Asia Pacific Science Press

Asia Pacific Journal of Clinical Medical Research
ISSN(O):3079-8345
ISSN(P):3079-8337
Vol. 2 No.5 (2026)

A Study on Multimodal Data- and Explainable AI-Based Risk Prediction and Intervention for Postpartum Depression in Nursing Care Settings

Yangyang Lin*

School of Smart Health Care, Zhejiang Dongfang Polytechnic, Wenzhou, Zhejiang, 325000, China

Corresponding author: Yangyang Lin, lu0101402@163.com

Copyright: 2026 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY-NC 4.0), permitting distribution and reproduction in any medium, provided the original author and source are credited, and explicitly prohibiting its use for commercial purposes.

Abstract: Objective: To evaluate a nursing-centered dynamic prediction and risk-stratified intervention system for postpartum depression (PPD) using multimodal longitudinal data and explainable artificial intelligence (AI). **Methods:** A quasi-experimental controlled study included 300 postpartum women enrolled within 6 hours after delivery and allocated by admission sequence to AI-assisted stratified nursing (n=150) or routine nursing care (n=150). Physiological, behavioral, psychological, and social-support indicators were collected, with EPDS assessments on postpartum days 7, 14, 28, and 42. A hybrid model combined temporal convolution, long short-term memory, text and voice representations, and multilayer fusion; SHAP and attention mechanisms supported case-level explanations. Outcomes included time to risk identification, EPDS trajectory, remission among high-risk women, nursing workload, and internally validated model performance. **Results:** Risk was identified earlier in the intervention group than in the control group (1.21 +/- 0.68 vs. 4.73 +/- 1.92 days, P<0.05). Mean EPDS reduction by day 42 was greater (-5.62 +/- 3.14 vs. -2.13 +/- 2.87, P<0.05), and remission was more frequent (78.3% vs. 42.6%; chi-square=36.41, P<0.05). High-risk identification reached 92.0% versus 64.0%. Time per follow-up decreased by 34.7%, and nurses reported fewer women requiring daily follow-up and lower work-related stress. The model achieved an AUC of 0.864, accuracy of 0.82, recall of 0.77, and F1-score of 0.79. **Conclusion:** A multimodal explainable AI system integrated with risk-stratified nursing care was associated with earlier risk recognition, improved depressive symptom trajectories, and lower nursing workload. External validation and a fully randomized multicenter evaluation are required before broad implementation.

Keywords: Postpartum Depression; Multimodal Data; Explainable Artificial Intelligence; Nursing Decision Support; Risk Stratification

Published: Oct 20, 2026

DOI: https://doi.org/10.62177/apjcmr.v2i5.1788

1. Introduction

1.1 Background and Clinical Importance

Postpartum depression (PPD) is a common and clinically important mental health condition after childbirth. Its effects extend beyond maternal mood and functioning to breastfeeding, mother-infant bonding, infant development, family relationships, and healthcare use. Meta-analytic evidence indicates that depressive symptoms affect a substantial proportion of postpartum women, although estimates vary by measurement method, timing, and population [1,2]. The World Health Organization has

therefore recommended that perinatal mental health be integrated into routine maternal and child health services [3].

The Edinburgh Postnatal Depression Scale (EPDS) is the most widely used screening instrument in postpartum care [4]. It is brief and useful for detecting women who may require further assessment, but it is not a diagnostic interview and a single score provides only a time-specific snapshot [5]. Postpartum risk changes with sleep disruption, feeding difficulties, pain, maternal-infant interaction, family support, and stressful events. Fixed-point screening may therefore overlook women whose symptoms emerge or escalate between scheduled visits.

Machine-learning studies have shown that clinical and psychosocial variables can support PPD risk prediction [6-9]. Wearable and mobile data may add information about sleep, activity, autonomic function, and daily rhythms [10]. However, models based on one-time questionnaires or limited predictor sets may not capture dynamic risk. Moreover, a clinically useful system must do more than maximize discrimination: it must communicate why risk has changed, fit the nursing workflow, and link the prediction to safe and proportionate care [11-14].

