

Effectiveness of an Integrated Health Education and Prenatal Yoga-Relaxation Program on Stress and Sleep Quality Among Pregnant Women: A Quasi-Experimental Repeated-Measures Study

Abstract

Background: Stress and poor sleep quality are common concerns during pregnancy and may adversely affect maternal well-being. An integrated approach combining health education with prenatal yoga-relaxation may provide a nonpharmacological strategy to address both problems. To analyze the effectiveness of integrated health education and prenatal yoga-relaxation on stress levels and sleep quality among pregnant women. **Methods:** This controlled quasi-experimental study with repeated measures involved 60 pregnant women in an intervention group (n = 30) and a control group (n = 30). The intervention group received health education and prenatal yoga-relaxation for 10 sessions over 5 weeks, while the control group received routine antenatal care. Stress was assessed using the DASS stress subscale and sleep quality using the PSQI at baseline (T0), after the fifth session (T1), and after the tenth session (T2). Data were analyzed using repeated-measures ANOVA or the Friedman test, as appropriate. **Results:** In the intervention group, mean stress scores decreased from 18.63 (T0) to 17.00 (T1) and 11.27 (T2), compared with 19.47, 20.07, and 19.70 in the control group. Mean PSQI scores decreased from 7.77 to 6.47 and 4.73 in the intervention group, compared with 8.57, 7.97, and 7.67 in the control group. Significant changes were also observed in perceived benefits, self-efficacy, activity-related affect, and commitment, with greater improvements in the intervention group. **Conclusion:** Integrated health education and prenatal yoga-relaxation progressively reduced stress and improved sleep quality among pregnant women over 10 sessions.

Keywords: Health education; prenatal yoga-relaxation; pregnant women; repeated measures; stress levels; sleep quality

Introduction

Pregnancy involves complex physiological, psychological, and social changes. Hormonal changes, physical discomfort, concerns about fetal health, preparation for childbirth, changes in family roles, and socioeconomic demands may increase pregnant women's vulnerability to stress. To a certain extent, stress represents an adaptive response. However, persistent stress may increase activation of the hypothalamic-pituitary-adrenal axis and the sympathetic nervous system, thereby affecting cortisol regulation, inflammatory responses, autonomic balance, and maternal psychological well-being. A meta-analysis showed that stress during pregnancy was associated with an increased risk of preterm birth, low birth weight, and small-for-gestational-age infants [1]. Sleep disturbance is a common maternal problem but does not always receive adequate attention in antenatal care. Increased urinary frequency, back and pelvic pain, fetal movement, difficulty finding a comfortable sleeping position, and pregnancy-related concerns may interfere with sleep duration, continuity, and efficiency. A meta-analysis of 24 studies involving 11,002 pregnant women found that approximately 45.7% of pregnant women experienced poor sleep quality. Sleep quality also tended to deteriorate with advancing gestational age, particularly from the second to the third trimester [2]. Another meta-analysis reported that the prevalence of poor sleep quality reached 44.5% during the perinatal period, indicating that sleep problems constitute an important component of the maternal health burden [3].

Stress and sleep disturbances have a bidirectional relationship. Increased psychological tension and physiological arousal may prolong sleep latency, increase the frequency of nighttime awakenings, and reduce sleep quality. Inadequate sleep may impair emotional regulation, increase fatigue, and exacerbate responses to stress. This relationship is particularly important during pregnancy because sleep disturbances not only affect maternal comfort but are also associated with maternal and perinatal complications. A systematic review and meta-analysis encompassing 120 studies found that sleep disturbances during pregnancy were associated with preeclampsia, gestational hypertension, gestational diabetes, cesarean delivery, preterm birth, and stillbirth. Sleep disturbances were reported to be associated with a 2.80-fold increased odds of preeclampsia and a 1.74-fold increased odds of gestational hypertension [4]. Stress management and improvement of sleep quality should therefore be incorporated into health promotion and complication prevention during pregnancy.

Nonpharmacological approaches based on mind-body interventions may represent safe and easily applicable strategies in maternal care. Prenatal yoga combines physical postures modified according to pregnancy conditions, breathing techniques, focused attention, body awareness, and relaxation [5]. This combination has the potential to reduce muscle tension, improve breathing patterns, decrease sympathetic

nervous system activity, increase parasympathetic activity, and help women achieve a state of physical and psychological relaxation. A randomized controlled trial demonstrated that prenatal yoga could reduce stress and cortisol levels and improve maternal immune function, as indicated by increased immunoglobulin A levels [6]. Broader evidence also supports the use of prenatal yoga to improve psychological health during pregnancy. The characteristics and effectiveness of prenatal yoga interventions indicate benefits for stress, anxiety, depression, pain, and several childbirth outcomes [7]. A meta-analysis of randomized controlled trials also found that mind-body interventions produced greater reductions in antenatal stress than control conditions [8].

Prenatal yoga combined with relaxation may influence stress and sleep quality by regulating both psychological and physiological responses. Relaxation following yoga practice may help maintain the calm state established during the exercise. In the present study, relaxation was conducted after prenatal yoga, allowing pregnant women an opportunity to reduce bodily tension, slow their breathing, and focus their attention on calming stimuli. A systematic review and meta-analysis showed that relaxation interventions during pregnancy, including muscle relaxation, breathing exercises, meditation, yoga, mindfulness, and music therapy, could reduce maternal stress and improve mental health. However, substantial variation remains in the forms, frequency, and intensity of these interventions, highlighting the need for studies employing more structured programs and repeated evaluations [9].

