoring, and treatment, yet little is known about how practicing clinicians in low- and middle-income countries experience this shift. This qualitative descriptive study explored clinicians' perceptions of AI in psychological practice and the ethical concerns they associate with its adoption. Eight clinicians (four men, four women) working in hospitals and private clinics in Islamabad and Rawalpindi, Pakistan, were recruited through purposive sampling and completed semi-structured online interviews. Transcripts were analyzed with Braun and Clarke's (2006) six-phase thematic analysis. Three themes and eight subthemes were developed: (a) Enhanced Efficiency with Persistent Uncertainty, (b) Erosion of the Human Touch in Psychological Care, and (c) Ethical Vulnerabilities in AI-Assisted Practice. Clinicians valued faster scoring, structured reports, and reduced administrative burden, but doubted the accuracy, transparency, and cultural validity of AI outputs. They regarded empathy and relational depth as irreplaceable and feared that AI could unsettle the therapeutic alliance and provoke client distrust. Ethical concerns centered on data privacy, informed consent, algorithmic bias, and professional autonomy and liability. The findings suggest that perceived ethical and relational risk can outweigh perceived usefulness in clinicians' appraisal of AI, extending technology acceptance models to psychological practice. Implications for training, clinical guidelines, regulation, and culturally responsive AI development in South Asia are discussed.

I. INTRODUCTION

Artificial intelligence (AI) has moved from the margins of mental health research into everyday clinical contexts. Screening algorithms, automated scoring platforms, symptom-tracking applications, natural language processing tools, and conversational agents now support the way clinicians assess, monitor, and treat psychological distress (Graham et al., 2019; Lee et al., 2021; Torous et al., 2021). Early conceptual work anticipated this shift and warned that psychologists would need to engage with its practical and ethical consequences (Luxton, 2014a, 2014b). The pace has quickened with generative AI. Large language models are now proposed as tools for documentation, training, and even therapeutic dialogue (Stade et al., 2024), and a recent randomized trial of a generative AI chatbot reported symptom reductions among adults with depression, anxiety, or eating-disorder risk (Heinz et al., 2025). Earlier rule-based chatbots showed encouraging but modest and short-term effects (Abd-Alrazaq et al., 2020; Fitzpatrick et al., 2017), and the wider evidence base remains uneven in quality (Torous et al., 2021).

The promise is practical. AI can shorten assessment, standardize scoring, and extend contact to people who would otherwise receive no service, which matters most in low- and middle-income countries where specialist workforces are small and the treatment gap is wide (Naslund et al., 2017; Rathod et al., 2017). The risks are equally practical. Scholars have raised concerns about the accuracy and opacity of algorithmic outputs, confidentiality and data governance, informed consent, the erosion of clinician autonomy, and harm from biased or culturally misaligned models (Fiske et al., 2019; Luxton, 2014b; Straw & Callison-Burch, 2020; World Health Organization [WHO], 2021). Because psychological assessment and treatment involve highly sensitive information and rest on a trusting relationship, these risks cannot be treated as technical details to be resolved after deployment.

Problem Statement

Clinicians are the people who decide whether, when, and how AI tools enter practice, and their perceptions shape whether those tools are adopted, resisted, or quietly worked around. Yet much of the literature on AI in mental health is technical, conceptual, or evaluative rather than grounded in practitioners' accounts of their own work. Survey research with psychiatrists shows considerable doubt that AI can replicate empathy and relational care (Blease et al., 2020; Doraiswamy et al., 2020), and recent interview studies with psychologists in Europe identify knowledge gaps, privacy worries, and fear of job displacement as barriers to use (de la Fuente Tambo et al., 2025). Evidence from South Asia is thin. A multicenter survey of psychiatrists in Pakistan found limited exposure to AI training and uneven familiarity with digital tools (Riaz et al., 2024), but we located no qualitative study that explores how Pakistani clinicians understand AI in psychological practice or which ethical issues they regard as most pressing. Without such evidence, tools developed and validated elsewhere risk being introduced in ways that weaken professional relationships, overlook local idioms of distress, or conflict with professional codes (American Psychological Association [APA], 2017).

Purpose of the Study and Research Questions

The purpose of this qualitative study was to explore how clinicians in Islamabad and Rawalpindi perceive the role of AI in psychological practice and which ethical concerns they associate with its use. Using thematic analysis of in-depth online interviews, we aimed to generate practice-based knowledge that can inform policy, training, and tool development. Three research questions guided the study:

1. How do clinicians perceive the impact of AI on psychological assessment, diagnosis, and intervention?
2. What ethical concerns do clinicians identify regarding the adoption of AI in psychological practice?
3. How do clinicians envision the future role of AI in clinical practice?