This study developed a nursing-centered framework that integrates multimodal longitudinal data, explainable AI, and risk-stratified intervention. The central premise was that nurses can use dynamic risk estimates as decision support while retaining clinical judgment, direct assessment, and referral responsibility.

1.2 Study Objective and Hypotheses

The objective was to evaluate whether a multimodal explainable AI system embedded in postpartum nursing care could identify emerging PPD risk earlier, improve depressive symptom outcomes, and reduce nursing workload compared with routine care. We hypothesized that the intervention would shorten the time to risk identification, produce a greater reduction in EPDS scores, increase remission among high-risk women, and demonstrate acceptable internal prediction performance.

2. Materials and Methods

2.1 Study Design and Setting

A prospective quasi-experimental controlled study was conducted at Wenzhou People's Hospital, Zhejiang, China, from January to September 2025. Participants were allocated to the intervention or control group according to a predefined admission sequence. The intervention group received AI-assisted dynamic risk prediction and stratified nursing care, whereas the control group received routine postpartum nursing, scheduled visits, standard EPDS screening, health education, supportive communication, and referral based on conventional clinical judgment.

2.2 Participants

The study included 300 women enrolled within 6 hours after delivery, with 150 participants in each group. Inclusion criteria were: age 18-45 years; ability to complete electronic assessments and follow-up; willingness to use the required data-collection devices; and written informed consent. Exclusion criteria were a history of severe psychiatric illness such as schizophrenia or bipolar disorder, a newborn requiring neonatal intensive care, acute severe depression requiring immediate psychiatric intervention, or missing three or more scheduled follow-up assessments.

The two groups were comparable in the available baseline continuous characteristics, including maternal age, body mass index, and gestational age (all P>0.05; Table 1). Categorical baseline variables were not included in the present analysis because the available records were incomplete.

Table 1. Available baseline characteristics of the study groups

Characteristic	AI-assisted nursing (n=150)	Routine care (n=150)	Statistic	P value
Age (years)	26.24 +/- 3.01	27.03 +/- 2.98	t=1.865	0.064
Maternal BMI (kg/m2)	24.30 +/- 2.05	23.99 +/- 2.09	t=1.059	0.291
Gestational age (weeks)	39.12 +/- 2.01	38.99 +/- 1.98	t=0.461	0.646

2.3 Multimodal Nursing Data

The data architecture included four domains. Physiological data comprised heart-rate variability (RMSSD and SDNN), sleep duration, sleep efficiency, nocturnal awakenings, activity, and resting heart rate. Behavioral data included mother-

infant interaction frequency, activity rhythm, feeding difficulties, breast discomfort, and other nursing-sensitive events. Psychological data included weekly EPDS scores, Beck Anxiety Inventory scores, and emotional features extracted from optional text or voice inputs. Social-support data included partner support, family participation in care, and changes in communication frequency.