Prenatal yoga and relaxation may produce beneficial psychophysiological responses; however, changes in health behavior are not determined solely by the physical experience of exercise. Knowledge, perceived benefits, self-efficacy beliefs, affective responses to activities, interpersonal support, and commitment to maintaining an action also play roles in sustaining health behaviors. Health education can help women understand the risks associated with stress and sleep disturbances, recognize the benefits of preventive actions, enhance self-efficacy, and strengthen their commitment to engaging in health-promoting behaviors. Research has shown that self-efficacy-based pregnancy education classes can reduce anxiety and depression and improve maternal preparedness for childbirth [10]. Structured prenatal education can also reduce health anxiety among primigravida pregnant women [11].

Based on these considerations, integrating health education with prenatal yoga-relaxation may provide more comprehensive benefits than interventions focusing on only one component. Health education addresses cognitive and affective aspects and the development of commitment, whereas prenatal yoga-relaxation provides direct experiences for managing physical tension and stress responses. Repeated exposure may also facilitate learning, adaptation, reinforcement of self-efficacy, and gradual habit formation. Although evidence regarding the benefits of prenatal yoga, relaxation, and antenatal education continues to grow, several research gaps remain. First, some studies have evaluated prenatal yoga or health education separately, resulting in limited evidence regarding programs that integrate both approaches. Second, prenatal yoga protocols vary considerably and often do not incorporate factors that contribute to behavioral commitment. Third, many studies use a single pre- and post-intervention measurement, making it difficult to determine when changes begin to occur and whether the benefits increase with greater exposure to the intervention. Reviews of pregnancy yoga have also emphasized the substantial heterogeneity of intervention protocols and the need for studies with more structured intervention designs and reporting [7].

To address these gaps, this study developed an integrated health education and prenatal yoga-relaxation program delivered over 10 sessions. Evaluations were conducted at three time points: before the intervention at the first session, after the fifth session, and after the tenth session. Repeated measurements enabled an evaluation of the pattern and progression of responses throughout the program rather than merely assessing differences in scores before and after the intervention. This study aimed to analyze the effectiveness of an integrated health education and prenatal yoga-relaxation program on stress levels and sleep quality among pregnant women. The research hypothesis was that there would be a time-by-group interaction, with reductions in stress levels and improvements in sleep quality in the intervention group occurring progressively and being greater than those observed in the control group.

Methods

Study design and allocation

This study employed a controlled quasi-experimental design with repeated measures using a control group with time-series measurements. The study was conducted at four Independent Midwife Practice Sites (*Tempat Praktik Mandiri Bidan*, TPMB) in Surabaya City, East Java, Indonesia. Two TPMBs were assigned to the intervention condition and two TPMBs to the control condition. The intervention condition received an integrated health education and prenatal yoga-relaxation program, whereas the control condition received routine antenatal care and standard pregnancy exercise.

Because the intervention was delivered in groups at the TPMB level, allocation was performed by site rather than by individual participant. Thus, this study did not use individual-level randomization. Two TPMBs were designated as intervention sites and two as control sites based on site eligibility and operational feasibility, including the availability of a sufficiently large space to accommodate approximately 15 pregnant women simultaneously for the prenatal yoga–relaxation sessions. The geographical separation of intervention and control sites was maintained to reduce the possibility of information and activity contamination between groups.

Within each eligible TPMB, participants were selected using simple random sampling. A sampling frame was prepared from the list of pregnant women receiving antenatal care at each participating TPMB who met the eligibility criteria. Eligible women were assigned numbers, and 15 participants were selected from each TPMB by drawing lots. If a selected woman declined participation or did not provide written informed consent, another eligible woman was randomly selected from the same sampling frame until the required number of participants was reached.

This procedure resulted in 60 participants, comprising 30 women in the intervention group and 30 women in the control group. Therefore, simple random sampling was used for participant selection within each site, whereas group allocation occurred at the TPMB level. To reduce selection bias, the same eligibility criteria and gestational-age range (12–20 weeks) were applied across intervention and control sites, and participant selection within each site was conducted using random selection from the eligible sampling frame. Potential site-related bias was addressed by including two sites in each study condition, maintaining comparable eligibility criteria across sites, and geographically separating intervention and control sites to minimize contamination.

Setting

The study was conducted at four Independent Midwife Practice Sites (*Tempat Praktik Mandiri Bidan*, TPMB) in Surabaya City, East Java, Indonesia, from April to May 2026. Two TPMBs served as intervention sites and two as control sites. The intervention sites were selected based on their suitability for group-based prenatal yoga–relaxation sessions, particularly the availability of an adequate space for approximately 15 participants. Intervention and control sites were geographically separated to minimize the risk of information exchange and intervention contamination between groups.

