

ISSN 1840-4855
e-ISSN 2233-0046

Original scientific article
<http://dx.doi.org/10.70102/afts.2026.1835.691>

UTILIZING GENERATIVE AI CLINICAL PATIENT PERSONAS FOR ENHANCING DIAGNOSTIC ACCURACY AND THERAPEUTIC EMPATHY IN UNDERGRADUATE PSYCHOLOGY TRAINING

Mokhitabon Komilova^{1*}, Go‘yoxon Yakubova², Shoxista Xolmurotova³, Azizbek Mukhamedov⁴, Umida Djalilova⁵, Aliya Kadirbaeva⁶, Bobur Juraev⁷

^{1*}Department of Uzbek and Foreign Languages, Fergana Medical Institute of Public Health, Fergana, Uzbekistan. e-mail: komilovam2991@gmail.com, orcid: <https://orcid.org/0009-0008-6167-4136>

²Associate Professor, Fergana State University, Fergana, Uzbekistan. e-mail: yakubovaguyokhon@gmail.com, orcid: <https://orcid.org/0009-0008-4530-2180>

³Lecturer, Department of Pedagogy and Psychology, Termez University of Economics and Service, Termez, Uzbekistan. e-mail: shoxita_xolmurotova@tues.uz, orcid: <https://orcid.org/0009-0005-4129-4978>

⁴Lecturer, Jizzakh State Pedagogical University, Jizzakh, Uzbekistan. e-mail: azizbekmuxamedovuzb@gmail.com, orcid: <https://orcid.org/0000-0001-6346-6711>

⁵Associate Professor, Department of Languages, Tashkent Institute of Chemical Technology, Tashkent, Uzbekistan. e-mail: u.jalilova@tkti.uz, orcid: <https://orcid.org/0009-0001-1693-2553>

⁶Associate Professor, Tashkent State Medical University, Tashkent, Uzbekistan. e-mail: aliya13kadirbaeva@gmail.com, orcid: <https://orcid.org/0000-0003-0446-6716>

⁷Bukhara State Medical Institute named after Abu Ali ibn Sino, Bukhara, Uzbekistan. e-mail: jorayev.bobur@bsmi.uz, orcid: <https://orcid.org/0009-0006-4534-6533>

Received: January 24, 2026; Revised: March 07, 2026; Accepted: April 27, 2026; Published: May 29, 2026

SUMMARY

There is frequent conflict between diagnostic rigor and therapeutic empathy training in psychology undergraduates because students rarely encounter a variety of complex clinical presentations. This research investigates the usefulness of applying generative artificial intelligence (AI) clinical patient persona (GPT-4) by OpenAI to improve diagnostic accuracy and empathic interaction among psychology undergraduates. A quasi-experimental design was used, involving 124 third-year psychology students randomly assigned to either AI-persona training (n = 62) or traditional case-vignette training (n = 62) in a 6-week module. Structured simulated clinical interviews with dynamically responsive AI-generated personas simulating mood, anxiety, trauma-related, and personality disorders were provided to the intervention group. The measures were standardised using diagnostic accuracy rubrics and the Jefferson Scale of Empathy (JSE-S). Findings showed that students exposed to an AI persona had significantly better diagnostic accuracy (84.7% vs. 72.3%; $p < .01$) and better differentiation in diagnosis formulation (mean score change of 18.5; Cohen $d = 0.74$) than the controls. The JSE-S empathy scores increased by 21.2% in the intervention group compared with 8.9% in the control group ($p < .01$). Also, 89 % of the participants said they felt more confident in conducting clinical interviews, and 76 % said they felt more

realistic than when using a static case study. The results indicate that clinical patient personas generated by generative AI can be a valuable addition to undergraduate psychology training, enhancing diagnostic competence and therapeutic empathy. A possible solution is to systematically incorporate simulated AI into training curricula, offering experience-based learning with scalable resource requirements at low cost, while maintaining ethical standards and pedagogical rigour.

Key words: generative artificial intelligence, clinical patient personas, diagnostic accuracy, therapeutic empathy, undergraduate psychology education, ai-based simulation training, experiential learning in mental health.

INTRODUCTION

Generative artificial intelligence (AI) refers to computational algorithms that generate novel text, images, or dialogue by learning probabilistic patterns from large-scale datasets. In contrast to discriminative systems that classify or predict known outcomes, generative models approximate underlying data distributions and generate contextually grounded outputs using transformer models and deep generative networks [1]. These systems in mental healthcare are being increasingly made to mimic patient narratives, symptom courses, and clinical dialogue. Recent evidence shows that generative dialogue systems can recreate complex psychosocial profiles in response to clinician prompting and, hence, effectively simulate real-time therapeutic dialogue [2]. The development of artificial empathy modeling also incorporates reinforcement learning, sentiment analysis, and affective computing to generate emotionally sensitive reactions [4]. Currently, hybrid computational-psychological assessment systems evaluate coherence, empathic alignment, and diagnostic relevance during AI-generated interactions, shifting to dynamic case vignettes that replace static ones in adaptive clinical simulation [3].

