

Materials and methods

Study design and participants

We did a two-arm, parallel-group, open-label, randomized controlled trial (RCT) in nine obstetric centers across China, primarily within the southeastern region. These obstetric centers were affiliated with the following healthcare institutions: Women's Hospital, School of Medicine, Zhejiang University; Gansu Provincial Maternal and Child Health Care Hospital; Hangzhou First People's Hospital; Yiwu Maternity and Children Hospital; The Second Affiliated Hospital of Wenzhou Medical University; Wenzhou Central Hospital; Hangzhou Linping District Maternity and Child Health Care Hospital; The Fourth Affiliated Hospital, Zhejiang University School of Medicine; and Tongxiang Maternal and Child Health Hospital. Eligible primiparous women were screened from April 2024 to May 2025. The complete inclusion and exclusion criteria for participation in the RCT were published previously in the study protocol (Li et al., 2025). Ethical approval was granted by the Ethics Committee of Women's Hospital, School of Medicine, Zhejiang University (Approval No. IRB-20230379-R).

Randomization and blinding

The eligible participants were randomly allocated to the control and intervention groups in a 1:1 ratio, using a center-stratified block randomization method. Randomization was performed with SPSS by an independent statistician (Version 25.0). Blinding was not implemented, as no feasible placebo was available for the hot-compress intervention (Li et al., 2025).

Procedures

The whole procedure was performed strictly in accordance with a previously published protocol (Li et al., 2025). Briefly, the control group received standard obstetrical care, and the intervention group received the same standard obstetrical care in addition to acupoint hot-compress therapy at $(42\pm 2)^\circ\text{C}$ for 4 h, starting 1 h after the onset of regular uterine contractions during the latent phase of labor. The intervention used a Hu-Chao-Nuan-Gong-Bao device, a licensed Class II item (licence No. 20192090292) that maintains a stable temperature. Two sizes of applicators were employed: model A (16 cm $\times$ 9 cm) with four heating cores (each 8 cm $\times$ 7 cm), and model B (13 cm $\times$ 10 cm) with one larger core (11.5 cm $\times$ 8 cm), applied to the designated acupoints. In accordance with our published protocol (Li et al., 2025), the intervention applied hot compress to three specific acupoints: Yongquan (KI1) on the plantar region (Hsieh et al., 2016; Zhu et al., 2022), Sanyinjiao (SP6) on the medial malleolar region (Feng et al., 2012; Lee et al., 2012; Du et al., 2017; Chen et al., 2025), and Baliao (BL31–34) on the lumbosacral region (Zhu et al., 2022). Baseline labor-pain intensity was assessed at 1 h after the onset of regular uterine contractions prior to intervention administration. The acupoint hot compress was removed if the pregnant woman was to undergo labor epidural analgesia with no contraindications. Assessment of labor-pain intensity was discontinued after the administration of epidural analgesia.

Outcomes

The primary outcome to be measured was labor pain, which we assessed starting from the latent phase of the first stage, at 3 and 5 h after the onset of regular uterine contractions, with labor pain at 1 h serving as the baseline. Pain intensity was assessed using the Visual Analog Scale (VAS), scored from 0–10. Women were asked to choose a score of 0 if they had no pain, between 1–3 for mild pain, 4–6 for moderate pain, and 7–10 for severe pain (Chang et al., 2014; Ghoneim et al., 2022).

The secondary outcomes were labor duration, postpartum blood loss, symptoms of depression, and newborn Apgar scores. Labor duration included the duration of the first, second, and third stages of labor, and was recorded in minutes by a blinded assessor using a digital stopwatch (de Sena Fraga et al., 2024). Postpartum blood loss was quantified at 0 (after placental delivery) and 2 h postpartum through the volume and weight of vaginal blood loss (ACOG, 2019). Depression symptoms were identified using the Edinburgh Postnatal Depression Scale (EPDS) within 48 h postpartum (Singla et al., 2025). The self-reported EPDS contained 10 items: pessimism, lack of interest, self-blame, worry, fear, impaired ability, sleep disorder, sadness, tearfulness, and self-injury or suicidal ideation. The neonatal Apgar score of five physical signs (pulse, respiration, appearance, muscle tone, and reflex irritability) was evaluated at 1 and 5 min after vaginal birth, serving as a predictor of neonatal mortality, morbidity, and long-term neurodevelopmental outcomes (Simon et al., 2025).