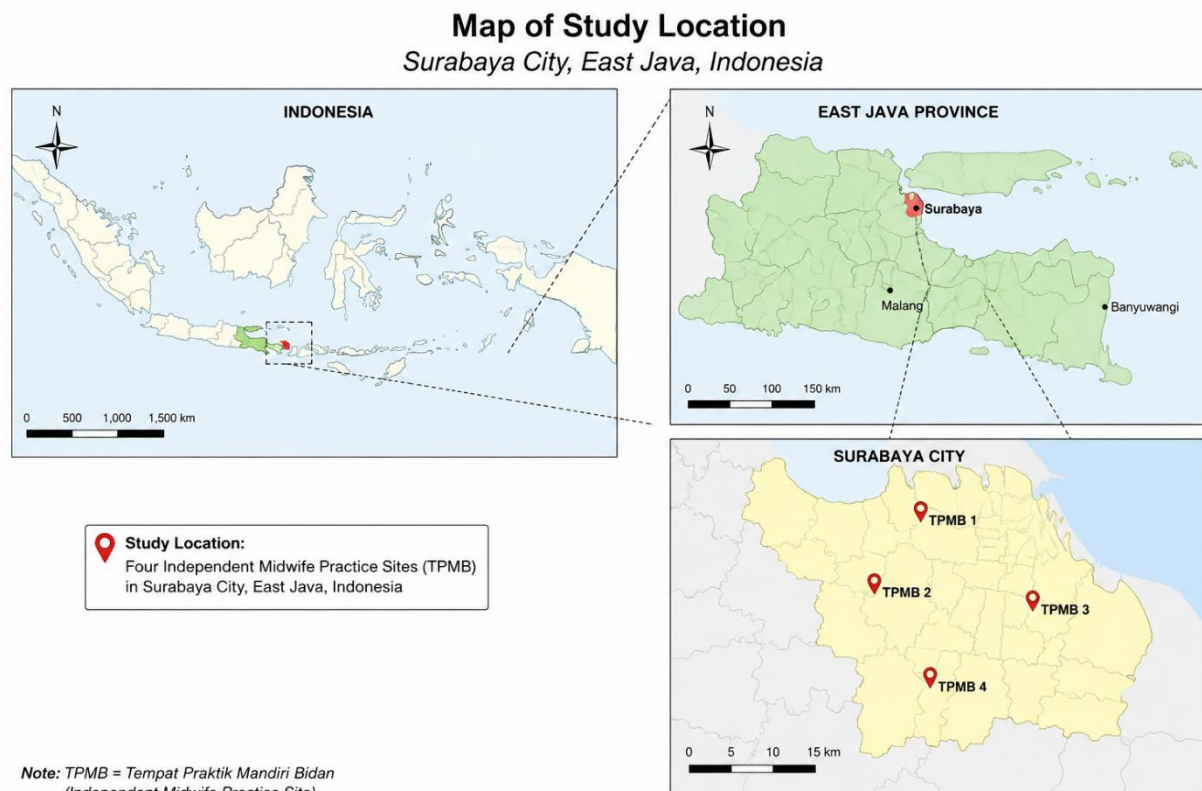


Figure 1. Map of the study location. The study was conducted at four Independent Midwife Practice Sites (TPMB) in Surabaya City, East Java, Indonesia

Participants

The study population consisted of pregnant women with a gestational age of 12–20 weeks who were receiving antenatal care at the participating TPMBs. The final sample comprised 60 pregnant women, with 30 participants in the intervention group and 30 in the control group. The inclusion criteria were: (1) pregnant women with a gestational age of 12–20 weeks; (2) able to communicate in Indonesian; (3) no medical contraindications to participation in prenatal yoga; (4) willing to participate in the study procedures; and (5) willing to provide written informed consent. The exclusion criteria were: (1) experiencing pregnancy complications that contraindicated participation in the intervention; (2) having musculoskeletal disorders or other medical conditions that could interfere with prenatal yoga; and (3) currently participating in a similar yoga program outside the study.

Sample size and sampling procedure

A total of 60 pregnant women were recruited, with 15 participants selected from each of the four participating TPMBs. The sample therefore comprised 30 participants from the two intervention sites and 30 participants from the two control sites.

Participant selection was performed using simple random sampling within each TPMB. The sampling frame consisted of eligible pregnant women identified from the antenatal-care records of each TPMB. Eligible women were assigned identification numbers, and 15 women per TPMB were selected by drawing lots. When a selected participant declined participation or did not provide informed consent, a replacement participant was selected randomly from the remaining eligible women in the same TPMB.

Data collection and intervention procedures

Eligible participants received an explanation of the study objectives, procedures, potential benefits and risks, confidentiality, and their right to withdraw from the study. Written informed consent was obtained before baseline data collection. Measurements were performed at three time points in both study groups: before the intervention at the first session (T0), at the end of the fifth session (T1), and at the end of the tenth session (T2). The intervention was conducted for five weeks, with two sessions per week, resulting in a total of 10 sessions.

Intervention group

Participants in the intervention group received an integrated health education and prenatal yoga–relaxation program. Each session lasted approximately 65–80 minutes and consisted of:

a) Health education and discussion (15–30 minutes). Each session began with health education, discussion, and question-and-answer activities. The educational content covered hypertension in pregnancy, complementary care, and the roles of family members, health professionals, and peers in supporting maternal health.

b) Prenatal yoga (approximately 30 minutes). Prenatal yoga was conducted under the supervision of a trained facilitator. The intervention included stretching and prenatal yoga postures, including bound angle pose, Sukhasana, Tadasana, Ardha Kati Chakrasana, Triangle Pose, Baddha Konasana, Ardha Pawanmuktasana, and side-lying positions. Participants were guided by the facilitator and supported by an instructional yoga video during the exercise session.

c) Relaxation with music (approximately 20 minutes). The relaxation component was performed in the side-lying position while participants listened to Mozart music for approximately 20 minutes.

The intervention facilitator had received prenatal-yoga training and was responsible for guiding participants, monitoring the implementation of the yoga postures, and ensuring that the activities were performed according to the intervention protocol. Participants received a study module containing educational materials related to hypertension in pregnancy and complementary care. The module was intended as supplementary reading at home to reinforce the educational content and was not provided as a manual for unsupervised prenatal yoga practice at home. Attendance was recorded at every intervention session. Participants were monitored throughout each session for discomfort or other adverse symptoms. If a participant experienced discomfort or symptoms requiring attention, the activity was stopped and the participant was assessed according to the study safety procedure.

Control group

Participants in the control group continued to receive routine antenatal care at their respective TPMBs and participated in standard pregnancy exercise provided at the site. They did not receive the integrated health education and prenatal yoga–relaxation program delivered to the intervention group during the study period. Because the control group received routine antenatal care and pregnancy exercise rather than the study intervention, the amount and content of participant contact were not identical

between groups. The control group nevertheless underwent the same outcome assessments at T0, T1, and T2.