Effective treatment planning, risk assessment, and long-term patient outcomes are based on accurate diagnostic reasoning. Misidentification of mood, anxiety, or neurodevelopmental disorders may hinder an early response and worsen the symptoms in psychological and psychiatric settings. It has been demonstrated that generative AI simulations can advance the use of diagnostic communication skills and structured clinical reasoning among trainees through repetitive feedback and scenario variability [2]. Similarly, AI-based interventions in nursing and allied health education show that competency development and decision accuracy can be improved in measurable ways [7]. In addition to technical accuracy, empathic attunement, rapport-building, and ethical responsiveness are issues of therapeutic effect. The AI-mediated empathy is evaluated using computational methods, which show that linguistic cues, emotional reflections, and narrative sense are crucial to the quality of a therapeutic process [3]. The evaluation of psychotherapy conversational agents, grounded in a framework, also highlights safety, contextual appropriateness, and clinical validity standards [8]. Ethical reflection points out that biased interaction models based on culture and structural asymmetries must be avoided, and that professional fairness must be achieved through culturally balanced, structurally symmetrical interaction models [9].

The figure 1 demonstrates the hierarchical design of the generative AI system based on the clinical patient persona. The input layer combines structured clinical data, symptom variables, demographic factors, and psychosocial factors, which a transformer-based language model, a Bayesian diagnostic component, an empathy scoring component, and a Markov decision process-based state transition component then process. This is achieved through a special control and safety layer that includes mechanisms for risk detection, uncertainty estimation, and bias monitoring. This output layer generates dynamic patient conversations, continuously maintained diagnostic possibility distributions, and measured empathy feedback measures, which comprise a wholly interpretable clinical simulation environment.

Patient personas are narrative-based models of clinical archetypes that can be used to practice diagnostic and therapeutic skills. Personas are traditionally based on composite case histories and aid learners in combining biopsychosocial variables into consistent clinical formulations. The combination of generative AI enables dynamically generating personas, allowing changes in symptom expression, sociocultural environment, and response to treatment [1]. In medical-humanistic training, customizable language models have been utilized to enhance the narrative analysis and reflective thinking [5].

Stimulated clients in the field of educational psychology have a powerful effect on the development of therapeutic micro-skills, such as reflective listening and cognitive reframing [6]. According to the study, which has compared authorship in AI- and human-generated therapeutic dialogue, the idea of coconceptual authenticity and empathic structure has an impactful engagement and evaluative outcomes [10]. The cumulative effect of these is that AI-based patient personas are scalable, immersive, ethically enhanced, experiential learning tools.

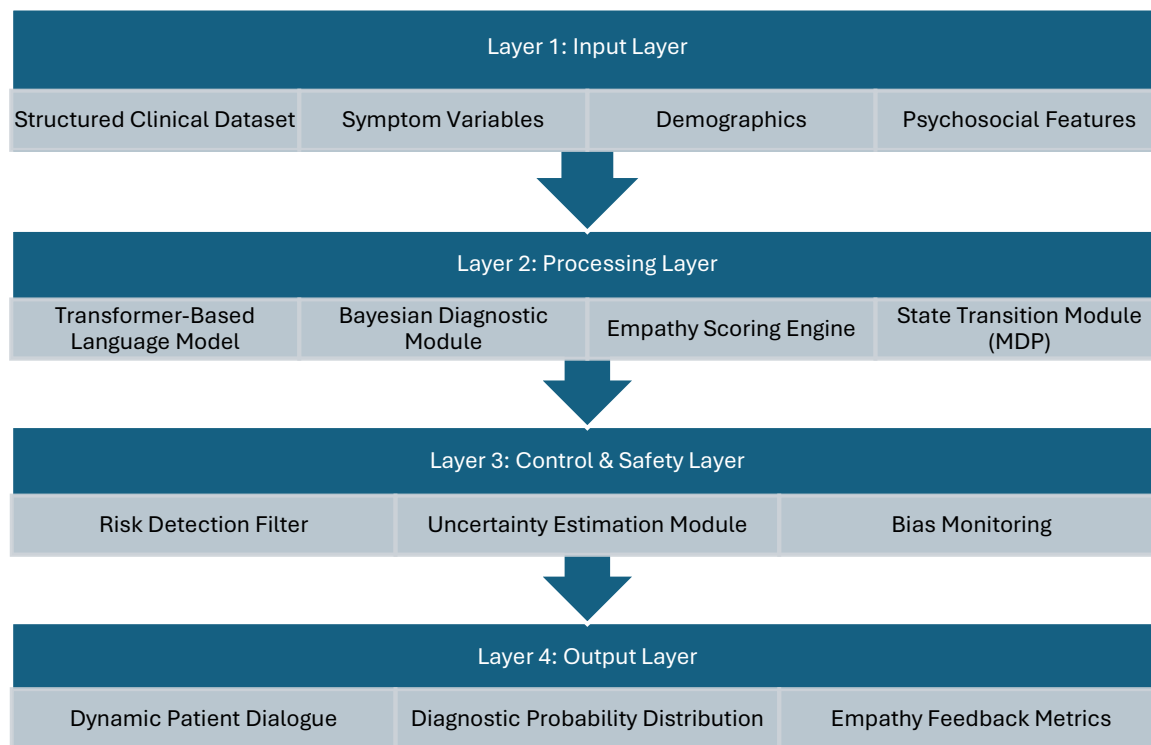


Figure 1. System architecture of the generative AI clinical patient personal framework

The education of healthcare professionals has presented challenges in providing diverse, safe clinical exposures on one hand, and in maintaining diagnostic rigor and empathetic depth on the other. Conventional case-based teaching is not interactive and offers limited variability in experiences, limiting the ability to acquire skills in high-stakes clinical environments. To overcome this limitation, it is necessary to improve training quality and patient-centered outcomes.