Statistical analysis

In our planned protocol (Li et al., 2025), the VAS score was considered a continuous variable. However, during the conduct of this trial, the majority of parturients were only able to assess their labor pain at the level of a categorical variable (mild, moderate, or severe). Hence, we collected the VAS scores accordingly, and following an expanded pilot study, the sample size was re-estimated. Based on preliminary data from 100 participants (50 in each group), a significant difference was observed between the intervention and control groups in the distribution of categorical VAS scores at 5 h after the onset of regular uterine contractions. In the intervention group, the proportions of mild, moderate, and severe pain were 12.00%, 78.00%, and 10.00%, respectively, whereas the corresponding proportions in the control group were 4.00%, 68.00%, and 28.00%. The effect size, calculated as Cramer's V from the chi-square test, was 0.26. With 95% power, a two-sided threshold of 0.05 and a 20% dropout rate, the required sample size was at least 275 per group, and the number of obstetric centers in the study was decreased accordingly.

All statistical analyses were performed on the modified intention-to-treat (ITT) population, which consisted of all randomly assigned patients who completed the assessment of baseline labor-pain intensity. None of the women in the modified ITT population received labor analgesia at any point. Comparisons are made between the population that completed the entire VAS-score assessment and the dropouts in terms of baseline characteristics. Normality assumptions were checked first. Continuous variables are presented as mean $\pm$ standard deviation (SD) or median (25th percentile, 75th percentile) according to the distribution of variables, and categorical variables are described as numbers and percentages. Continuous variables are compared using independent t-tests or the Wilcoxon-Mann-Whitney test, whereas categorical variables are compared with the chi-squared test or Fisher's exact test.

For the primary outcome, ordinal logistic regression was used to further explore the relationship between labor pain and acupoint hot compress, considering the significant differences in baseline labor-pain intensity between the intervention and control groups. The model was adjusted for labor-pain intensity 1 h after the onset of regular uterine contractions to enhance the precision of treatment-effect estimation. The results of the ordinal logistic regression analysis are presented as B coefficients and odds ratios (OR) with a 95% confidence interval (95% CI), and the proportional odds assumptions were met. Sensitivity analyses of the primary outcome were carried out in the per-protocol population who completed the entire VAS-score assessment. Multiple imputation by chained equations accounted for missing data on the primary outcome, and estimates from 50 imputed datasets were generated under the missing-at-random assumption. Model estimates were pooled across imputations following the Rubin rules, with results reported as OR and 95% CI. Moreover, a worst-case imputation was applied, under which all missing primary outcomes were assumed to represent severe pain in both groups.

For the secondary outcomes, we performed the primary analysis in the per-protocol population to explore the effect of acupoint hot compress on outcomes related to vaginal delivery. Sensitivity analysis was conducted in the modified ITT population limited to vaginal delivery. In addition, an exploratory subgroup analysis based

on the use of labor epidural analgesia was performed to investigate whether the effect of the acupoint hot compress on secondary outcomes was related to labor epidural analgesia.

Independent statistical analyses were conducted using SPSS (Version 25.0), with multiple imputation performed via the MICE package in R (Version 4.5.2). Statistical significance was set at P value of <0.05 and was not adjusted for multiple testing (Sweeting et al., 2024). The trial is registered with the Chinese Clinical Trial Registry (ChiCTR2300079244).