Measurement instruments

Data were collected using structured questionnaires administered in Indonesian. The questionnaire included measures of participant characteristics, previous pregnancy experience, cognitive and affective factors, interpersonal support, commitment to action, stress, and sleep quality.

a) Participant characteristics and previous experience. Participant characteristics included education and household income. Previous experience included parity and pregnancy history.

b) Cognitive and affective variables. Perceived benefits, including experience, understanding, and belief, were measured using a questionnaire adapted from Puspita et al., consisting of six items, with reported item validity coefficients ranging from 0.550 to 0.816 [12]. Self-efficacy was assessed using the General Self-Efficacy Scale (GSES), consisting of 10 items, with reported validity coefficients ranging from 0.654 to 0.888 [13]. Activity-related affect, including positive and negative affect, was measured using an instrument adapted from Bahabadi et al., consisting of six items, with reported validity coefficients ranging from 0.677 to 0.891 [14].

c) Interpersonal support. Family support was assessed using a 10-item questionnaire covering three dimensions, with reported validity coefficients ranging from 0.490 to 0.808 [15]. Support from health professionals was assessed using the Health Care Climate Questionnaire (HCCQ), consisting of four items, with reported validity coefficients ranging from 0.488 to 0.804 [16]. Peer support was assessed using the Perceived Social Support from Friends (PSSF) scale, consisting of 10 items, with reported validity coefficients ranging from 0.685 to 0.890.

d) Commitment to action. Commitment to action was assessed using an eight-item questionnaire measuring awareness to act, with reported validity coefficients ranging from 0.538 to 0.909 [17].

e) Stress. Stress was assessed using the stress subscale of the Depression Anxiety Stress Scales-42 (DASS-42) [18,19]. The stress subscale consists of 14 items, with reported validity coefficients ranging from 0.451 to 0.864. The Indonesian-language version was used in this study. For each item, participants indicated the extent to which the statement applied to them. The stress score was calculated by summing the 14 stress items. In accordance with the scoring method used in this study, the reported DASS-42 stress scores were raw scores and were not multiplied by two. Therefore, the possible raw stress score ranged from 0 to 42, with higher scores indicating greater perceived stress. [If your actual analysis multiplied the score by two, replace this sentence with the appropriate multiplied-score procedure and score range of 0–84.]

f) Sleep quality. Sleep quality was assessed using the Pittsburgh Sleep Quality Index (PSQI) [20]. The Indonesian version of the PSQI was administered to participants. The instrument used in this study consisted of 15 questionnaire items, with reported validity coefficients ranging from 0.361 to 0.766 [21].

For all instruments adapted for this study, the Indonesian-language versions were reviewed and, where applicable, subjected to validity testing before use. [INSERT: the reliability coefficients for each instrument in the present study or the published Indonesian validation study, as appropriate.] The score ranges and interpretation of each instrument were based on the respective instrument's scoring guidelines.

Outcome assessment

The primary outcomes were stress level and sleep quality. Stress was measured using the DASS-42 stress subscale, while sleep quality was measured using the PSQI. Both outcomes were assessed at baseline (T0), after the fifth intervention session (T1), and after the tenth intervention session (T2) in both the intervention and control groups.

Statistical analysis

Descriptive analysis was used to calculate the mean and standard deviation. Data normality was assessed using the Shapiro–Wilk test. Statistical analyses were performed using SPSS version 29. The equivalence of respondent characteristics between the control and intervention groups at baseline, before the intervention, was assessed using the chi-square test. To assess changes in outcomes across the pretest and posttest measurements, repeated-measures analysis of variance (ANOVA) was used for normally distributed data, whereas the Friedman test was used for data that were not normally distributed. For repeated-measures ANOVA, Mauchly's test of sphericity was performed to assess the assumption of sphericity. When this assumption was violated ($p < 0.05$), the Greenhouse–Geisser correction was applied. The significance level was set at 0.05, with a 95% confidence interval. The intervention was considered to have a statistically significant effect when the p-value was less than 0.05, whereas a p-value greater than 0.05 was interpreted as indicating no statistically significant effect of the intervention.

When repeated-measures ANOVA indicated a statistically significant difference across measurement time points, post hoc pairwise comparisons were conducted using the Bonferroni correction. For non-normally distributed data, when the Friedman test indicated a statistically significant difference, post hoc pairwise comparisons between time points were conducted using the Wilcoxon signed-rank test. The significance level for post hoc comparisons was adjusted to 0.017 (0.05/3), based on the three measurement time points and the corresponding pairwise comparisons.

Ethical practices

This study was approved by the Health Research Ethics Committee, Faculty of Nursing, Universitas Airlangga, Indonesia (Ethical Approval No. 3712-KEPK). All participants were provided with information regarding the study objectives, procedures, potential benefits and risks, confidentiality of their data, and their rights as participants. Written informed consent was obtained from all participants before the study commenced. Participants were informed that participation was voluntary and that they had the right to withdraw from the study at any stage without any consequences to their antenatal care or other services. All participant information and research data were kept confidential and used solely for research purposes.

Results

The findings showed that, in the control group, the majority of respondents had junior high school and senior high school education, an average income below the regional minimum wage (UMR), were primigravida, and had low-risk pregnancies (Table 1). In the intervention group, the majority of respondents had a senior high school education, an income below the UMR, were multigravida, and all participants had low-risk pregnancies. The homogeneity test indicates that the control group and the intervention group are comparable. The Pearson chi-square test shows no significant differences between the groups in terms of education ($p = 0.425$), income ($p = 0.121$), parity ($p = 0.121$), and pregnancy history ($p = 0.492$). All respondent characteristics had p -values > 0.05 , indicating that there were no statistically significant differences in baseline characteristics between the control and intervention groups before the intervention, demonstrating that the groups were homogeneous and comparable.