This study contributes to the growing literature on AI and mental health by foregrounding practitioners, who are the primary actors in how technologies take hold in clinical settings. The findings can inform policymakers, developers, and regulators about concrete ethical problems clinicians expect to face. They can also guide professional training, particularly in digital ethics, algorithmic literacy, and technology-assisted intervention. These insights are especially relevant in Pakistan, where mental health resources are scarce and AI-based services are expanding, and where culturally and ethically responsible design cannot be borrowed wholesale from high-income settings.

II. LITERATURE REVIEW

AI refers to computational systems that perform tasks traditionally associated with human cognition, such as pattern recognition, language processing, prediction, and decision support (Topol, 2019). In mental health settings, these capabilities appear as automated screening tools, natural language processing models that detect risk in speech or text, conversational agents that deliver psychoeducation or structured therapeutic exercises, and models that predict course of illness or treatment response (Bendig et al., 2022; Graham et al., 2019; Vaidyam et al., 2019). Clinicians' perceptions, meaning their subjective appraisals, expectations, and emotional reactions to such tools, influence whether technologies are taken up, resisted, or adapted to clinical routines (Lee et al., 2021). The ethical issues most often associated with AI in healthcare include confidentiality and data security, informed consent, transparency and explainability, fairness, professional autonomy and accountability, and the potential for harm from incorrect or biased outputs (Fiske et al., 2019; Jobin et al., 2019; WHO, 2021). In psychological practice these issues are intensified by the sensitivity of the data and by the centrality of the therapeutic relationship, which makes them a distinctive set of concerns rather than a generic subset of medical AI ethics.

A scoping review of machine learning in mental health found that most applications focused on detection and diagnosis, with fewer addressing prognosis, treatment, and support (Shatte et al., 2019). Reviews of conversational agents describe a fast-growing field with promising but preliminary evidence (Abd-Alrazaq et al., 2020; Vaidyam et al., 2019). In one early trial, a fully automated cognitive behavioral agent reduced depressive symptoms in young adults over two weeks (Fitzpatrick et al., 2017), and a more recent randomized trial of a generative AI chatbot reported clinically meaningful symptom reductions (Heinz et al., 2025). Even so, systematic reviews caution that study quality is variable, follow-up is short, and safety reporting is limited (Abd-Alrazaq et al., 2020; Lee et al., 2021). Evidence on harms and bias is equally important. Algorithms trained on unrepresentative data can reproduce or amplify inequities, as demonstrated in a widely cited analysis of a commercial health-risk algorithm that systematically underestimated the needs of Black patients (Obermeyer et al., 2019). In mental health specifically, language-based models have been shown to encode biases related to attributes such as gender, religion, and race (Straw & Callison-Burch, 2020). These findings are directly relevant to settings where clients' idioms of distress differ from the populations on which most tools are trained.

Qualitative and survey studies suggest that clinicians hold ambivalent views. In a global survey of psychiatrists, respondents anticipated that AI could assist with documentation, information synthesis, and risk prediction, yet most doubted that it could replicate empathic, relational care (Blease et al., 2020; Doraiswamy et al., 2020). In interviews with 14 private-practice psychologists in Spain, de la Fuente Tambo et al. (2025) identified 18 factors shaping acceptance of generative AI, including knowledge gaps, data privacy concerns, fear of peer judgment and job displacement, and anticipated client resistance, alongside enablers such as AI training and administrative assistance. Reviews of the field note that clinician attitudes, training, and trust are among the main barriers and facilitators of implementation (Lee et al., 2021). Research from South Asia is scarce. A multicenter survey of psychiatrists in Pakistan reported limited AI training and uneven familiarity with digital mental health tools (Riaz et al., 2024). The wider literature on digital mental health in low- and middle-income countries points to promise for extending services but also to infrastructural, linguistic, and cultural barriers (Naslund et al., 2017). Few studies, however, have asked practicing clinicians in these contexts to describe in their own words how AI is changing their work and what worries them. That gap motivates the present study.

This exploratory study is conceptually anchored in the Technology Acceptance Model (TAM; Davis, 1989) and in principle-based professional ethics. TAM proposes that perceived usefulness and perceived ease of use shape intentions to adopt a technology, and later models such as the unified theory of acceptance and use of technology added social influence and facilitating conditions (Venkatesh et al., 2003). In psychological practice, clinicians' judgments of an AI tool's clinical relevance and fit with their workflow may therefore influence whether they integrate it into assessment or treatment decisions. Acceptance models, however, were developed for general information systems and say little about ethical risk or relational harm, which are salient in clinical work. We therefore complemented TAM with principle-based ethics. The APA's (2017) *Ethical Principles of Psychologists and Code of Conduct* articulates beneficence and nonmaleficence, fidelity and responsibility, integrity, justice, and respect for people's rights and dignity, and these principles parallel the widely used biomedical ethics framework of Beauchamp and Childress (2019). We also treated the therapeutic alliance as a sensitizing concept, given robust evidence that the alliance is reliably associated with psychotherapy outcomes (Flückiger et al., 2018; Norcross & Lambert, 2018). Together these frameworks guided the interview questions but did not constrain the inductive development of themes.