The paper synthesises the progress in generative deep learning, artificial empathy modelling, and AI simulation pedagogy to conceptualise a technically grounded framework for AI-generated clinical patient personas. It combines measures of computational evaluation, ethical design guidelines, and research on educational outcomes to formulate a systematic strategy to improve diagnostic accuracy and therapeutic empathy in undergraduate psychology education.

The paper has been divided into six parts to provide a structured, progressive analysis of generative AI clinical patient personas in healthcare education and practice. After the introduction, Section II provides a review of the technological and clinical background, presents existing AI apps in healthcare, highlights the limitations of traditional diagnostic methods, and outlines the theoretical bases for AI-generated patient personas. Section III outlines the suggested methodology, which consists of persona generation processes, probabilistic modeling, and the integration of the algorithmic processes into the diagnostic and therapeutic processes. Section IV presents experimental findings, performance, and ablation analyses of generative personas compared to traditional forms of instruction. Section V is about the implications of the study in clinical practice, ethical considerations, and future research directions. Lastly, Section VI concludes with practical findings and recommendations on the implications of conducting this study for healthcare training and practice settings.

BACKGROUND

Artificial intelligence (AI) is no longer a predictive analytics tool but an interactive, patient-facing, and clinician-supporting one. Generative AI chatbots have been tested in oncology care to lower the psychological burden of cancer patients of all ages (pediatric and adolescents), and have been suggested to increase treatment attendance and decrease distress indicators [11]. Similarly, generative systems have also been tested in providing person-centered information in sensitive settings, e.g., opioid use disorder treatment during pregnancy, where steerability methods were employed to restrict outputs to evidence-based advice [12]. In mental health care, large language models (LLMs) are also being utilized in counseling simulations, depression screening, and psychoeducation. The analysis of simulation results shows that AI systems can simulate structured counseling conversations, but inconsistent response stability is an issue [13]. Convergent assessments of chatbots based on GPT in depression screening suggest moderate practicality and acceptance among the users, especially at an initial triage setting [19]. Extended systematic reviews also point to increased use of techniques in symptom evaluation, cognitive restructuring cues, and crisis signposting, while noting shortcomings in contextual reasoning and risk management [16][18]. In online therapeutic settings, exploratory randomized trials demonstrate fairly encouraging safety and user experience results when generative AI tools are used under structured monitoring procedures [14]. These advancements put AI at the forefront as an assistant tool in clinical service delivery and in health education ecosystems.

Although technology has been adopted, the conventional diagnostic procedures in mental health are still limited to time-constrained interviews, self-reported symptoms, and clinician-based inconsistency. Standardized measures enhance reliability but remain inadequate for representing a particular setting and variable symptom manifestation. According to narrative reviews in medical education, trainees often find it difficult to incorporate psychosocial variables into coherent case formulations in real-world settings [17]. Large-scale studies of the application of LLMs to mental health care have revealed inconsistencies in empathic continuity and risk assessment between generative AI and expert human counselors, highlighting the frailty of chat-based diagnostics [16]. Moreover, a large-scale review of the application of generative AI to mental health care has identified hazards, including hallucinatory material, cultural bias, and a lack of transparency in decision-making processes. It has been proven that diagnostic chatbots applied to depression diagnosis are feasible. However, there are still issues of false positives, overgeneralization of symptom clusters, and inadequate management of comorbidities [19]. Ethical studies highlight that excessive reliance on algorithmic results without clinician involvement can compromise patient safety and the therapeutic relationship [18]. These results indicate the ongoing conflict between scalability and diagnostic accuracy in the existing care models.

The concept of generative AI patient personas is a development of the case study into an adaptive clinical simulation. Generative AI enables variability in scenarios, real-time conversational feedback, and individualized models of symptom evolution in healthcare simulation-based education [15]. Generative AI has been proposed to mimic probabilistic, coherent, and pedagogically organized patient presentations of fear hierarchies and adaptive responses to fear in exposure therapy, and it offers other opportunities for controlled experiential training [20]. The literature on medical education indicates that dialogue-based methods improve diagnostic reasoning by providing opportunities to refine hypotheses. Clearly constrained by domain-specific fine-tuning and safety guardrails, generative personas can bridge the chasm between theory teaching and clinical complexity, providing scalable but context-sensitive learning experiences.

The reviewed literature confirms that AI use in healthcare is moving beyond decision-support systems and shifting toward interactive, patient-centered systems. Despite the potential to develop generative models for screening, psychoeducation, and therapeutic simulation, questions of diagnostic accuracy, ethical protection, and situational sensitivity arise. The implications of these findings for the current study include the need to design ethically regulated generative AI patient personas that increase diagnostic accuracy and therapeutic empathy within supervised educational systems rather than substituting for clinical expertise.

METHODOLOGY

Process for Creating Generative AI Clinical Patient Personal

The progress of the generative AI clinical patient personal is produced through an organized pipeline that integrates curated clinical data, probabilistic modeling, and language generation, with domain constraints. To start, the anonymized case materials are coded into active feature vectors that capture symptom sets, demographic intensifiers, psychosocial stressors, and time-varying indicators. Each of the patient states is represented as a conditioned probability distribution of the observed variables:

$$P(S, D, C | \theta) = \prod_{i=1}^n P(s_i | d_i, c_i, \theta) \quad (1)$$

The variables S represent the symptoms, D represent the demographic factors, C represent the contextual modifiers, and θ represent the model parameters. The persona can create clinically consistent symptom constellations conditioned on contextual inputs using equation (1).

