

HUMAN-AI COLLABORATION IN MENTAL HEALTH NURSING: BALANCING EMPATHY AND AUTOMATION

*Jennifer C. Carulla

**Sylvia O. Lopez

***Melody M. Gatdula

****Kieth Joseph E. Valeriano

*****Jennie Jasmin F. Biongan

*****Alvin L. Acosta

*****Elenor C. Santos

Paper Received: 02.11.2025 / **Paper Revised:** 27.11.2025 / **Paper Accepted:** 02.12.2025 / **Paper Published:** 04.12.2025

Corresponding Author: Jennifer C. Carulla; doi:10.46360/cosmos.et.620261001

Abstract

Background: Integration of artificial intelligence (AI) in mental health nursing is accelerating, offering opportunities for improving clinical efficiency, diagnostic precision, and care coordination. However, this technological transformation raises concerns regarding the preservation of human empathy and the nurse-patient therapeutic relationship.

Aim: This research investigates how human-AI collaboration can be optimized in mental health nursing such that it enhances clinical workflows without compromising empathic care.

Methods: We conducted a narrative literature synthesis drawing on recent studies examining AI applications in mental health nursing, nurse perceptions of AI, and empirical evaluations of empathy outcomes in AI-augmented care contexts.

Results: Findings indicate that AI systems can support clinical decision-making, risk assessment, and monitoring while freeing nurses' time for deeper interpersonal engagement. Conversely, AI's limitations in authentic empathy and concerns about depersonalization, ethical considerations, and bias must be systematically addressed.

Conclusion: Effective incorporation of AI in mental health nursing requires a hybrid human-AI model. This model preserves empathic engagement by nurses while leveraging automated tools for routine or cognitively demanding tasks. Education, ethical frameworks, and co-design practices between clinicians and developers are crucial.

Keywords: Human-AI Collaboration; Mental Health Nursing; Empathy; Automation; Therapeutic Relationship; Ethical AI.

Introduction

Mental health nursing is a specialty that demands not only clinical judgment and therapeutic skills but also deep interpersonal connection and empathy. Empathy has been consistently identified as a cornerstone of effective mental health interactions, facilitating trust, emotional safety, and therapeutic outcomes (Yu, 2025). However, the growing complexity of healthcare systems and shortages of qualified professionals have accelerated interest in AI-based tools to enhance care efficiency and accessibility. (Yıldız, 2025; Stones, 2025)

AI technologies-spanning machine learning, natural language processing, and digital phenotyping-hold promise for enhancing assessment, risk detection, and personalized monitoring in mental health contexts (Stones, 2025; Li, 2024). Yet such tools inherently lack the lived experience and emotional authenticity of human clinicians. The challenge lies in integrating AI innovations in a way that strengthens, rather than undermines, the empathic nurse-patient relationship (Ly, 2025).

This research article explores the tension between automation and authentic empathic care, proposing a balanced framework for human-AI collaboration in mental health nursing.

Methods

We conducted a narrative literature synthesis of peer-reviewed journal articles, qualitative studies, and integrative reviews focused on AI applications in mental health nursing and related themes of empathy, automation, and clinician perspectives. The search included databases such as PubMed, CINAHL, PsycINFO, and Scopus with terms like "AI in mental health nursing", "empathy and AI", and "human-AI collaboration in healthcare." Articles published within the last three years were prioritized for relevance and currency.

Data extraction focused on:

1. **Types of AI applications** (e.g., diagnostic tools, conversational agents).

*-*****College of Nursing, Our Lady of Fatima University, Valenzuela City Campus, Philippines.

2. **Clinical outcomes** related to efficiency and risk detection.
3. **Empathy and relational dynamics**, particularly where human clinicians and AI overlapped.
4. **Ethical concerns** (e.g., bias, privacy, depersonalization).

Qualitative data (e.g., nurse attitudes) were thematically analyzed for recurring patterns related to readiness, barriers, and facilitators for AI acceptance.

Results

AI Utilization in Mental Health Nursing

AI technologies are increasingly embedded in mental health settings, supporting a broad range of tasks including early detection of clinical risk factors, automated monitoring, and documentation assistance. Integrative reviews show that AI applications in mental health nursing span machine learning algorithms, natural language processing, digital phenotyping, and conversational agents designed to complement clinical practice rather than replace human clinicians. These tools can be used to identify patterns indicative of symptom escalation, automate time-consuming documentation tasks, and improve overall workflow efficiency.

For instance, machine learning and natural language processing have shown promise in detecting subtle changes in verbal and written communication that may forecast worsening mood or risk of self-harm - information that might otherwise go unnoticed in busy clinical environments (Yinen, 2025; Kampman, 2024). Moreover, by taking on routine or analytical tasks (e.g., risk flagging or alert generation), AI can help nurses allocate more time to direct patient care.

Despite these advances, the literature also emphasizes that technological tools are most effective when integrated with human clinical judgment. Excessive reliance on AI without adequate clinician oversight risks depersonalization and may reduce nurses' attention to contextual details that algorithms cannot capture.