III. METHOD

Participants

We used a qualitative descriptive design to explore clinicians' perceptions of AI in psychological practice and the ethical concerns surrounding its use. Qualitative inquiry is well suited to understanding subjective experiences, meanings, and interpretations (Creswell & Poth, 2018). Because the aim was to capture clinicians' perspectives in their own words, semi-structured interviews were the primary source of data, allowing participants to elaborate on issues they considered important while keeping the conversation anchored to the research questions. We recruited practicing clinicians through purposive sampling, which allows researchers to select information-rich participants who can speak in depth to the phenomenon of interest (Patton, 2015). Inclusion criteria were (a) being a clinical

psychologist, (b) currently working in a hospital or private clinic, and (c) having direct or indirect exposure to digital or AI-assisted mental health tools. The sample comprised eight clinicians (four men, four women) working in hospitals and private clinics in Islamabad and Rawalpindi, Pakistan. Participants were aged 29 to 45 years (*M* = 36.6) and had 6 to 18 years of clinical experience (*M* = 10.9); five worked in hospitals and three in private clinics. All had encountered AI-based tools in their work settings, including online assessment platforms, symptom trackers, chatbot-based psychoeducation, or predictive screening systems. Participant characteristics are shown in Table 1. Participants are identified by codes (P1–P8) throughout to protect confidentiality.

Table I: Demographic Characteristics of Participants (N = 8)

Participant	Gender	Age (years)
P1	Female	34
P2	Male	40
P3	Female	37
P4	Male	31
P5	Female	29
P6	Male	45
P7	Female	41
P8	Male	36

Note. Years of clinical experience ranged from 6 to 18 (*M* = 10.9). Five participants worked in hospitals and three in private clinics.

Procedure and Data Collection

Data collection began after approval from the responsible institutional ethics review committee. Participants received an information sheet describing the study purpose, confidentiality safeguards, and their right to withdraw at any time, and each gave written informed consent before the interview. Interviews were conducted individually on secure video-conferencing software, an established and acceptable mode of qualitative data collection (Archibald et al., 2019). Each lasted approximately 40 to 60 minutes and was audio-recorded with permission. A flexible interview guide covered clinicians' experiences with AI-based tools, perceived benefits and risks, ethical concerns, and expectations for the future of practice, while leaving room for unanticipated issues. The interviewer kept field notes on context and emerging ideas. Recruitment and interviewing continued until thematic saturation was judged to have been reached. Recordings were transcribed verbatim, anonymized, and stored securely.

Data Analysis

We followed the six-phase approach to thematic analysis described by Braun and Clarke (2006) to identify, organize, and interpret patterns of meaning across the dataset.

1. **Familiarization.** The lead researcher read the transcripts repeatedly and listened again to the recordings to attend to tone, pauses, and emphasis. First impressions and reflective memos were recorded to inform later coding.
2. **Generating initial codes.** Transcripts were coded line by line. Early codes were descriptive and semantic, capturing meaningful units such as perceived usefulness of AI, data security concerns, and perceived changes to therapeutic processes. Across the eight interviews, 142 initial codes were generated.
3. **Searching for themes.** Codes were collated into candidate themes: perceived clinical benefits, concerns about accuracy, effects on the therapeutic relationship, and ethical and professional dilemmas. Visual mapping and thematic matrices were used to examine relationships among codes.
4. **Reviewing themes.** Candidate themes were checked against the coded extracts and the full dataset for coherence and distinctiveness. Some themes were merged or refined, and others were dropped for lack of support, leaving a smaller set of well-supported themes and subthemes.
5. **Defining and naming themes.** Each theme was defined around a central organizing concept, and a short analytic narrative was written to specify what it captured and how it addressed the research questions. Final names were Enhanced Efficiency with Persistent Uncertainty, Erosion of the Human Touch in Psychological Care, and Ethical Vulnerabilities in AI-Assisted Practice.
6. **Producing the report.** Themes, participant quotations, and interpretation were woven into a coherent account linked to the research questions and the wider literature, keeping participants' voices central.