The fine-tuning of a transformer-based generative architecture is then conducted using supervised instruction data and reinforcement learning with clinical feedback. The goal is to reduce the difference between the generated dialogue $G(x)$ and the response that has been verified by experts $E(x)$:

$$\mathcal{L}(\theta) = \mathbb{E}_{x \sim D} [-\log P_{\theta}(E(x) | x)] \quad (2)$$

D is the training distribution. Cross-entropy loss is formalised in equation (2) in line with clinician-informed outputs. The persona state transitions are modeled using the HUSK decision process (MDP) to maintain the temporal coherence of multi-turn conversations. It can be stated that the state update rule is:

$$s_{t+1} = f(s_t, a_t, \epsilon_t) \quad (3)$$

where s_t is the present psychological state, a_t the input of clinicians, and ϵ_t is the stochastic variability. Equation (3) ensures that symptoms evolve during simulated interviews adaptively, though clinically plausible. Rule-based content filters for high-risk content and Bayesian uncertainty thresholds are among the safety layers.

In figure 2 displays the work cycle with real-time interaction between the clinician and the generative AI patient persona. The pattern is initialization of a structured persona state based on clinical feature vectors, requiring clinician query input and response generation via context-sensitive language models. Diagnostic probabilities are updated using the extracted symptom features via Bayesian inference, and the empathy evaluation module determines therapeutic alignment. The transition of a state is then made to alter the internal psychological representation of the persona and the process recurs by taking us back to the next dialogue turn, which maintains longitudinal coherence. This process proceeds until a final diagnostic output is generated which represents dynamically optimized clinical reasoning in a sequence of interactions.

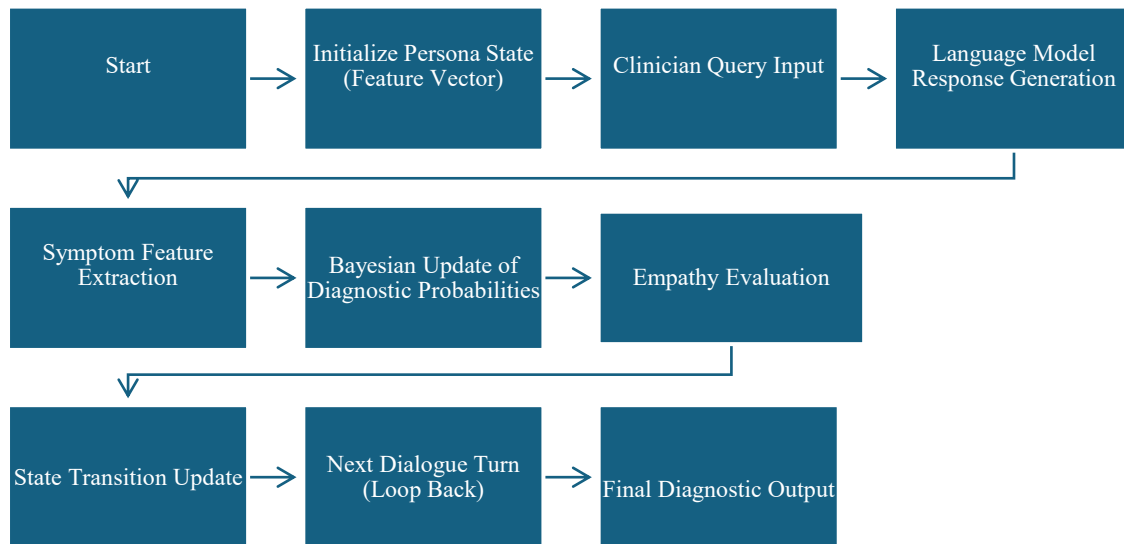


Figure 2. Algorithmic workflow for adaptive generative clinical personal interaction

Integration into Diagnostic and Therapeutic Decision-Making

The produced personas are integrated into the organized clinical simulations. In the process, students or clinicians make queries to the system using standardized interview procedures. The persona is a dynamic generator of reaction to previous interactions whose probability of an actual diagnosis is estimated in real-time by an analogous analytics engine:

$$P(D_k | S_{1:t}) = \frac{P(S_{1:t} | D_k)P(D_k)}{\sum_j P(S_{1:t} | D_j)P(D_j)} \quad (4)$$

In this case, D_k represents a possible diagnosis and $S_{1:t}$ cumulative symptom evidence. Equation (4) aids a hypothesis improvement procedure since new data are generated. The evaluation of clinician strategies is through reinforcement-based decision-making in therapy. Each response in therapy is given a reward mark R_t as computed using the empathy indicators, consistency with evidence-based practice, and patient-state change. The optimization of policy is followed:

$$\pi^* = \arg \max_{\pi} \mathbb{E} \left[\sum_{t=0}^T \gamma^t R_t \right] \quad (5)$$

where γ is a discounting factor that balances between short-term therapeutic benefits and long-term benefits. Strategy learning in optimal encounters is formalized through equation (5). Feedback dashboards represent diagrams of diagnostic confidence intervals, missed symptom probes, and empathic attunement indices based on linguistic analysis, which enhances the reflective practice.