Empathy and Human-Machine Collaboration

A critical finding across studies is that AI systems can augment - but not replace - empathic engagement. Research in human-AI collaboration outside clinical settings indicates that AI support can enhance the quality of empathic interactions when used as a feedback tool. For example, AI-in-the-loop systems providing real-time empathy prompts in peer support contexts improved empathic responses by nearly 20%, particularly for users who initially struggled to respond empathetically.

However, the consensus in the healthcare literature is that AI lacks genuine lived experience and affective empathy, meaning it cannot *authentically* experience emotions or build a therapeutic alliance in the way humans do. This distinction is significant: while AI can mimic empathetic language and support caregivers with prompts or phrasing suggestions, it does not *feel* empathy - an attribute tied closely to human presence and relational depth. Theoretical work also shows that current AI systems struggle with multimodal emotional cues (e.g., tone, facial expression, body language), which limits their ability to achieve the same nuance as human clinicians in empathic engagement.

Taken together, these findings suggest that AI may *support* empathic communication (e.g., by prompting clinicians or improving text-based empathy in constrained settings), but cannot independently sustain the full emotional and relational dimensions central to therapeutic nursing care.

Perceptions of Nurses and Barriers to AI Implementation

Nurses' perceptions of AI integration are mixed, shaped by perceived opportunities as well as fears related to workflow, depersonalization, and professional identity. Qualitative studies of psychiatric nurses indicate that while many recognize AI's potential to support clinical tasks (e.g., decision support and administrative burden reduction), there are persistent concerns that automation could undermine holistic, patient-centered care.

Common barriers include privacy concerns, workflow disruption, and fear of depersonalization. Nurses often view AI tools as potentially beneficial for routine tasks but worry that they might inadvertently shift focus away from patient-clinician interactions. In some qualitative work, nurses also expressed uncertainty about how AI could emotionally support patients, especially in contexts that demand nuanced psychosocial engagement.

Additionally, individual factors such as nurses' emotional regulation skills and attitudes toward change influence how they perceive and adopt technology. High emotional regulation has correlated with greater openness and fewer negative attitudes toward AI integration, suggesting that clinicians' emotional competencies play a role in how comfortably they incorporate AI into practice.

Ethical, Bias, and Trust Considerations

AI systems introduce a suite of ethical challenges, including algorithmic bias, data privacy concerns, accountability ambiguity, and potential erosion of

human-centered care. Systematic reviews of nursing ethics in AI implementation highlight that algorithmic bias and inequity are major risks when training data do not represent diverse populations - a problem that could reinforce existing disparities in mental health care if not carefully addressed.

Patient privacy and data security are also pressing concerns, as AI often relies on large datasets containing sensitive information. Nurses, as patient advocates, must be vigilant in protecting confidentiality and ensuring that consent processes for AI use are transparent and comprehensible.

Liabilities related to accountability and responsibility further complicate ethical implementation. Clinicians fear unclear lines of responsibility when AI decisions lead to adverse outcomes, underscoring the need for transparent governance frameworks that define legal and professional boundaries.

Finally, there is a risk that overreliance on AI may erode clinicians' critical thinking and clinical judgment, especially if automated recommendations are accepted uncritically. Such "automation bias" could diminish nurses' role as autonomous, context-sensitive caregivers.

Discussion

This review reveals both the promise and limitations of AI in mental health nursing, underscoring the need for balanced human-AI collaboration that preserves the distinctively human elements of care while harnessing technological capabilities.

Balancing Automation with Empathy

Empathy is not a peripheral feature of mental health care - it is central to building therapeutic alliance, fostering trust, and enabling patient engagement (Milasan, 2025; Luny, 2024). Authentic empathic engagement encompasses *affective resonance* and *reflective listening*, which enable patients to feel understood and supported. While AI can simulate or suggest empathic phrasing, it does not *experience* emotions and thus cannot fully replicate human-to-human therapeutic presence.

The hybrid models of care emerging in the literature propose that AI be positioned as a support tool that enhances clinicians' capacity to focus on relational work by automating administrative and analytical tasks. When AI reduces cognitive workload and streamlines routine activities, nurses may have greater bandwidth for interpersonal engagement - a dynamic consistent with models of "augmented intelligence" rather than standalone automation.

However, AI's inability to fully interpret multimodal emotional cues or sustain autonomous

empathic engagement limits its role in direct therapeutic interactions. The implication is that AI should be designed to *complement* clinician empathy, not to replace it.

Education and Training

Effective human-AI collaboration requires structured training that extends beyond technical proficiency to include ethical reasoning, data literacy, and critical appraisal of AI outputs. Nursing education must prepare clinicians to interpret algorithmic recommendations prudently, question outputs that conflict with clinical judgment, and understand AI's limitations in emotional and contextual interpretation (Sharma, 2022).