Trustworthiness and Ethical Considerations

We drew on the criteria for qualitative quality described by Tracy (2010). Rigor was supported by repeated engagement with recordings and transcripts, line-by-line coding, and the use of reflective memos and thematic matrices as a record of analytic decisions. Credibility was supported by returning to the full dataset when reviewing themes and by presenting thick description with verbatim quotations so that readers can judge the fit between data and interpretation. Ethical practice, described next, was treated as integral to quality. The study received ethical approval from an institutional ethics review committee. Participants were informed of the study's aims, the confidentiality of their responses, the voluntary nature of participation, and their right to withdraw at any time, and they provided electronic informed consent. Transcripts were anonymized, identifying details were removed, and participants were identified only by codes (P1–P8). Audio files and transcripts were kept in encrypted folders accessible only to the research team. Because interviews touched on workplace technologies and organizational policies, these topics were handled sensitively so that participation could not jeopardize participants' employment.

IV. RESULTS

Thematic analysis produced three themes and eight subthemes describing how clinicians perceive AI to be changing psychological practice and the ethical concerns they associate with its adoption (see Table 2). Each theme is presented with representative quotations; participants are identified by code, gender, and age.

Table II: Themes and Subthemes Describing Clinicians' Perspectives on AI in Psychological Practice

Theme	Subtheme
1. Enhanced Efficiency with Persistent Uncertainty	1.1 Streamlining assessment and administrative work
	1.2 Questioning accuracy and clinical judgment
2. Erosion of the Human Touch in Psychological Care	2.1 Threat to relational depth and empathy
	2.2 Client distrust and miscommunication
3. Ethical Vulnerabilities in AI-Assisted Practice	3.1 Privacy and data security risks
	3.2 Informed consent and transparency challenges
	3.3 Algorithmic bias and cultural insensitivity
	3.4 Professional autonomy and liability concerns

Theme I: Enhanced Efficiency with Persistent Uncertainty

Clinicians widely acknowledged that AI tools brought efficiencies to parts of psychological practice. These benefits were consistently accompanied by uncertainty about accuracy, reliability, and contextual appropriateness.

Subtheme I.I: Streamlining Assessment and Administrative Work

Participants described AI-supported assessment platforms and symptom-tracking applications as reducing paperwork, generating quicker summaries, and automating scoring. P1 (female, 34) said, *"AI definitely speeds up the initial screening. Instead of manually scoring everything, I get structured reports instantly."* P2 (male, 40) added, *"It saves time, especially with large caseloads. The automated interpretations give me a head start."* Clinicians emphasized that these tools let them redirect time toward therapeutic interaction rather than administrative tasks.

Subtheme I.II: Questioning Accuracy and Clinical Judgment

Despite these gains, clinicians were skeptical that AI could capture nuanced human experience. P3 (female, 37) explained, *"These tools are good for structure, but they miss the emotional and cultural layers that only a clinician can understand."* P4 (male, 31) worried about over-reliance: *"I worry about over-trusting an algorithm. What if it interprets something wrongly and we rely on it?"* Uncertainty was strongest around diagnostic accuracy and predictive analytics, which some clinicians considered *"too black box"* for responsible decision-making.

Theme II: Erosion of the Human Touch in Psychological Care

Clinicians frequently voiced concern that AI could harm the therapeutic relationship, which they regarded as foundational to effective practice.

Subtheme II.I: Threat to Relational Depth and Empathy

Participants were apprehensive about AI's inability to empathize, attune, or respond to clients' emotional nuance. P5 (female, 29) said, *"You can't replace human empathy. Clients open up because they feel understood, not processed."* P6 (male, 45) noted, *"Therapy is relational. AI can't pick up on micro-expressions, tone changes, or the overall atmosphere in the room."* Several clinicians stressed that the therapeutic alliance could weaken if clients perceived AI as taking precedence over human judgment.

Subtheme II.II: Client Distrust and Miscommunication

Participants reported that clients sometimes became confused or anxious when told that digital tools were used in their assessment or monitoring. P7 (female, 41) recalled, *"One client asked me if the AI would 'judge' him. There is a lot of fear because people don't know how the system works."* P8 (male, 36) observed, *"Clients think AI means surveillance, and that makes them uncomfortable."* Clinicians feared that this unease could unintentionally disrupt rapport building.

Theme III: Ethical Vulnerabilities in AI-Assisted Practice

Ethical concerns ran through clinicians' accounts and centered on privacy, consent, professional autonomy, and fairness.

Subtheme III.I: Privacy and Data Security Risks

Participants consistently voiced concern about confidentiality and data safety, especially in online data collection systems. P1 (female, 34) said, *"I don't know where the client data is going. The companies don't always explain their storage or encryption policies clearly."* P2 (male, 40) warned, *"There is a real risk of breaches. These tools collect sensitive psychological information."* Clinicians were particularly uneasy about third-party software operated outside healthcare settings.