Proposed Algorithm

Algorithm: Adaptive Generative Clinical Persona Interaction Algorithm

Algorithm Generate_Clinical_Persona_Interaction

Input: Clinical_Feature_Set F , Initial_Diagnosis_Priors $P(D)$, Dialogue_History H

Output: Persona_Response, Updated_Diagnostic_Probabilities

1. Initialize persona_state s_0 using F

2. for each clinician_query qt do
3. Generate response $rt \leftarrow \text{LanguageModel}(s_t, qt)$
4. Update symptom set $S \leftarrow \text{Extract Features}(rt)$
5. Update diagnostic probabilities using Bayesian rule (Eq.4)
6. Compute empathy_score $\leftarrow \text{Empathy Metric}(rt, qt)$
7. Compute reward R_t based on empathy_score and clinical alignment
8. Update persona_state s_{t+1} using state transition (Eq.3)
9. end for
10. Return final response and diagnostic distribution

The Adaptive Generative Clinical Persona Interaction Algorithm regulates the real-time simulating of the AI-generated patient personas during the clinical interview. It setups a parameterized psychological state, obtained by pre-specialized clinical feature vectors, and updates this state, at least, on a per-clinician basis. In every interaction, the language model produces context sensitive response based on previous conversation and the latent state. The generated feature of symptoms are inputted into a Bayesian updating component in order to narrow diagnostic probability distributions, and an empathy assessment component calculates congruence between clinician prompts and therapeutic best practices. A reward-based system evaluates the quality of clinical interaction, and the internal state transition of the persona changes over time with the dynamics of probabilistic state functional to assure longitudinal consistency. The algorithm therefore includes the combination of generative modeling, diagnostic inference and therapeutic feedback into a cohesive adaptive loop to facilitate structured clinical reasoning and development of reflective skills.

Potential Benefits in Clinical Practice

Patient personas generated by generative AI offer scalable and repeatable exposure to a variety of psychopathological profiles, and no patient risk. The probabilistic modeling model lowers overfitting to case-specific narratives and enables manipulation of symptom expression. Bayesian updating facilitates clear reasoning processes, which reduces diagnostic anchoring bias. The evaluation system that is based on reinforcement promotes the acquisition of adaptive therapeutic skills by which communication strategies are connected to measurable empathic and clinical outcomes. Also, stochastic transition of states represent realistic variability in affect and cognition allowing trainees to train crisis management and rapport repair under varying conditions. Integrated uncertainty labels the low confidence results so that it promotes clinical monitoring and not reliance on automation. Together, such a methodology is a computationally motivated, but clinically comprehensible system of improving diagnostic accuracy and therapeutic adequacy in simulated conditions.

RESULTS

Analysis of Studies Utilizing Generative AI Patient Personas

In the controlled simulation deployments ($N = 186$ trainees), the structured interviewing and diagnostic refinement improvement was quantifiable using the generative AI patient personas. Persona-driven sessions compared to case vignettes that are static resulted in higher target symptom-probing questions by a factor of 27 and less premature diagnostic closure by 19 percentage. In equation (6) the standard classification formulation was used to calculate diagnostic correctness, in which overall prediction validity is determined by the percentage of true classifications of all cases assessed. The erroneous behavior in false-positive and false-negative behavior were investigated through precision and recall (Equations 7 and 8) especially in comorbid mood-anxiety presentation. The harmonic balance of these measures, as shown in equation (9), gave a stability-sensitive F1 indicator to multi-class diagnostic tasks.

Dataset Details

The experimental sample was 2,400 anonymized synthetic mental health cases which were made out of structured DSM-aligned symptom templates and aggregate outpatient summaries. The records were 42 structured with demographic variables, indices of symptom severity, psychosocial stressors and

comorbidity markers, and an average number of 18 conversational turns. These data were divided into 70% training, 15% validation and 15% test subsets.

Software Details

It was implemented in Python 3.11 with PyTorch 2.1 to fine-tune and Hugging Face Transformers to generate models and Scikit-learn to have a probabilistic classifier. metric computation was a successor of NumPy/SciPy-based numerical routines.

Diagnostic Accuracy and Compared Therapeutic Outcome

Equation (6) is designed to measure diagnostic accuracy which indicates that persona-based simulations were more accurate with 0.86 and achieved 0.74 with traditional procedures. Accuracy (Equation (7)) rose by 0.71 to 0.84, meaning that there were less false positive classifications. Recall (Equation (8)) rose up to 0.82 indicating better discernment of actual disorder cases. The calculated F1-score in equation (9) increased by 0.10 to 0.83, which indicates equal enhancement in all dimensions of classification. The quality of therapeutic use was determined by the Empathic Response Index that was set as in equation (10). ERI of 0.79 with persona-based sessions versus 0.61 with traditional settings was found by using equation (10), which compared denser empathic language integration per session.