References

- ACOG, 2019. Quantitative blood loss in obstetric hemorrhage: Acog committee opinion, number 794. *Obstet Gynecol*, 134(6):e150-e156. <https://doi.org/10.1097/aog.0000000000003564>
- Chang CB, Park KW, Kang YG, et al., 2014. Coexisting lumbar spondylosis in patients undergoing TKA: how common and how serious? *Clin Orthop Relat Res*, 472(2):710-717. <https://doi.org/10.1007/s11999-013-3298-7>
- Chen J, Li Q, Yu W, et al., 2025. Ventral anterior-lateral complex of the thalamus mediates chronic stress-induced pain hypersensitivity and underlies electroacupuncture analgesia. *Brain Behav*, 15(9):e70855. <https://doi.org/10.1002/brb3.70855>
- de Sena Fraga CD, de Araújo RC, de Sá L, et al., 2024. Use of a peanut ball, positioning and pelvic mobility in parturient women shortens labour and improves maternal satisfaction with childbirth: a randomised trial. *J Physiother*, 70(2):134-141. <https://doi.org/10.1016/j.jphys.2024.02.017>
- Du K, Wang X, Chi L, et al., 2017. Role of Sigma-1 receptor/p38 MAPK inhibition in acupoint catgut embedding-mediated analgesic effects in complete Freund's adjuvant-induced inflammatory pain. *Anesth Analg*, 125(2):662-669. <https://doi.org/10.1213/ANE.0000000000001857>
- Feng XM, Mi WL, Xia F, et al., 2012. Involvement of spinal orexin A in the electroacupuncture analgesia in a rat model of post-laparotomy pain. *BMC Complement Altern Med*, 12:225. <https://doi.org/10.1186/1472-6882-12-225>
- Ghoneim NI, Mansour NA, Elmaghraby SA, et al., 2022. Treatment of temporomandibular joint disc displacement using arthrocentesis combined with injectable platelet rich fibrin versus arthrocentesis alone. *J Dent Sci*, 17(1):468-475. <https://doi.org/10.1016/j.jds.2021.07.027>
- Hsieh YL, Chou LW, Hong SF, et al., 2016. Laser acupuncture attenuates oxaliplatin-induced peripheral neuropathy in patients with gastrointestinal cancer: a pilot prospective cohort study. *Acupunct Med*, 34(5):398-405. <https://doi.org/10.1136/acupmed-2016-011112>
- Lee SH, Kim SY, Kim JH, et al., 2012. Phosphoproteomic analysis of electroacupuncture analgesia in an inflammatory pain rat model. *Mol Med Rep*, 6(1):157-162. <https://doi.org/10.3892/mmr.2012.879>
- Li X, Pu Y, Li N, et al., 2025. Effect of acupoint hot compress on relieving pain in primiparous women during the latent phase of the first stage of labour: a study protocol for a prospective, multicentre, randomised controlled clinical trial. *BMJ Open*, 15(5):e094226. <https://doi.org/10.1136/bmjopen-2024-094226>
- Simon LV, Shah M, Bragg BN, 2025. Apgar score. Statpearls. StatPearls Publishing. Copyright © 2025, StatPearls Publishing LLC., Treasure Island (FL) ineligible companies. Disclosure: Manan Shah declares no relevant financial relationships with ineligible companies. Disclosure: Bradley Bragg declares no relevant financial relationships with ineligible companies.
- Singla DR, Silver RK, Vigod SN, et al., 2025. Task-sharing and telemedicine delivery of psychotherapy to treat perinatal depression: a pragmatic, noninferiority randomized trial. *Nat Med*, 31(4):1214-1224. <https://doi.org/10.1038/s41591-024-03482-w>
- Sweeting A, Enticott J, Immanuel J, et al., 2024. Relationship between early-pregnancy glycemia and adverse outcomes: findings from the TOBOGM study. *Diabetes Care*, <https://doi.org/10.2337/dc23-2214>
- Zhu Y, Wang F, Zhou J, et al., 2022. Effect of acupoint hot compress on postpartum urinary retention after vaginal delivery: a randomized clinical trial. *JAMA Netw Open*, 5(5):e2213261. <https://doi.org/10.1001/jamanetworkopen.2022.13261>