Table 1. Demographic characteristics of pregnant women

Characteristic	Control n (%)	Intervention n (%)	Homogeneity test (Chi-square)
Education			
Elementary school	4 (13.3)	3 (10.0)	2.792 ($p = 0.425$)
Junior high school	11 (36.7)	6 (20.0)	
Senior high school/Vocational high school	11 (36.7)	14 (46.7)	
Higher education	4 (13.3)	7 (23.3)	
Income			
< UMR	16 (53.3)	17 (56.7)	2.403 ($p = 0.121$)
UMR	7 (23.3)	4 (13.3)	
> UMR	7 (23.3)	9 (30.0)	
Parity			
Primigravida	18 (60.0)	11 (36.7)	2.403 ($p = 0.121$)
1–4 times	12 (40.0)	19 (63.3)	
Pregnancy History			
Low risk	28 (93.3)	30 (100.0)	2.069 ($p = 0.492$)
High risk.	2 (6.7)	0 (0.0)	

Perceived benefits

In the control group (Table 2), the Friedman test revealed a statistically significant difference in perceived benefit scores across T0, T1, and T2 ($\chi^2=11.565$, $p=0.003$, $p<0.005$). Post hoc analysis using the Wilcoxon signed-rank test showed no statistically significant difference between T0 and T1 ($p=0.034$, $p>0.017$). Similarly, no statistically significant difference was found between T1 and T2 ($p=0.059$, $p>0.017$). However, a statistically significant difference was observed between T0 and T2 ($p=0.014$, $p<0.017$). In the intervention group, the Friedman test revealed a statistically significant difference in perceived benefit scores across T0, T1, and T2 ($\chi^2=56.214$, $p<0.001$, $p<0.005$). Post hoc analysis using the Wilcoxon signed-rank test showed statistically significant differences between T0 and T1 ($p<0.001$, $p<0.017$), between T1 and T2 ($p<0.001$, $p<0.017$), and between T0 and T2 ($p<0.001$, $p<0.017$). The analysis revealed a statistically

significant time-by-group interaction for perceived benefit scores ($F=32.351$, $p<0.001$), indicating that the patterns of change from T0 to T1 and T2 differed between the intervention and control groups.

In Press

Table 2. Differences in perceived benefits, self-efficacy, activity-related affect, at T0, T1, and T2 in the control and intervention groups

Variable	Statistic	T0	T1	T2	Statistical test	p-value	T0 vs. T1 (p-value)	T1 vs. T2 (p-value)	T0 vs. T2 (p-value)	Time × Group (F)
Perceived benefits										
Control	Median	18.00	18.00	18.00	$\chi^2 = 11.565$	0.003	0.034	0.059	0.014	32.351
Intervention	Median	18.00	19.00	20.00	$\chi^2 = 56.214$	<0.001	<0.001	<0.001	<0.001	
Self-efficacy										
Control	Mean (SD)	30.63 (3.378)	30.93 (3.216)	31.03 (3.135)	W = 23.190	<0.001	0.052	0.249	0.015	51.461
	Greenhouse-Geisser corrected F					7.490 (0.006)				
Intervention	Mean (SD)	30.30 (4.580)	31.70 (3.993)	33.83 (3.185)	$\chi^2 = 53.626$	<0.001	<0.001	<0.001	<0.001	
Activity-related affect										
Control	Mean (SD)	16.03 (3.846)	16.40 (3.663)	16.67 (3.437)	W = 12.635	0.002	<0.001	0.027	<0.001	38.411
	Greenhouse-Geisser corrected F					16.597 (<0.001)				
Intervention	Mean (SD)	16.27 (4.653)	17.30 (3.870)	19.23 (2.725)	W = 50.155	<0.001	<0.001	<0.001	<0.001	
	Greenhouse-Geisser corrected F					38.966 (<0.001)				

Note. T0 = baseline; T1 = first follow-up; T2 = second follow-up; SD = standard deviation; χ^2 = Friedman's test; W= Mauchly's test; F = F-statistic. Post hoc comparisons were performed using the Wilcoxon signed-rank test for Friedman analyses and the Bonferroni correction for repeated-measures ANOVA, as applicable. The time × group interaction indicates whether changes over time differed between the control and intervention groups.

Self-efficacy

In the control group, repeated-measures ANOVA yielded a Mauchly's test of sphericity statistic of 23.190, with $p=0.000$, indicating that the assumption of sphericity was violated. Therefore, the Greenhouse–Geisser correction was applied, yielding $F=7.490$ and $p=0.006$, indicating a statistically significant difference in self-efficacy scores over time ($p<0.05$). Post hoc analysis using the Bonferroni correction showed that the self-efficacy score at T1 ($M=30.933$, $SD=0.587$) was slightly higher than that at T0 ($M=30.633$, $SD=0.617$). However, this increase was not statistically significant ($p=0.052$, $p>0.017$). The self-efficacy score at T2 ($M=31.033$, $SD=0.572$) was higher than that at T1, but the difference was not statistically significant ($p=0.249$, $p>0.017$). However, the score at T2 was significantly higher than that at T0 ($p=0.015$, $p<0.017$). In the intervention group, the Friedman test revealed a statistically significant difference in self-efficacy scores across T0, T1, and T2 ($\chi^2=53.626$, $p<0.001$, $p<0.05$). Post hoc analysis using the Wilcoxon signed-rank test showed statistically significant differences between T0 and T1 ($p<0.001$, $p<0.017$), between T1 and T2 ($p<0.001$, $p<0.017$), and between T0 and T2 ($p<0.001$, $p<0.017$). The analysis revealed a statistically significant time-by-group interaction for self-efficacy scores ($F=51.461$, $p<0.001$), indicating that the patterns of change from T0 to T1 and T2 differed between the intervention and control groups.