2.4 Data Preprocessing and Multimodal Fusion

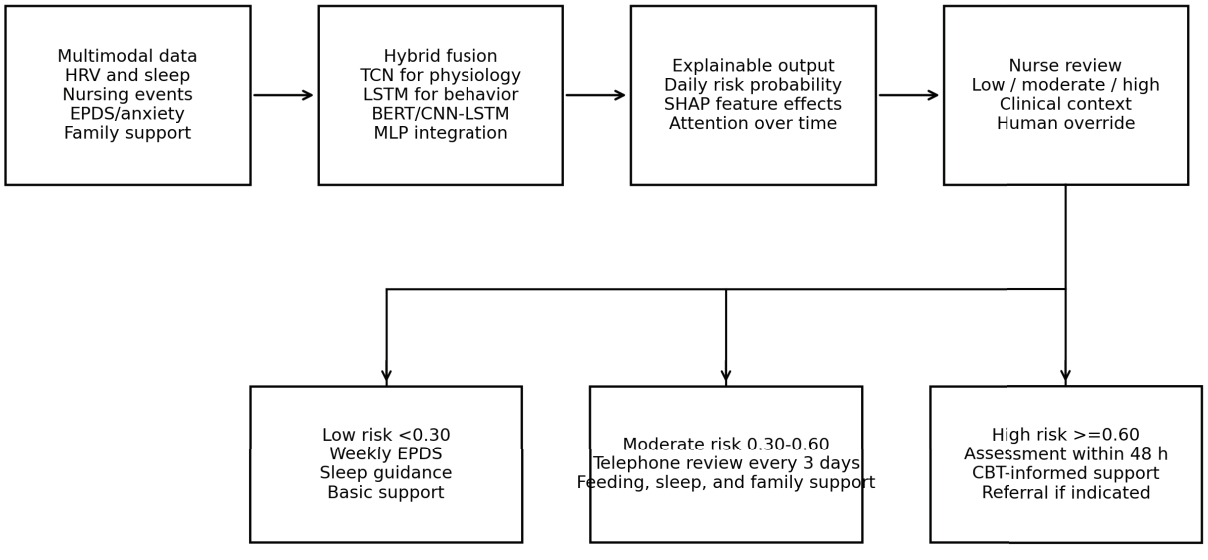
Continuous missing values were addressed through multiple imputation by chained equations, categorical variables through mode imputation, and variables with more than 30% missingness were excluded. A nursing expert panel selected 48 clinically relevant features: 12 physiological, 18 behavioral or nursing-event, 10 psychological, and 8 social-support features. Continuous variables were standardized, Spearman correlations were reviewed, and 24-hour and 72-hour rolling-window summaries were derived for time-varying measures.

A hybrid early- and late-fusion strategy was used. Physiological time series were modeled with temporal convolutional components, behavioral sequences with long short-term memory networks, text features with a BERT-based representation, and voice features with a CNN-LSTM module. Submodel outputs and social-support variables were integrated in a multilayer perceptron to produce a daily probability of clinically relevant PPD risk. The data were divided into training, validation, and test sets in a 70:15:15 ratio. Model optimization used Adam, a batch size of 32, and early stopping.

2.5 Explainability and Nursing Intervention

SHAP values were used to identify the contribution of individual features, while attention weights highlighted time points and nursing events associated with changes in risk. The output was designed as a clinical decision aid rather than an autonomous diagnostic system. Nurses reviewed the score, explanation, recent EPDS results, and clinical context before selecting or escalating care.

Figure 1. Nursing-centered multimodal prediction, explainability, and risk-stratified intervention framework.



The model generated a daily risk score from 0 to 1. Low risk (<0.30) triggered weekly EPDS monitoring, sleep guidance, and basic psychological support. Moderate risk (0.30-0.60) triggered nurse telephone follow-up every 3 days, joint sleep-emotion monitoring, prompt management of feeding difficulties, and strategies to strengthen family support. High risk (>=0.60) triggered psychological assessment within 48 hours, CBT-informed nursing support, safety evaluation, and psychiatric consultation when indicated. Any concern about self-harm or severe deterioration overrode the algorithmic threshold and required immediate clinical assessment.

2.6 Outcome Measures

The primary clinical outcomes were the change in EPDS score from postpartum day 7 to day 42 and remission, defined in the

study as EPDS<10 at the final follow-up. Other outcomes were time from risk emergence to identification, the proportion of high-risk women identified, time per follow-up, number of women requiring daily follow-up, and nurse self-rated work stress on a 0-10 scale. Prediction performance was summarized by area under the receiver operating characteristic curve (AUC), accuracy, recall, and F1-score.

2.7 Statistical Analysis

Analyses were conducted using SPSS version 26.0. Continuous variables were summarized as mean +/- standard deviation and compared using independent-samples t tests where appropriate. Categorical outcomes were compared with chi-square tests. Repeated EPDS measurements were examined descriptively across the four follow-up points, and between-group differences in change were assessed at the prespecified endpoint. Statistical significance was set at two-sided P<0.05. Prediction metrics were calculated on the held-out test subset. Because exact confidence intervals and full repeated-measures model outputs were not available in the archived analysis output, the results are reported using the prespecified summary statistics without adding unobserved estimates.

2.8 Ethical Considerations

Written informed consent was obtained from participants. Data were de-identified for analysis, and access to multimodal records was restricted to authorized study personnel. The name of the Approving Ethics Committee: Ethics Committee of Wenzhou People's Hospital. Official Ethics Approval Number: KY-202507-028.