Subtheme III.II: Informed Consent and Transparency Challenges

Clinicians stressed the importance of explaining AI tools to clients in understandable terms. P3 (female, 37) asked, *"Most clients don't understand AI. So how do we ensure they truly consent?"* P4 (male, 31) acknowledged, *"We must tell them what data is used and how it's analyzed, but sometimes I myself don't fully know the algorithmic details."* Limited algorithmic transparency made ethical disclosure difficult.

Subtheme III.III: Algorithmic Bias and Cultural Insensitivity

Participants worried that models trained in Western contexts might misread local linguistic and cultural expressions. P5 (female, 29) said, *"Many tools are trained on foreign samples. They don't understand our cultural expressions of distress."* P6 (male, 45) added, *"An algorithm may flag something as abnormal simply because it's unfamiliar with our norms."*

Subtheme 3.4: Professional Autonomy and Liability Concerns

Clinicians feared that AI could influence clinical decisions in ways that undermine professional expertise. P7 (female, 41) asked, *"If an AI suggests something different from my clinical judgment, what do I follow? Who is responsible if something goes wrong?"* P8 (male, 36) said, *"There is pressure to accept AI outputs, but that puts us in a tricky legal and ethical position."* The tension between human discretion and algorithmic recommendation emerged as a central ethical dilemma.

Summary of Findings

Clinicians perceived AI as both a facilitator and a disruptor of psychological practice. They acknowledged efficiency gains but expressed strong reservations about accuracy, relational depth, privacy, consent, cultural appropriateness, and professional control. Across all themes, participants emphasized keeping human-centered values at the core of care as technology becomes more integrated into mental health services. Figure 1 shows the most frequent words in the interview data.

Figure I: Word Cloud of the Most Frequent Words in Participants' Interview Data



V. DISCUSSION

This study explored how clinicians in Islamabad and Rawalpindi perceive AI in psychological practice and what ethical concerns they associate with it. Three themes captured their accounts: Enhanced Efficiency with Persistent Uncertainty, Erosion of the Human Touch in Psychological Care, and Ethical Vulnerabilities in AI-Assisted Practice. Together they describe a profession that sees real practical value in AI but is wary of its limits, its effect on the therapeutic relationship, and the ethical exposure it creates. In answer to the first research question, clinicians perceived AI as improving the speed and structure of assessment while leaving diagnostic and interpretive judgment with the clinician. In answer to the second, privacy, consent, bias, and accountability emerged as the central ethical problems. In answer to the third, participants envisioned AI as support for human judgment, handling scoring and administration so that more time is available for clients, rather than as a substitute for the clinician.

Efficiency Without Certainty

Participants' appreciation of automated scoring, structured reports, and reduced paperwork is consistent with reviews that identify workflow support and decision assistance as promising and comparatively low-risk applications of AI in mental health care (Graham et al., 2019; Lee et al., 2021) and with findings that administrative assistance is an enabler of acceptance among psychologists (de la Fuente Tambo et al., 2025). Yet efficiency was never presented as sufficient. Clinicians doubted that tools could capture emotional and cultural nuance and described some models as "too black box" for responsible decisions. This mirrors survey evidence that psychiatrists expect AI to support, not replace, clinical judgment (Blease et al., 2020; Doraiswamy et al., 2020) and the broader concern that opaque models undermine clinicians' ability to explain and defend their decisions (Fiske et al., 2019; WHO, 2021). Clinicians in this study appeared willing to use AI for preliminary and administrative tasks but not to rely on it for delicate clinical determinations.

Protecting the Human Touch

The second theme reflects the primacy clinicians place on relationship. Participants described empathy, attunement, and the reading of subtle cues as irreplaceable, and they worried that clients' reactions to AI could damage rapport. These views are in line with decades of evidence that the therapeutic alliance is reliably associated with treatment outcomes (Flückiger et al., 2018; Norcross & Lambert, 2018) and with psychiatrists' doubts that technology can provide empathic care (Doraiswamy et al., 2020). They also complement trial evidence that some users form a working bond with well-designed chatbots (Heinz et al., 2025): the clinicians interviewed here were less concerned with whether clients can relate to a tool than with whether AI in the room changes how clients relate to the clinician. The client distrust participants described, including fears that AI means judgment or surveillance, suggests that how AI is introduced may matter as much as what it does. This is an area where clients' own voices are needed.