Activity-Related Affect

In the control group, repeated-measures ANOVA yielded a Mauchly's test of sphericity statistic of 12.635, with $p=0.002$, indicating that the assumption of sphericity was violated. Therefore, the Greenhouse–Geisser correction was applied, yielding $F=16.597$ and $p<0.001$. This indicated a statistically significant difference in activity-related affect scores across the different measurement times ($p<0.05$). Post hoc analysis using the Bonferroni correction showed that the mean score at T1 ($M=16.40$, $SD=3.663$) was higher than that at T0 ($M=16.03$, $SD=3.846$). This increase was statistically significant ($p<0.001$, $p<0.017$). Furthermore, the activity-related affect score at T2 ($M=16.667$, $SD=3.437$) was higher than that at T1 ($M=16.40$, $SD=3.663$). This increase was also statistically significant ($p<0.001$, $p<0.017$). In the intervention group, repeated-measures ANOVA yielded a Mauchly's test of sphericity statistic of 50.155, with $p<0.001$, indicating that the assumption of sphericity was violated. Therefore, the Greenhouse–Geisser correction was applied, yielding $F=38.966$ and $p<0.001$, indicating a statistically significant difference in activity-related affect scores over time ($p<0.05$). Post hoc analysis using the Bonferroni correction showed that the score at T1 ($M=17.30$, $SD=3.870$) was higher than that at T0 ($M=16.27$, $SD=4.653$). This increase was statistically significant ($p<0.001$, $p<0.017$). The activity-related affect score at T2 ($M=19.23$, $SD=2.725$) was higher than that at T1, and this increase was statistically significant ($p<0.001$, $p<0.017$). Similarly, the score at T2 was significantly higher than that at T0 ($p<0.001$, $p<0.017$). The analysis revealed a statistically significant time-by-group interaction for activity-related affect scores ($F=38.411$, $p<0.001$), indicating that the patterns of change from T0 to T1 and T2 differed between the intervention and control groups.

Commitment

In the control group, the Friedman test for commitment yielded a chi-square statistic of 55.861 and a p-value of less than 0.001, indicating a statistically significant difference in commitment levels across the measurement time points. Post hoc analysis using the Wilcoxon signed-rank test revealed statistically significant differences between T0 and T1 ($p<0.001$, $p<0.017$), between T1 and T2 ($p<0.001$, $p<0.017$), and between T0 and T2 ($p<0.001$, $p<0.017$). In the intervention group, the Friedman test yielded a chi-square statistic of 60.000 and a p-value of less than 0.001, indicating a statistically significant difference in commitment levels across the measurement time points. Post hoc analysis using the Wilcoxon signed-rank test also revealed statistically significant differences between T0 and T1 ($p<0.001$, $p<0.017$), between T1 and T2 ($p<0.001$, $p<0.017$), and between T0 and T2 ($p<0.001$, $p<0.017$). The analysis revealed a statistically significant time-by-group interaction for commitment scores ($F=148.960$, $p<0.001$), indicating that the patterns of change from T0 to T1 and T2 differed between the intervention and control groups. These findings indicate that commitment scores increased over time in both groups, with a substantially greater increase observed in the intervention group than in the control group. The median commitment score in the control group increased from 9.50 at baseline to 12.00 at T2, whereas the median score in the intervention group increased from 10.00 to 19.00 over the same period. The intervention, which consisted of a pregnancy hypertension prevention module based on the Health Promotion Model and complementary care, was effective in increasing participants' commitment to practicing complementary care, specifically mind–body interventions.

Stress Levels

In the control group, repeated-measures ANOVA yielded a Mauchly's test of sphericity statistic of 61.080, with $p=0.000$, indicating that the assumption of sphericity was violated. Therefore, the Greenhouse–Geisser correction was applied, yielding $F=0.875$ and $p=0.363$. This result indicated no statistically significant difference in stress scores over time ($p>0.05$). No post hoc tests were conducted because the repeated-measures ANOVA result was not statistically significant. In the intervention group, the Friedman test revealed a statistically significant difference in stress scores across T0, T1, and T2 ($\chi^2=56.165$, $p<0.001$, $p<0.05$). Post hoc analysis using the Wilcoxon signed-rank test showed statistically significant differences between T0 and T1 ($p<0.001$, $p<0.017$), between T1 and T2 ($p<0.001$, $p<0.017$), and between T0 and T2 ($p<0.001$, $p<0.017$). The analysis revealed a statistically significant time-by-group interaction for stress scores ($F=69.179$, $p<0.001$), indicating that the patterns of change from T0 to T1 and T2 differed between the intervention and control groups.

Sleep Quality

In the control group, the Friedman test yielded a chi-square statistic of 15.462 and a p-value of 0.001, indicating a statistically significant difference in sleep quality scores across the measurement time points. Post hoc analysis using the Wilcoxon signed-rank test revealed a statistically significant difference between T0 and T1 ($p=0.008$, $p<0.017$). A statistically significant difference was also observed between T1 and T2 ($p=0.006$, $p<0.017$), as well as between T0 and T2 ($p=0.002$, $p<0.017$). In the intervention group, the Friedman test yielded a chi-square statistic of 41.407 and a reported p-value of 0.001, indicating a statistically significant difference in sleep quality scores across the measurement time points. Post hoc analysis using the Wilcoxon signed-rank test revealed statistically significant differences between T0 and T1 ($p<0.001$, $p<0.017$), between T1 and T2 ($p<0.001$, $p<0.017$), and between T0 and T2 ($p<0.001$, $p<0.017$).

The analysis revealed a statistically significant time-by-group interaction for sleep quality scores ($F=13.862$, $p<0.001$), indicating that the patterns of change from T0 to T1 and T2 differed between the intervention and control groups. Sleep quality improved in both groups, with a substantially greater improvement observed in the intervention group. The median score in the control group decreased from 8.50 at baseline to 8.00 at T2, whereas the median score in the intervention group decreased from 8.00 to 5.00. Assuming that lower scores indicate better sleep quality, these findings suggest that the intervention had a stronger effect on improving sleep quality than the absence of the intervention.