3. Results

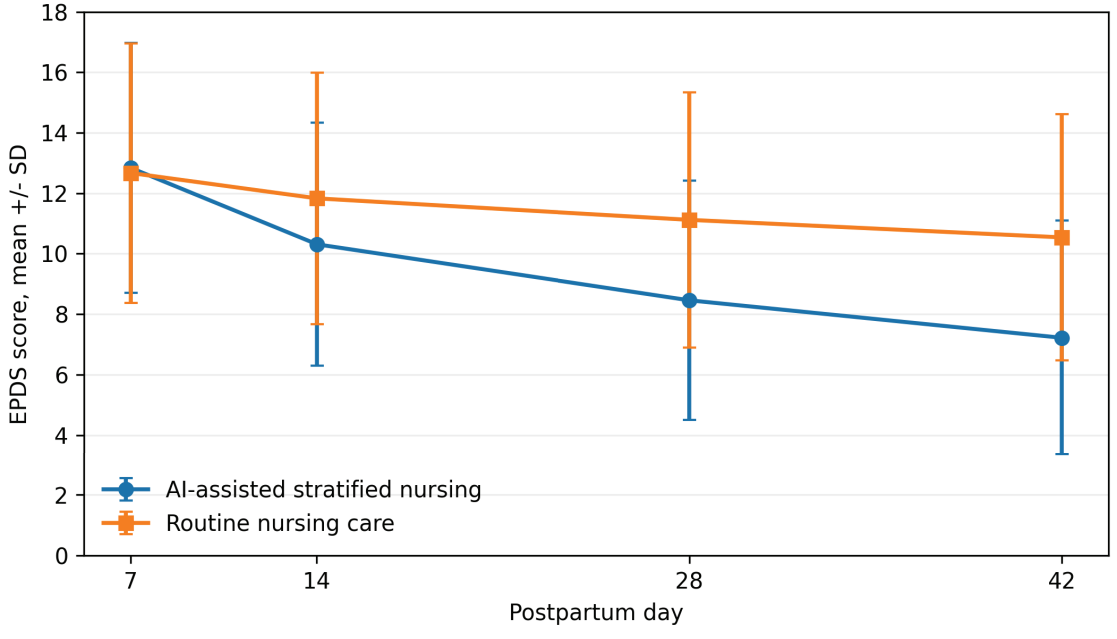
3.1 EPDS Trajectories

EPDS scores were similar at postpartum day 7 (12.84 +/- 4.13 in the intervention group and 12.67 +/- 4.29 in the control group). Scores then declined more rapidly under AI-assisted stratified nursing. By day 42, the mean scores were 7.22 +/- 3.87 and 10.54 +/- 4.07, respectively. The mean reduction was -5.62 +/- 3.14 in the intervention group compared with -2.13 +/- 2.87 in the control group (P<0.05; Table 2 and Figure 2).

Table 2. EPDS scores across postpartum follow-up

Group	Day 7	Day 14	Day 28	Day 42	Change, day 42-day 7
AI-assisted nursing (n=150)	12.84 +/- 4.13	10.31 +/- 4.02	8.46 +/- 3.95	7.22 +/- 3.87	-5.62 +/- 3.14
Routine care (n=150)	12.67 +/- 4.29	11.83 +/- 4.16	11.12 +/- 4.22	10.54 +/- 4.07	-2.13 +/- 2.87

Figure 2. EPDS trajectories at postpartum days 7, 14, 28, and 42. Error bars represent standard deviations.