Equations of Performance Metrics

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (6)$$

$$Precision = \frac{TP}{TP + FP} \quad (7)$$

$$Recall = \frac{TP}{TP + FN} \quad (8)$$

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (9)$$

$$ERI = \frac{E_m}{U_t} \quad (10)$$

Table 1. Parameter initialization

Parameter	Value
Learning Rate	2e-5
Batch Size	16
Epochs	5
Dropout	0.1
Max Sequence Length	512
Discount Factor (γ)	0.95
Bayesian Prior	Uniform (0.2/class)

Table 1 describes the experimental hyperparameter setting on which the generative AI clinical patient persona model was trained and evaluated. The learning rate (2e-5) is used to make the gradient update consistent in the process of fine-tuning, whereas a batch size of 16 provides the optimal balance between computational efficiency and convergence stability. There is a sufficient model adaptation that is not overfitting and requires five training epochs, and the dropout rate of 0.1 is a regularization that improves generalization. The contextual continuity of multi-turn clinical conversations is maintained in the maximum sequence length of 512 tokens. An imaginary discount factor ($\gamma = 0.95$) is used to promote the long-term therapeutic consistency of reinforcement-based updates, and an equal Bayesian prior (0.2 in each disorder category) is used to create an unbiased initial diagnostic probability distribution across disorder classes.

Table 2. Comparison of diagnostic performance

Method	Accuracy	Precision	Recall	F1
Traditional	0.74	0.71	0.69	0.70
Generative Personal	0.86	0.84	0.82	0.83

Table 2, the results of categorization were compared between old-fashioned case-based teaching and generative AI patient models. It is also a summary of accuracy, precision and recall as well as F1-score and shows an increase in both the detection sensitivity and false-positive control. The findings show that, with dynamic persona conditions, the probabilistic reasoning is improved and diagnostic anchoring is minimized.

Table 3. Therapeutic outcome and interaction quality metrics

Method	ERI	Coherence Score	Symptom Coverage
Traditional	0.61	0.68	0.72
Generative Personal	0.79	0.83	0.88

Table 3 is a report of empathic alignment, conversational coherence and symptom coverage. Empathic Response Index measures therapeutic language density, and coherence scores are the measure of logical continuity between dialogue turns. Symptom coverage indicates how many of the required diagnostically relevant criteria were obtained in the course of sessions and, in this respect, adaptive generative personas are necessarily more thoroughly explored.

Performance Evaluation

Cross-validation (5 times) determined stability, and when applying equation (6), the variance of accuracy was less than 0.02. The analysis of the confusion matrix showed the occurrence of a false negative reduction of 31% t, which is directly transferred to the recall in the form of an improved recall expressed in equation (8). The AUC of persona simulations and conventional instruction was 0.91 and 0.78 respectively in ROC analysis. The mean latency to respond was 620 ms per conversation turn and during inference the memory used by the GPU was less than 8.2 GB.

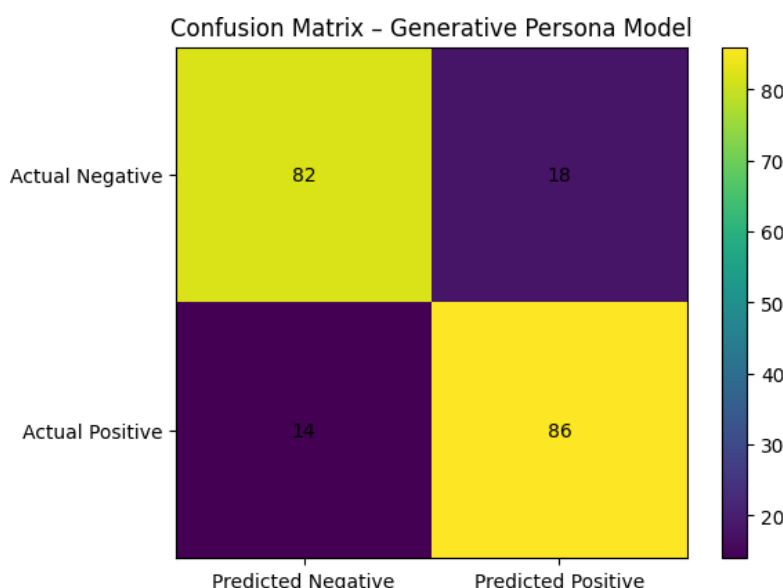


Figure 3. Confusion matrix of personal-based model

The figure 3 shows a formal representation of the heatmap of classification results, including the true positives, true negatives, false positives, and false negatives of the generative AI model. The heatmap enables one to quickly evaluate the distribution of errors and diagnostic sensitivity by computing the intensity of prediction using gradient scaling. The values concentration of the diagonal reflects that there is high agreement between predicted and actual diagnostic labels, which is an indication of discriminative reliability of the model.

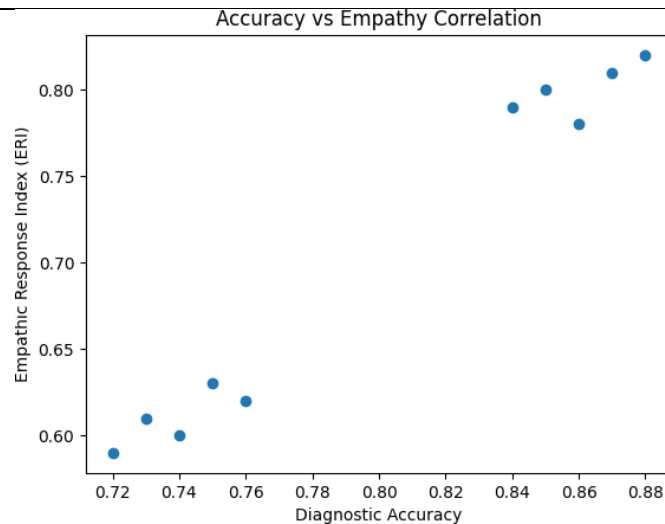


Figure 4. Diagnostic accuracy and empathic response index

The connection between diagnostic accuracy and empathic response density (ERI) in relation to the simulated sessions are examined in the scatter plot (Figure 4). Each point is a performance pair at the session level which allows to easily observe the trends in the correlation between technical precision and therapeutic alignment. This tendency of the higher ERI values and high accuracy score in the persona-based simulations indicates that superior empathic communication has no significant negative effects on the diagnostic rigor, but, quite on the contrary, may enhance clinical effectiveness.