Table 3. Differences in commitment scores at T0, T1, and T2 in the control and intervention groups

Group	Statistic	T0	T1	T2	Friedman test (χ^2)	p-value	T0 vs. T1 (p-value)	T1 vs. T2 (p-value)	T0 vs. T2 (p-value)	Time × Group (F)
Control	Median	9.50	11.00	12.00	$\chi^2 = 55.861$	< 0.001	< 0.001	< 0.001	< 0.001	148.960
Intervention	Median	10.00	13.00	19.00	$\chi^2 = 60.00$	< 0.001	< 0.001	< 0.001	< 0.001	

Note. T0 = baseline; T1 = first follow-up; T2 = second follow-up; χ^2 = chi-square statistic. Post hoc comparisons were performed using the Wilcoxon signed-rank test. The time × group interaction indicates whether changes in commitment scores over time differed between the control and intervention groups.

Table 4. Differences in stress levels at T0, T1, and T2 in the control and intervention groups

Group	Statistic	T0	T1	T2	Test	p-value	T0 vs. T1 (p-value)	T1 vs. T2 (p-value)	T0 vs. T2 (p-value)	Time × Group (F)
Control	Mean (SD)	19.47 (6.08)	20.07 (4.96)	19.70 (4.92)	W = 61.080 (0.000)	F = 0.875 (0.363)	—	—	—	69.179
Intervention	Median	19.00	18.50	11.00	$\chi^2 = 56.165$	0.001	< 0.001	< 0.001	< 0.001	

Note. T0 = baseline; T1 = first follow-up; T2 = second follow-up; SD = standard deviation; F = F-statistic. Post hoc comparisons were not performed for the control group because the repeated-measures ANOVA result was not statistically significant. For the intervention group, post hoc comparisons were performed using the Wilcoxon signed-rank test.

Table 5. Differences in sleep quality at T0, T1, and T2 in the control and intervention groups

Group	Statistic	T0	T1	T2	Friedman test (χ^2)	p-value	T0 vs. T1 (p-value)	T1 vs. T2 (p-value)	T0 vs. T2 (p-value)	Time × Group (F)
Control	Median	8.50	8.00	8.00	$\chi^2 = 15.462$	0.001	0.008	0.006	0.002	13.862
Intervention	Median	8.00	5.00	5.00	$\chi^2 = 41.407$	0.001	< 0.001	< 0.001	< 0.001	

Note. T0 = baseline; T1 = first follow-up; T2 = second follow-up; χ^2 = chi-square statistic; F = F-statistic. Post hoc comparisons were performed using the Wilcoxon signed-rank test. The time × group interaction indicates whether changes in sleep quality over time differed between the control and intervention groups.

Discussion

This study evaluated changes in stress levels and sleep quality among pregnant women receiving an integrated Health Education and Prenatal Yoga–Relaxation intervention at three measurement time points, compared with women receiving standard antenatal care supplemented with pregnancy exercise e. The findings indicated more favorable descriptive changes in both primary outcomes in the intervention group. The significant time $\times$ group interactions for stress levels ($F = 69.179$, $p < 0.001$) and sleep quality ($F = 13.862$, $p < 0.001$) indicated that the patterns of change over time differed between the groups. However, because this study used a quasi-experimental, site-based design, the findings should be interpreted as associations rather than definitive evidence of causality. Repeated sessions provided opportunities for participants to become familiar with the educational materials, yoga movements, breathing techniques, and relaxation procedures. These experiences may have supported engagement with the program, although the internalization of educational materials and the application of learned skills in daily life were not directly assessed. These considerations are relevant to pregnancy, which involves physical, hormonal, psychological, social, and family-role changes [22].

Stress levels were one of the primary outcomes. The intervention group demonstrated statistically significant changes in stress scores across T0, T1, and T2, whereas the control group showed no statistically significant within-group change over time (Greenhouse–Geisser corrected $F = 0.875$, $p = 0.363$). The significant time $\times$ group interaction provides evidence that the trajectories of stress scores differed between the groups. These findings are consistent with systematic reviews and meta-analyses reporting potential benefits of prenatal yoga for maternal stress, anxiety, depression, and emotional well-being [7,23,24]. However, differences in intervention duration, frequency, program components, and participant characteristics limit direct comparisons across studies. The present study also combined health education, prenatal yoga, and relaxation accompanied by Mozart music. Therefore, the observed differences cannot be attributed exclusively to yoga or to any single intervention component. Moreover, the magnitude of benefit relative to previous studies cannot be established reliably without comparable effect estimates and measurement scales.

Several explanations for the observed stress-related changes have been proposed in previous literature. Breathing exercises, controlled movements, focused attention, and relaxation may help participants manage perceived tension and respond to stressful situations. Neurophysiological, neuroendocrine, psychological, and behavioral pathways have also been discussed as possible explanations for the relationship between mind–body practices and stress [25–27]. Previous studies have reported changes in subjective and physiological stress indicators following prenatal yoga [6,28]. Nevertheless, this study did not measure autonomic nervous system activity, hypothalamic–pituitary–adrenal axis function, or cortisol concentrations. Consequently, these mechanisms cannot be confirmed as explanations for the observed findings. Similarly, the specific contribution of Mozart music was not evaluated separately. These possibilities should therefore be considered hypotheses informed by previous literature rather than mechanisms demonstrated by the present study.