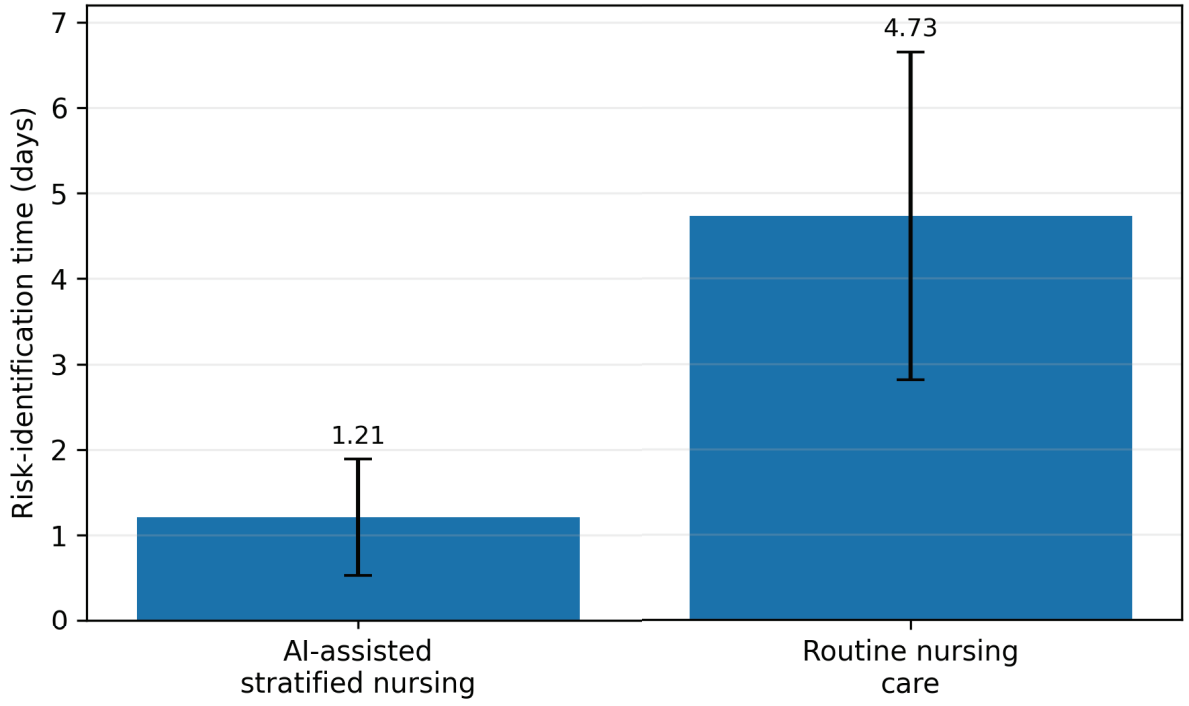
3.2 Remission and Timeliness of Risk Identification

At the final follow-up, remission among women classified as high risk was 78.3% in the intervention group and 42.6% in the control group (chi-square=36.41, P<0.05). The mean time to risk identification was 1.21 +/- 0.68 days with AI-assisted monitoring and 4.73 +/- 1.92 days with routine care (P<0.05). The high-risk identification rate was 92.0% versus 64.0% (P<0.05; Table 3 and Figure 3).

Table 3. Risk identification and remission outcomes

Outcome	AI-assisted nursing	Routine care	P value
High-risk remission at day 42	78.3%	42.6%	<0.05
Risk-identification time (days)	1.21 +/- 0.68	4.73 +/- 1.92	<0.05
High-risk identification rate	92.0%	64.0%	<0.05

Figure 3. Time from risk emergence to identification in the two care groups.