Ethical Vulnerability in Context

The third theme shows that ethical concern was not peripheral but integral to clinicians' thinking. Worries about data storage and third-party software, doubts about whether clients can give meaningful consent to processes that clinicians themselves do not fully understand, and anxiety about liability when AI and clinical judgment diverge correspond closely to the issues highlighted in ethical analyses of AI in mental health care and in international governance guidance (Fiske et al., 2019; Jobin et al., 2019; Luxton, 2014b; WHO, 2021). They also map onto the APA's (2017) principles of fidelity and responsibility, justice, and respect for people's rights and dignity, which suggests that existing professional codes provide a sound starting point but may need explicit guidance for algorithmic tools.

The concern about cultural misalignment deserves emphasis. Participants worried that tools trained on Western samples would misread local expressions of distress and flag normal variation as abnormal. Empirical work shows that language-based models can encode demographic and cultural biases (Straw & Callison-Burch, 2020) and that algorithmic bias can produce systematic inequities in health care (Obermeyer et al., 2019). The clinicians' fears therefore rest on a documented mechanism rather than on technophobia, and they are heightened in a setting where regulation of digital mental health is still developing and where AI training for clinicians is limited (Riaz et al., 2024). Compared with research from Europe and North America, which emphasizes privacy, transparency, and the alliance (de la Fuente Tambo et al., 2025; Fiske et al., 2019), this study adds that cultural validity and local regulatory uncertainty are prominent concerns in South Asian practice. Because the sample was small and drawn from one region, this comparison should be treated as a hypothesis for further testing rather than a conclusion.

Theoretical and Clinical Implications

The findings are consistent with technology acceptance theory in that perceived usefulness, particularly efficiency, was acknowledged by all participants (Davis, 1989; Venkatesh et al., 2003). However, perceived usefulness alone did not translate into endorsement. Perceived ethical and relational risk appeared to constrain acceptance even where usefulness was recognized. This suggests that acceptance models applied to psychological practice should incorporate constructs such as perceived threat to the therapeutic alliance, perceived ethical and legal exposure, and perceived cultural validity. Testing such an extended model quantitatively would be a useful next step. Clinicians need structured training in digital literacy, algorithmic behavior, and the responsible use of AI-assisted tools, including how to explain these tools to clients. Clinical guidelines should protect the therapeutic alliance, require transparency, and preserve clinicians' final authority over decisions. AI outputs should be framed as decision support, not as a substitute for professional judgment. Practically, services adopting AI could use plain-language consent procedures that tell clients what data are collected, where they are stored, and how outputs are used, and could establish routine review of cases in which AI outputs and clinical impressions diverge.

Policy and Development Implications

For policymakers and regulators, the findings point to an urgent need for:

- Clear data protection rules and standards for the responsible use of AI in health care, including requirements on storage, encryption, and third-party data sharing;
- Standards for informed consent specific to algorithmic tools;
- Mechanisms to monitor and reduce cultural and linguistic bias in AI tools used with Pakistani populations; and
- Locally grounded model development and validation that reflects regional psychological norms and idioms of distress.

For developers, clinician input should be built into the design process so that tools are usable, acceptable, transparent about how outputs are generated, and ethically aligned with professional codes.

Limitations and Future Directions

Several limitations should be considered. First, the sample was small (eight clinicians) and drawn from one geographic area, so the findings are not statistically generalizable, and perspectives from other provinces, public-sector community services, or rural practice may differ. Sample size in qualitative research is best judged against the study's aim and the specificity of the sample (Malterud et al., 2016), and some methodological work suggests that richer understanding of heterogeneous topics can require larger samples than the number needed to identify the main codes (Hennink et al., 2017; see also Guest et al., 2006). Second, participants differed in their direct experience with AI, and those with limited exposure may have drawn more on expectations than on sustained use. Third, interviews were conducted online, which may have affected rapport and the richness of responses. Fourth, the data

were self-reported and may be shaped by social desirability or recall bias, particularly when participants discussed workplace technologies. Finally, coding was carried out primarily by a single researcher, so the themes reflect one analytic perspective even though memos and thematic matrices documented decisions.

Future research should include larger and more diverse samples across regions to capture cultural and institutional variation, and should explore how clients perceive AI in therapy, since client comfort and trust are central to the process of treatment. Longitudinal qualitative studies could trace how clinicians' perceptions change as AI becomes more embedded in mental health services, including with the rapid arrival of generative AI (Stade et al., 2024). Mixed-methods and quantitative studies could test an extended acceptance model that includes relational and ethical risk. Studies are also needed on culturally specific AI tools developed and validated with South Asian populations. Finally, interdisciplinary collaboration among data scientists, psychologists, ethicists, and policymakers is essential for building AI systems that are ethical, transparent, and clinically useful.