Sleep quality was the second primary outcome, and its interpretation requires consideration of the results in both groups. The median sleep quality score decreased from 8.50 at T0 to 8.00 at T2 in the control group and from 8.00 to 5.00 in the intervention group. Assuming that lower scores indicate better sleep quality, the descriptive changes suggest a greater improvement in the intervention group. Nevertheless, the control group also demonstrated statistically significant changes across measurement times (Friedman $\chi^2 = 15.462$, $p = 0.001$), with significant pairwise comparisons between T0 and T1 ($p = 0.008$), T1 and T2 ($p = 0.006$), and T0 and T2 ($p = 0.002$). The intervention group likewise demonstrated significant changes over time (Friedman $\chi^2 = 41.407$, reported $p = 0.001$), with all three pairwise comparisons reported as $p < 0.001$. These findings show that improvement was not exclusive to the intervention group. The significant time $\times$ group interaction supports a difference in the patterns of change between the groups, but it does not establish that every observed change resulted from the intervention itself.

Reduced physical discomfort, opportunities for relaxation, and engagement in structured activities may be relevant to understanding changes in sleep quality. Previous research on relaxation interventions during pregnancy, including yoga, music, breathing exercises, and mindfulness, provides context for the potential benefits of these approaches [9,29,30]. Sleep and stress may also be related, because stress can interfere with sleep, while adequate sleep may support emotional regulation and daily coping [31]. However, the direction of this relationship was not examined in the present study, and no mediation analysis was conducted. It is therefore not possible to determine whether changes in stress contributed to changes in sleep quality, whether sleep-related changes influenced stress, or whether both outcomes were associated with other aspects of the intervention or participants' circumstances. In addition, the reported sleep scores should be interpreted according to the scoring direction and measurement properties of the instrument used.

Changes in perceived benefits, self-efficacy, activity-related affect, and commitment provide secondary theoretical context for the primary outcomes. In the intervention group, median perceived benefit scores increased from 18.00 at T0 to 20.00 at T2, while the control group's median remained at 18.00. Perceived benefits may reflect participants' increasingly positive evaluations of complementary care. According to the Health Promotion Model, perceived benefits are associated with health-promoting behaviors, and previous research among pregnant women has identified perceived benefits and self-efficacy as relevant predictors of such behaviors [14]. Mean self-efficacy scores in the intervention group increased from 30.30 to 33.83, compared with an increase from 30.63 to 31.03 in the control group. This pattern may reflect increased confidence in performing Prenatal Yoga-Relaxation. Information, demonstrations, guided practice, and facilitator feedback are possible contributors, consistent with previous studies of Health Promotion Model-based self-care interventions among pregnant women [32,33]. However, the individual contributions of these components were not evaluated.

Activity-related affect also increased in the intervention group, with mean scores rising from 16.27 at T0 to 19.23 at T2, compared with an increase from 16.03 to 16.67 in the control group. Positive experiences during structured activities may be associated with willingness to repeat health-promoting behaviors. The supportive group environment, interactions with the facilitator, and social support may also have contributed to participants' experiences. Although self-efficacy and activity-related affect increased in parallel, their reciprocal relationship and temporal sequence were not tested. These variables should therefore be interpreted as associated theoretical outcomes rather than established mediators of changes in stress or sleep quality.

Commitment scores also increased in both groups, with a larger descriptive increase in the intervention group. The median score increased from 10.00 to 19.00 in the intervention group and from 9.50 to 12.00 in the control group. The significant time $\times$ group interaction for commitment ($F = 148.960$, $p < 0.001$) indicated that changes differed between the groups. Within the Health Promotion Model, perceived benefits, self-efficacy, and positive affective experiences may support commitment to planned health-promoting actions. Nevertheless, commitment was not tested as a mediator between cognitive-affective changes and the primary outcomes. The findings do not establish that increased commitment led to reduced stress or improved sleep, nor do they demonstrate that participants maintained these behaviors after the intervention ended.

Several limitations should be considered when interpreting the findings. First, the quasi-experimental design and allocation of participants across four independent midwifery practices (TPMBs) may have introduced site-related confounding associated with healthcare personnel, facilities, service provision, and local practices. Random selection of respondents within each practice and similar gestational-age eligibility criteria did not ensure equivalence in all baseline characteristics. Second, the control group received standard antenatal care supplemented with pregnancy exercise, meaning that the comparison involved two different intervention packages. Differences in contact intensity, facilitator attention, participant expectations, and other activities during the study period may have contributed to the observed outcomes. Third, the relatively small sample size and five-week intervention period limit generalizability and prevent conclusions about the durability of the changes. Fourth, self-reported measures of stress and sleep quality may be affected by recall, reporting, and social desirability biases. Finally, the study did not assess objective physiological indicators or test the proposed mediating relationships among the theoretical variables and primary outcomes.

Despite these limitations, the findings suggest that the integrated Health Education and Prenatal Yoga-Relaxation program was associated with more favorable changes in stress levels and sleep quality than the comparison condition. The significant time $\times$ group interactions support further evaluation of this program as a complementary approach to antenatal care, rather than establishing its independent causal effect. Future research should employ randomized designs across a larger number of sites, use comparison conditions with similar levels of contact and attention, and extend follow-up beyond the intervention period. Where feasible, objective stress and sleep measures, prespecified effect estimates, and analyses of potential mediators should be included. Component-specific comparisons would also help determine the contributions of education, yoga, relaxation, and music. Such studies would strengthen causal inference, clarify the mechanisms that may be involved, and establish whether improvements are sustained over time.

Conclusion

The integrated Health Education and Prenatal Yoga-Relaxation program was associated with improvements in stress levels and sleep quality among pregnant women. Significant time $\times$ group interactions indicated that changes in both outcomes over time differed between the intervention and control groups. However, given the quasi-experimental design, these findings should be interpreted cautiously. The intervention may be considered a complementary approach to antenatal care to support

maternal psychological well-being and sleep quality, although further randomized studies with longer follow-up are needed to confirm its effects and sustainability.

In Press