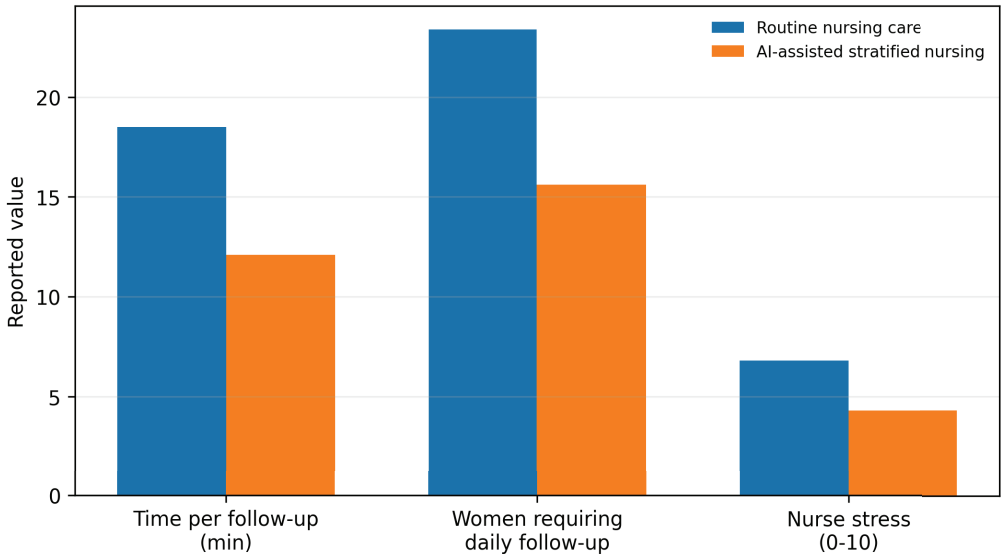
3.3 Nursing Workload

AI-assisted risk stratification reduced the mean time required for one follow-up from 18.5 +/- 6.2 minutes to 12.1 +/- 4.7 minutes, corresponding to a 34.7% reduction. The number of women requiring daily follow-up was lower (15.6 +/- 6.3 vs. 23.4 +/- 7.8), and nurse self-rated work stress was also lower (4.3 +/- 1.6 vs. 6.8 +/- 1.9; all P<0.05; Table 4 and Figure 4).

Table 4. Nursing workload indicators

Indicator	Routine care	AI-assisted nursing	P value
Time per follow-up (min)	18.5 +/- 6.2	12.1 +/- 4.7	<0.05
Relative reduction in follow-up time	-	34.7%	<0.05
Women requiring daily follow-up	23.4 +/- 7.8	15.6 +/- 6.3	<0.05
Nurse stress score (0-10)	6.8 +/- 1.9	4.3 +/- 1.6	<0.05

Figure 4. Nursing workload and self-rated stress under routine and AI-assisted care.



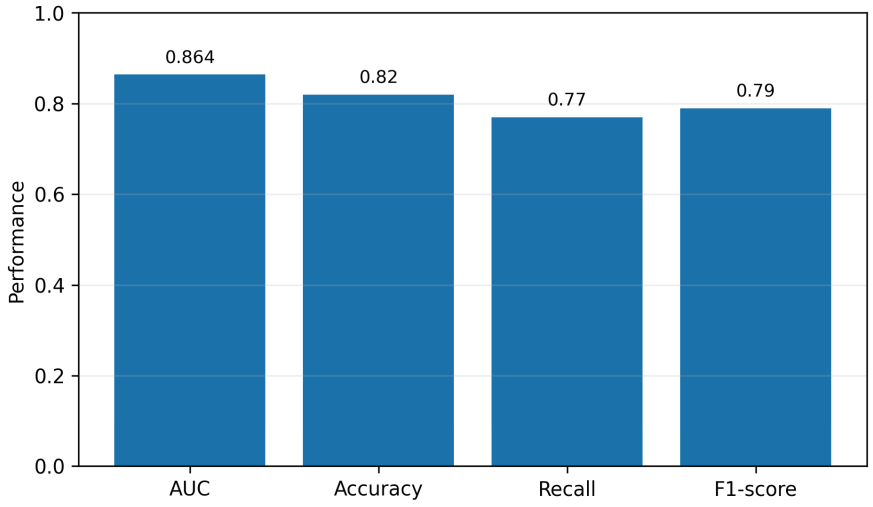
3.4 Prediction Performance and Explainable Outputs

The multimodal model achieved an AUC of 0.864, accuracy of 0.82, recall of 0.77, and F1-score of 0.79 on internal testing (Table 5 and Figure 5). Case-level explanation displays commonly highlighted changes in nocturnal awakening, HRV, negative emotional language, mother-infant interaction, and partner support. These outputs allowed nurses to review the likely drivers of a risk alert before initiating or escalating care.

Table 5. Internal prediction performance

AUC	Accuracy	Recall	F1-score
0.864	0.82	0.77	0.79

Figure 5. Summary of internally validated prediction performance metrics.