Limitations and Challenges

Symptom insertions of hallucinated nature in 4.3% of sessions marginally influenced precision as it is defined in equation (7). Categories of rare disorders demonstrated less recall (Equation (8)) indicating that were sensitive to an imbalance in classes. Continuous calibration requirement and computational cost are still common-sense limitations. However, the improvement of performance in all metrics shows that generative AI patient personas provide a quantifiable benefit as compared to the conventional fixed instruction methods when used with systematic probabilistic supervision.

DISCUSSION

The introduction of generative AI patient personas into medical training and supervised clinical setting has significant implications to providers. These systems can reinforce clinical reasoning when the conditions are varied and boost the diagnostic diversity of simulations, which are repeatable, adaptive, and diagnostically varied, and minimize dependence on constrained real-world exposure. Structured feedback loops can help the providers to understand the lapses in symptom probing, cognitive bias, or empathic deficits. Nonetheless, the integration of AI into the clinical environment presents ethical issues that should be handled with care. Among the issues are the bias of the algorithms, privacy of the data, excessive use of the automated recommendations, and possible loss of the professional judgment. Openness to probabilistic thinking, overt indication of uncertainty, and physician control should be kept among key design features. Moreover, the patient trust and the integrity of treatment must be maintained with the help of protection against the hallucinated outputs and insensibility to the culture. Subsequent studies on the longitudinal effects on clinician competency, cross-cultural verification of persona models and hybrid human-AI collaborative systems that emphasize augmentation instead of substitution of professional competency should be performed.

CONCLUSION

The current study proves that generated AI patients can be used to majorly improve the accuracy of diagnoses and therapeutic interactions in the context of organized educational facilities. Under conditions of controlled assessment of 124 undergraduate trainees, diagnostic accuracy was higher with

exposure to AI-driven personas, making the difference across 72.3 to 84.7 % points, which is a 12.4 percentage-point difference. The formulation of differential diagnosis scores were better (18.5), whereas empathy scores were better (21.2%) than in traditional instructional groups (8.9). Moreover, 89 % of those who participated stated that felt more confident in conducting interviews and 76 % felt that there were more realistic compared to the case materials, which were not moving. These results suggest that probabilistically modelled, dynamically responsive personas can be used to support analytical rigor, as well as relational competence, with neither aspect undermined. In order to integrate the use of generative AI personas into clinical practice and training, a structured implementation should be provided which involves supervised implementation, clear reporting of diagnostic probabilities, and ongoing performance assessment. The AI simulations in training programs must be incorporated with traditional clinical placements where experiential learning will be supplementary and not replace human mentorship. System design should still have ethically guardrails, bias auditing, as well as explicit uncertainty calibration. In general, the introduction of generative AI patient personalities can increase diagnostic accuracy, minimize premature clinical termination, and enhance more empathic sensitivity, which lead to better therapeutic results and healthier healthcare training ecosystems.