VI. CONCLUSION

This study examined how practicing clinicians in Islamabad and Rawalpindi perceive AI in psychological practice and the ethical issues it raises. Thematic analysis identified three overarching themes: enhanced efficiency with persistent uncertainty, erosion of the human touch, and ethical vulnerabilities involving privacy, consent, bias, and professional autonomy. Clinicians saw AI as a double-edged technology: it can streamline assessment and reduce administrative load, but it also carries risks to the relational and ethical foundations of psychological care that, in their view, computational systems cannot replicate. Concerns about accuracy, cultural sensitivity, and transparency were the main obstacles to wider adoption. As AI becomes more common in mental health services, its integration should be guided by human-centered values, rigorous training, culturally validated tools, and clear ethical and regulatory safeguards, with clinicians' voices informing each step.

Disclosure Statement

No potential conflict of interest was reported by the authors.

Funding

The authors received no funding from any organizations.

VII. REFERENCES

Abd-Alrazaq, A. A., Rababeh, A., Alajlani, M., Bewick, B. M., & Househ, M. (2020). Effectiveness and safety of using chatbots to improve mental health: Systematic review and meta-analysis. *Journal of Medical Internet Research, 22*(7), Article e16021. https://doi.org/10.2196/16021

American Psychological Association. (2017). *Ethical principles of psychologists and code of conduct* (2002, amended effective June 1, 2010, and January 1, 2017). https://www.apa.org/ethics/code

Archibald, M. M., Ambagtsheer, R. C., Casey, M. G., & Lawless, M. (2019). Using Zoom videoconferencing for qualitative data collection: Perceptions and experiences of researchers and participants. *International Journal of Qualitative Methods, 18*. https://doi.org/10.1177/1609406919874596

Beauchamp, T. L., & Childress, J. F. (2019). *Principles of biomedical ethics* (8th ed.). Oxford University Press.

Bendig, E., Erb, B., Schulze-Thuesing, L., & Baumeister, H. (2022). The next generation: Chatbots in clinical psychology and psychotherapy to foster mental health: A scoping review. *Verhaltenstherapie, 32*(Suppl. 1), 64–76. https://doi.org/10.1159/000501812

Blease, C., Locher, C., Leon-Carlyle, M., & Doraiswamy, M. (2020). Artificial intelligence and the future of psychiatry: Qualitative findings from a global physician survey. *Digital Health, 6*. https://doi.org/10.1177/2055207620968355

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology, 3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Creswell, J. W., & Poth, C. N. (2018). *Qualitative inquiry and research design: Choosing among five approaches* (4th ed.). SAGE Publications.

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13*(3), 319–340. https://doi.org/10.2307/249008

de la Fuente Tambo, D., Iglesias Moreno, S., & Armayones Ruiz, M. (2025). Barriers and enablers for generative artificial intelligence in clinical psychology: A qualitative study based on the COM-B and theoretical domains framework (TDF) models. *BMC Psychology, 13*, Article 1181. https://doi.org/10.1186/s40359-025-03500-7

Doraiswamy, P. M., Blease, C., & Bodner, K. (2020). Artificial intelligence and the future of psychiatry: Insights from a global physician survey. *Artificial Intelligence in Medicine, 102*, Article 101753. https://doi.org/10.1016/j.artmed.2019.101753

Fiske, A., Henningsen, P., & Buyx, A. (2019). Your robot therapist will see you now: Ethical implications of embodied artificial intelligence in psychiatry, psychology, and psychotherapy. *Journal of Medical Internet Research, 21*(5), Article e13216. https://doi.org/10.2196/13216

Fitzpatrick, K. K., Darcy, A., & Vierhile, M. (2017). 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, 4*(2), Article e19. https://doi.org/10.2196/mental.7785

Flückiger, C., Del Re, A. C., Wampold, B. E., & Horvath, A. O. (2018). The alliance in adult psychotherapy: A meta-analytic synthesis. *Psychotherapy, 55*(4), 316–340. https://doi.org/10.1037/pst0000172

Graham, S., Depp, C., Lee, E. E., Nebeker, C., Tu, X., Kim, H.-C., & Jeste, D. V. (2019). Artificial intelligence for mental health and mental illnesses: An overview. *Current Psychiatry Reports, 21*(11), Article 116. https://doi.org/10.1007/s11920-019-1094-0

Guest, G., Bunce, A., & Johnson, L. (2006). How many interviews are enough? An experiment with data saturation and variability. *Field Methods, 18*(1), 59–82. https://doi.org/10.1177/1525822X05279903

Heinz, M. V., Mackin, D. M., Trudeau, B. M., Bhattacharya, S., Wang, Y., Banta, H. A., Jewett, A. D., Salzhauer, A. J., Griffin, T. Z., & Jacobson, N. C. (2025). Randomized trial of a generative AI chatbot for mental health treatment. *NEJM AI, 2*(4). https://doi.org/10.1056/AIoa2400802