4. Discussion

4.1 Principal Findings

This study found that a nursing-centered multimodal and explainable AI system was associated with earlier identification of postpartum depression risk, larger improvement in EPDS scores, higher remission among high-risk women, and reduced nursing workload. The model showed useful internal discrimination and balanced recall and F1 performance. Importantly, the intervention was not limited to generating a probability: the score was linked to a predefined stepped-care pathway and reviewed by nurses in context.

4.2 Contribution of Multimodal Dynamic Data

The findings support the use of complementary modalities rather than reliance on a single questionnaire. EPDS captures self-reported symptoms at a specific time, whereas sleep, HRV, activity, nursing-sensitive events, language, and social support may change between formal assessments. Multimodal learning is intended to exploit such complementary signals [11]. In this study, rolling temporal features and nursing-event data helped the system respond to patterns that static screening may miss. The observed reduction in time to risk identification is clinically relevant because timely psychosocial support and targeted follow-up can reduce the occurrence or persistence of postpartum depressive symptoms [15].

4.3 Explainability and Nursing Decision Support

Clinical adoption of AI depends on more than numerical performance. Clinicians need explanations that relate model output to recognizable patient factors and support appropriate action [12]. The SHAP and attention displays in this study translated risk estimates into plausible nursing concerns, such as worsening sleep fragmentation, reduced mother-infant interaction, and declining family support. This design is consistent with the principle that AI should augment, rather than replace, professional judgment [16,17]. It also aligns with current reporting guidance emphasizing transparent model development, evaluation, and clinical context [13].

4.4 Implications for Nursing Practice

The intervention may improve continuity between inpatient maternity care and post-discharge follow-up. Previous work has linked machine-learning-based PPD risk assessment to the evaluation of preventive interventions [18]. Low-risk women can receive routine education without unnecessary intensive contact, moderate-risk women can receive structured telephone support and rapid management of feeding or sleep problems, and high-risk women can be prioritized for timely psychological assessment and referral. This targeted allocation may explain the lower follow-up burden and work stress observed among nurses. The approach also operationalizes the WHO recommendation to integrate perinatal mental health assessment and response into maternal and child health services [3].

4.5 Strengths and Limitations

Strengths include the use of repeated postpartum assessments, integration of nursing-sensitive and digital indicators, explicit explainability methods, and evaluation of both patient outcomes and nursing workload. The intervention also connected prediction to a practical risk-stratified pathway rather than treating model accuracy as the sole endpoint.

Several limitations require caution. First, allocation by admission sequence was not fully randomized and may have introduced selection or temporal bias. Second, the study was conducted at one hospital, and the model received only internal split-sample validation; the test subset was small for a high-dimensional architecture. Third, EPDS is a screening measure rather than a diagnostic interview, and the study did not report clinician-confirmed PPD diagnoses. Fourth, multimodal data may be missing non-randomly when women are distressed or unable to use devices. Fifth, the available categorical baseline records were internally inconsistent and require correction from the original case report forms. Finally, exact ethics approval information, confidence intervals, calibration measures, and decision-curve analysis should be added before submission to a peer-reviewed clinical journal. External validation, calibration assessment, and a multicenter randomized trial are needed.

5. Conclusion

A multimodal explainable AI system embedded in a nursing-led, risk-stratified care pathway was associated with earlier recognition of postpartum depression risk, greater improvement in EPDS scores, higher high-risk remission, and lower nursing workload than routine postpartum care. The system provides a feasible model for connecting longitudinal digital assessment with actionable nursing follow-up while preserving human oversight. The findings should be regarded as preliminary until the model is externally validated and the intervention is tested in a rigorously randomized multicenter study.

Funding

This study was supported by the 2025 Institutional Research Project of Zhejiang Dongfang Polytechnic (Project No. DF2025AIGC39).

Conflict of Interests

The authors declare that there is no conflict of interest regarding the publication of this paper.

Reference

[1] Shorey, S., Chee, C. Y. I., Ng, E. D., Chan, Y. H., Tam, W. W. S., & Chong, Y. S. (2018). Prevalence and incidence of postpartum depression among healthy mothers: A systematic review and meta-analysis. Journal of Psychiatric Research, 104, 235–248. https://doi.org/10.1016/j.jpsychires.2018.08.001

[2] Meltzer-Brody, S., Howard, L. M., Bergink, V., et al. (2018). Postpartum psychiatric disorders. Nature Reviews Disease Primers, 4, Article 18022. https://doi.org/10.1038/nrdp.2018.22

[3] World Health Organization. (2022). Guide for integration of perinatal mental health in maternal and child health services. World Health Organization.