Given that clinicians' attitudes toward AI are linked to emotional regulation and cognitive flexibility, training programs that strengthen emotional intelligence and digital competencies can mitigate resistance and enhance collaborative adoption.⁴

Ethical and Regulatory Frameworks

Robust ethical and regulatory frameworks are essential to govern the integration of AI in mental health nursing. This includes patient data protection protocols, transparent consent processes, algorithm audits to detect and mitigate bias, and clear guidelines on professional accountability (Sharma, 2022; Stones, 2025). Ethical considerations should be embedded in every stage of AI development and implementation, involving clinicians, patients, and community stakeholders to ensure equitable and human-centered outcomes.

Implications for Future Research

Several research gaps remain. Longitudinal studies are needed to assess how human-AI collaboration impacts therapeutic alliance, patient satisfaction, and long-term clinical outcomes. Comparative trials that juxtapose traditional care models with AI-augmented workflows would provide empirical evidence on whether hybrid systems confer measurable benefits without detracting from empathic care.

Future work should also explore co-design methodologies, where nurses and patients actively participate in developing AI tools tailored to specific clinical contexts and relational goals (Hacıaloğlu, Boyraz Şeker, & Kaya, 2025).

Conclusion

Human-AI collaboration in mental health nursing offers transformative potential for enhancing care quality and efficiency. However, the preservation of empathetic, person-centered care remains paramount. Hybrid care models - where AI supports clinical tasks while clinicians maintain primacy in human connection - represent the most promising

pathway forward. Sustained investment in human-centered design, robust ethics frameworks, and educational programs will be essential to strike the appropriate balance between automation and empathy in mental health nursing.

Conflict of Interest

There is no conflict of interest between the authors in this manuscript. Results have been deduced objectively; nothing influenced this result by any individual, monetary, or institutional interests. In effect, utmost endeavors were in order to conduct a bias-free operation of gathering and analyzing the data into interpretation and producing a study based on merit, away from research inconsistencies.

References

1. Yıldız, E. (2025). Artificial Intelligence in Mental Health Nursing: Balancing Clinical Efficiency and the Human Touch - A Quest for a New Synthesis. *J Psychiatr Ment Health Nurs*.
2. Yu, I. (2025). Integrative review on AI in mental health nursing practice. *Int J Mental Health Nurs*.
3. Stones, A.Y. (2025). Nursing perspectives on AI in psychiatric care. *PubMed*.
4. Li, W. & et al. (2025). AI meets empathy: mental health nurses' views on digital tool integration. *J Mental Health Nurs*, 2(2).
5. Luny, C. (2024). Ethical considerations in AI-driven therapy. *JMIR Mental Health*.
6. Singh, A.K. (2025). AI applications in mental health care: diagnosis and intervention. *Psychological Medicine*.
7. Li, W., Chen, R., Zhang, M. (2025). Artificial intelligence meets empathy: Integrating digital tools in mental health nursing practice. *J Ment Health Nurs*.
8. Ly, Y. (2024). The Exploration of Psychiatric Nurses' Perspectives on the Applications of Artificial Intelligence in Supporting Care. *PubMed*.
9. Yinen, Y. (2025). Ethical and Institutional Readiness for Artificial Intelligence in Nursing: An Umbrella Review. *PubMed*.
10. Hacıaloğlu, N., Boyraz Şeker, E. & Kaya, F. (2025). Nurses' attitudes towards artificial intelligence: relationship between cognitive flexibility and emotion regulation. *BMC Psychol*, 13, 1121.
11. Mohammed, SAAQ, Osman, YMM, Ibrahim, AM, Shaban, M. (2025). Ethical and regulatory considerations in the use of AI and machine learning in nursing: A systematic review. *Int Nurs Rev.*, 72(1), e70010. doi: 10.1111/inr.70010.
12. Rubin, M., Arnon, H., Huppert, J.D., Perry, A. (2024). Considering the Role of Human Empathy in AI-Driven Therapy. *JMIR Ment Health*, 11, e56529.
13. Milasan, L.H., Scott-Purdy, D. (2025). The Future of Artificial Intelligence in Mental Health Nursing Practice: An Integrative Review. *Int J Ment Health Nurs*, 34(1) e70003. doi: 10.1111/inm.70003.
14. Sharma, A. & et. al. (2022). Human-AI Collaboration Enables More Empathic Conversations in Text-based Peer-to-Peer Mental Health Support. *arXiv*.
15. Kampman, O. P., Phang, Y. S., Han, S., Xing, M., Hong, X., Hoosainsah, H., ... & Morris, R. J. (2024). A multi-agent dual dialogue system to support mental health care providers. *arXiv preprint arXiv:2411.18429*.
16. Hassanein, S, El Arab, RA, Abdrbo, A, Abu-Mahfouz, MS, Gaballah, MKF, Seweid, MM, Almari, M and Alzghoul, H (2025). Artificial intelligence in nursing: an integrative review of clinical and operational impacts. *Front. Digit. Health*, 7, 1552372. doi:10.3389/fdgth.2025.1552372