Hennink, M. M., Kaiser, B. N., & Marconi, V. C. (2017). Code saturation versus meaning saturation: How many interviews are enough? *Qualitative Health Research, 27*(4), 591–608. https://doi.org/10.1177/1049732316665344

Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence, 1*(9), 389–399. https://doi.org/10.1038/s42256-019-0088-2

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

Luxton, D. D. (2014a). Artificial intelligence in psychological practice: Current and future applications and implications. *Professional Psychology: Research and Practice, 45*(5), 332–339. https://doi.org/10.1037/a0034559

Luxton, D. D. (2014b). Recommendations for the ethical use and design of artificial intelligent care providers. *Artificial Intelligence in Medicine, 62*(1), 1–10. https://doi.org/10.1016/j.artmed.2014.06.004

Malterud, K., Siersma, V. D., & Guassora, A. D. (2016). Sample size in qualitative interview studies: Guided by information power. *Qualitative Health Research, 26*(13), 1753–1760. https://doi.org/10.1177/1049732315617444

Naslund, J. A., Aschbrenner, K. A., Araya, R., Marsch, L. A., Unützer, J., Patel, V., & Bartels, S. J. (2017). Digital technology for treating and preventing mental disorders in low-income and middle-income countries: A narrative review of the literature. *The Lancet Psychiatry, 4*(6), 486–500. https://doi.org/10.1016/S2215-0366(17)30096-2

Norcross, J. C., & Lambert, M. J. (2018). Psychotherapy relationships that work III. *Psychotherapy, 55*(4), 303–315. https://doi.org/10.1037/pst0000193

Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. *Science, 366*(6464), 447–453. https://doi.org/10.1126/science.aax2342

Patton, M. Q. (2015). *Qualitative research and evaluation methods* (4th ed.). SAGE Publications.

Rathod, S., Pinninti, N., Irfan, M., Gorczynski, P., Rathod, P., Gega, L., & Naeem, F. (2017). Mental health service provision in low- and middle-income countries. *Health Services Insights, 10*. https://doi.org/10.1177/1178632917694350

Riaz, M. M. A., Abrar, T., Tebha, S. S., Tariq, H., Fatima, R., Hamad, M., Khan, A. A., Ahadi, A., Ul Sabah, N., Rasool, W., Latif, K., Umar, M., Siddique, I., & Nawaz, F. A. (2024). Psychiatrists' readiness for digital psychiatry in Pakistan: A multicenter cross-sectional study with regression analysis [Conference abstract]. *European Psychiatry, 67*(S1), S548–S549. https://doi.org/10.1192/j.eurpsy.2024.1138

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

Stade, E. C., Stirman, S. W., Ungar, L. H., Boland, C. L., Schwartz, H. A., Yaden, D. B., Sedoc, J., DeRubeis, R. J., Willer, R., & Eichstaedt, J. C. (2024). Large language models could change the future of behavioral healthcare: A proposal for responsible development and evaluation. *npj Mental Health Research, 3*, Article 12. https://doi.org/10.1038/s44184-024-00056-z

Straw, I., & Callison-Burch, C. (2020). Artificial intelligence in mental health and the biases of language based models. *PLOS ONE, 15*(12), Article e0240376. https://doi.org/10.1371/journal.pone.0240376

Topol, E. J. (2019). High-performance medicine: The convergence of human and artificial intelligence. *Nature Medicine, 25*(1), 44–56. https://doi.org/10.1038/s41591-018-0300-7

Torous, J., Bucci, S., Bell, I. H., Kessing, L. V., Faurholt-Jepsen, M., Whelan, P., Carvalho, A. F., Keshavan, M., Linardon, J., & Firth, J. (2021). The growing field of digital psychiatry: Current evidence and the future of apps, social media, chatbots, and virtual reality. *World Psychiatry, 20*(3), 318–335. https://doi.org/10.1002/wps.20883

Tracy, S. J. (2010). Qualitative quality: Eight "big-tent" criteria for excellent qualitative research. *Qualitative Inquiry, 16*(10), 837–851. https://doi.org/10.1177/1077800410383121

Vaidyam, A. N., Wisniewski, H., Halamka, J. D., Kashavan, M. S., & Torous, J. B. (2019). Chatbots and conversational agents in mental health: A review of the psychiatric landscape. *The Canadian Journal of Psychiatry, 64*(7), 456–464. https://doi.org/10.1177/0706743719828977

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly, 27*(3), 425–478. https://doi.org/10.2307/30036540

World Health Organization. (2021). *Ethics and governance of artificial intelligence for health: WHO guidance*. https://www.who.int/publications/i/item/9789240029200