[4] Cox, J. L., Holden, J. M., & Sagovsky, R. (1987). Detection of postnatal depression: Development of the 10-item Edinburgh Postnatal Depression Scale. British Journal of Psychiatry, 150, 782–786. https://doi.org/10.1192/bjp.150.6.782

[5] Levis, B., Negeri, Z., Sun, Y., et al. (2020). Accuracy of the Edinburgh Postnatal Depression Scale for screening to detect major depression among pregnant and postpartum women: Systematic review and meta-analysis of individual participant data. BMJ, 371, Article m4022. https://doi.org/10.1136/bmj.m4022

[6] Qi, W., Wang, Y., Li, C., et al. (2023). Predictive models for predicting the risk of maternal postpartum depression: A systematic review and evaluation. Journal of Affective Disorders, 333, 107–120. https://doi.org/10.1016/j.jad.2023.04.026

[7] Shin, D., Lee, K. J., Adeluwa, T., & Hur, J. (2020). Machine learning-based predictive modeling of postpartum depression. Journal of Clinical Medicine, 9(9), Article 2899. https://doi.org/10.3390/jcm9092899

[8] Lin, Y., & Zhou, D. (2025). A method for predicting postpartum depression via an ensemble neural network model. Frontiers in Public Health, 13, Article 1571522. https://doi.org/10.3389/fpubh.2025.1571522

[9] Lilhore, U. K., Dalal, S., Varshney, N., et al. (2024). Prevalence and risk factors analysis of postpartum depression at early stage using a hybrid deep learning model. Scientific Reports, 14, Article 4533. https://doi.org/10.1038/s41598-024-54927-8

[10] Hurwitz, E., Butzin-Dozier, Z., Master, H. T., et al. (2024). Harnessing consumer wearable digital biomarkers for individualized recognition of postpartum depression using the All of Us Research Program data set: Cross-sectional study. JMIR mHealth and uHealth, 12, Article e54622. https://doi.org/10.2196/54622

[11] Baltrusaitis, T., Ahuja, C., & Morency, L. P. (2019). Multimodal machine learning: A survey and taxonomy. IEEE Transactions on Pattern Analysis and Machine Intelligence, 41(2), 423–443. https://doi.org/10.1109/TPAMI.2018.2798607

[12] Tonekaboni, S., Joshi, S., McCradden, M. D., & Goldenberg, A. (2019). What clinicians want: Contextualizing explainable machine learning for clinical end use. Proceedings of Machine Learning Research, 106, 359–380.

[13] Collins, G. S., Moons, K. G. M., Dhiman, P., et al. (2024). TRIPOD+AI statement: Updated guidance for reporting clinical prediction models that use regression or machine learning methods. BMJ, 385, Article e078378. https://doi.org/10.1136/bmj-2023-078378

[14] Desai, P. M., Harkins, S., Rahman, S., et al. (2024). Visualizing machine learning-based predictions of postpartum depression risk for lay audiences. Journal of the American Medical Informatics Association, 31, 289–297. https://doi.org/10.1093/jamia/ocad198

[15] Dennis, C. L., & Dowswell, T. (2013). Psychosocial and psychological interventions for preventing postpartum depression. Cochrane Database of Systematic Reviews, (2). https://doi.org/10.1002/14651858.CD001134.pub3

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

[17] Wiens, J., Saria, S., Sendak, M., et al. (2019). Do no harm: A roadmap for responsible machine learning for health care. Nature Medicine, 25, 1337–1340. https://doi.org/10.1038/s41591-019-0548-6

[18] Liu, H., Dai, A., Zhou, Z., et al. (2023). An optimization for postpartum depression risk assessment and preventive intervention strategy based machine learning approaches. Journal of Affective Disorders, 328, 163–174. https://doi.org/10.1016/j.jad.2023.02.028