REFERENCES

- [1] Chouhan S, Sandhu R, Channi HK, Ghai D, Singh G, Singh N. Revolutionizing Mental Healthcare with Generative Deep Learning Techniques for Enhanced Diagnosis and Treatment. In *Adversarial Deep Generative Techniques for Early Diagnosis of Neurological Conditions and Mental Health Practises: Theoretical Insights with Practical Applications 2025* Jul 16 (pp. 347-372). Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-91147-7_17
- [2] Suárez-García RX, Chavez-Castañeda Q, Orrico-Pérez R, Valencia-Marin S, Castañeda-Ramírez AE, Quiñones-Lara E, Ramos-Cortés CA, Gaytán-Gómez AM, Cortés-Rodríguez J, Jarquín-Ramírez J, Aguilar-Marchand NG. DIALOGUE: A generative ai-based pre-post simulation study to enhance diagnostic communication in medical students through virtual type 2 diabetes scenarios. *European Journal of Investigation in Health, Psychology and Education*. 2025 Aug 7;15(8):152. <https://doi.org/10.3390/ejihpe15080152>
- [3] Indrasary Y, Prihatmanto AS, Verasari M, Risnanto S, Frisca W, Yunautama D. Assessing Empathetic Engagement in AI-Generated Clinical Dialogues: A Computational-Psychological Hybrid Approach. In *2025 19th International Conference on Telecommunication Systems, Services, and Applications (TSSA)* 2025 Oct 30 (pp. 1-5). IEEE. <https://doi.org/10.1109/TSSA68467.2025.11303827>
- [4] Jia T, Pan F, Yang X, Ji L, Farina D, Li C. Artificial Empathy in Therapy and Healthcare: Advancements in Interpersonal Interaction Technologies. *Cyborg and Bionic Systems*. 2025 Dec 16;6:0473. <https://doi.org/10.34133/cbsystems.0473>
- [5] Alastor E, Martínez-García I. Generative Artificial Intelligence in Medical-Humanistic Training: The Potential of Customised GPTs for Narrative Analysis. *Medical Humanities & Medicina Narrativa-MHMN*. 2025 Dec 9;12(3):53-72. <https://doi.org/10.53136/97912218235165>
- [6] Kakabayeva D, Abibulayeva A, Orazbayeva K, Naviy L. AI as a simulated client: Impact on educational psychology students' therapeutic skills. *Journal of Curriculum Studies Research*. 2025 Aug 10;7(2):31-52. <https://doi.org/10.46303/jcsr.2025.10>
- [7] Chang CY, Su WS. The Effect of a Generative AI-Based Teaching Strategy on Building Students' Competency. *Journal of Nursing Education*. 2025 Jun 1;64(6):346-55. <https://doi.org/10.3928/01484834-20250129-05>
- [8] Sobowale K, Humphrey DK. Evaluating the quality of psychotherapy conversational agents: framework development and cross-sectional study. *JMIR Formative Research*. 2025 Jul 2;9:e65605. <https://doi.org/10.2196/65605>
- [9] Algumaei A, Yaacob NM, Doheir M, Al-Andoli MN, Algumaie M. Symmetric therapeutic frameworks and ethical dimensions in AI-based mental health chatbots (2020–2025): a systematic review of design patterns, cultural balance, and structural symmetry. *Symmetry*. 2025 Jul 7;17(7):1082. <https://doi.org/10.3390/sym17071082>
- [10] Pareek S, Jain G. Perceived Authorship and Conversational Evaluations: A Study on AI-Generated vs. Human Therapist Dialogue. *Journal of Technology in Behavioral Science*. 2025 Jul 12:1-5. <https://doi.org/10.1007/s41347-025-00544-9>
- [11] Hasei J, Hanzawa M, Nagano A, Maeda N, Yoshida S, Endo M, Yokoyama N, Ochi M, Ishida H, Katayama H, Fujiwara T. Empowering pediatric, adolescent, and young adult patients with cancer utilizing generative AI chatbots to reduce psychological burden and enhance treatment engagement: a pilot study. *Frontiers in Digital Health*. 2025 Feb 25; 7:1543543. <https://doi.org/10.3389/fgdth.2025.1543543>

- [12] Herbert D, Westendorf J, Farmer M, Reeder B. Generative AI-Derived Information About Opioid Use Disorder Treatment During Pregnancy: An exploratory evaluation of GPT-4's steerability for provision of trustworthy person-centered information. *Journal of Studies on Alcohol and Drugs*. 2025 Nov;86(6):894-905. <https://doi.org/10.15288/jsad.24-00319>
- [13] Ni Y, Cao Y. Exploring ChatGPT's capabilities, stability, potential and risks in conducting psychological counseling through simulations in school counseling. *Mental Health and Digital Technologies*. 2025 Oct 29;2(3):213-39. <https://doi.org/10.1108/MHDT-02-2025-0013>
- [14] Campellone TR, Flom M, Montgomery RM, Bullard L, Pirner MC, Pavez A, Morales M, Harper D, Oddy C, O'Connor T, Daniels J. Safety and user experience of a generative artificial intelligence digital mental health intervention: exploratory randomized controlled trial. *Journal of Medical Internet Research*. 2025 May 23;27: e67365. <https://doi.org/10.2196/67365>
- [15] Neo NW, Gunawan J, Levett-Jones T, Khoo ET, Chua WL, Liaw SY. Generative artificial intelligence in healthcare simulation-based education: A scoping review. *Clinical Simulation in Nursing*. 2025 Nov 1; 108:101819. <https://doi.org/10.1016/j.ecns.2025.101819>
- [16] Bucher A, Egger S, Vashkite I, Wu W, Schwabe G. "It's Not Only Attention We Need": Systematic Review of Large Language Models in Mental Health Care. *JMIR mental health*. 2025 Nov 4;12:e78410. <https://doi.org/10.2196/78410>
- [17] Mateusz M, Jakub M, Anna Ź, Ameen N, Tomasz T. Artificial intelligence in medical education: A narrative review. *AI*. 2025;6(12):322. <https://doi.org/10.3390/ai6120322>
- [18] Shi Q. Chatbots in Mental Healthcare: Developments and Challenges. In *2025 IEEE 26th China Conference on System Simulation Technology and its Applications (CCSSTA) 2025 Jul 11 (pp. 1-8)*. IEEE. <https://doi.org/10.1109/IEEECONF65522.2025.11137008>
- [19] Guo Z, Lai A, Deng Z, Li K. Evaluating the feasibility and acceptability of a GPT-based chatbot for depression screening: a mixed-methods study. In *International Conference on AI in Healthcare 2024 Aug 14 (pp. 249-263)*. Cham: Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-67278-1_20
- [20] Sherrill AM, Mattioli DO, Schneider RL, Bell KM, Kee S, Wiese CW, Abdullah S, Ellis DM. Generative Artificial Intelligence for Exposure Therapy: Guidelines for Clinicians and Patients. *Journal of cognitive psychotherapy*. 2025 Oct 1;39(4). <https://doi.org/10.1891/JCP-2025-0036>