Table S1 Comparison of baseline characteristics between the participant population that completed all VAS-score assessments and the dropout population

Characteristics	Participant group	Dropout group	<i>P</i> value
<i>n</i>	495	85	
Mothers			
Age (years)	28.92±2.90	28.78±3.50	0.727
Pre-pregnancy BMI (kg/m ²)	21.60±3.26	21.41±2.77	0.612
BMI at delivery (kg/m ²)	26.66±3.29	26.00±3.27	0.091
Marital status			1.000
Married	480 (97.0%)	83 (97.6%)	
Unmarried	15 (3.0%)	2 (2.4%)	
Place of residence			0.403
City	461 (93.1%)	77 (90.6%)	
Village	34 (6.9%)	8 (9.4%)	
Maternal education			0.766
Less than college education	159 (32.1%)	28 (32.9%)	
College education	273 (55.2%)	44 (51.8%)	
Higher than college education	63 (12.7%)	13 (15.3%)	
Occupation			0.441
Paid activity	421 (85.1%)	75 (88.2%)	
Unpaid activity	74 (14.9%)	10 (11.9%)	
Cigarette smoking			1.000
Yes	3 (0.6%)	0	
No	492 (99.4%)	85 (100.0%)	
Alcohol intake			1.000
Yes	3 (0.6%)	0	
No	492 (99.4%)	85 (100.0%)	
Health insurance			0.157
Yes	459 (92.7%)	75 (88.2%)	
No	36 (7.3%)	10 (11.8%)	
Mode of conception			0.894
Spontaneous	462 (93.3%)	79 (92.9%)	
ART	33 (6.7%)	6 (7.1%)	
Induction of labor			0.956
Yes	316 (63.8%)	54 (63.5%)	
No	179 (36.2%)	31 (36.5%)	
Labor epidural analgesia			0.003
Yes	438 (88.5%)	65 (76.5%)	
No	57 (11.5%)	20 (23.5%)	
Pregnancy complications [#]			0.046
Yes	166 (33.5%)	38 (44.7%)	
No	329 (66.5%)	47 (55.3%)	
Gestational age at delivery (weeks)	39.23±1.03	39.24±0.90	0.966

Offspring			
Gender			0.731
Male	223 (45.1%)	40 (47.1%)	
Female	272 (54.9%)	45 (52.9%)	
Birth weight (g)	3248.86±341.51	3350.29±344.00	0.012
Baseline labor pain intensity based on VAS*			0.275
<i>n</i>	495	69	
Mild	229 (46.3%)	25 (36.2%)	
Moderate	240 (48.5%)	39 (56.5%)	
Severe	26 (5.3%)	5 (7.2%)	

Data was presented as mean±SD, median (25th percentile, 75th percentile), or number (percentage). Statistical analyzed by the t-test, Wilcoxon-Mann-Whitney test, Pearson Chi-Square test or Fisher's exact test. *Pregnancy complications: Included hypertensive disorders of pregnancy, intrahepatic cholestasis of pregnancy, gestational diabetes, hypothyroidism, Hashimoto thyroiditis, antiphospholipid, polycystic ovary syndrome and obesity. *VAS at 1 h after the onset of regular uterine contractions, prior to the intervention. BMI: body mass index; ART: assisted reproductive technology; SD: standard deviation; VAS: Visual Analog Scale.

Table S2 Secondary outcomes in modified intention-to-treat population limited to vaginal delivery

Outcomes	Intervention group	Control group	<i>P</i> value
<i>n</i>	252	243	
Labor duration (min)			
First stage	560 (380, 829)	530 (380, 700)	0.088
Second stage	71 (41, 120)	70 (39, 105)	0.397
Third stage	6 (4, 8)	5 (4, 8)	0.727
Total	649 (460, 885)	615 (457, 790)	0.069
Blood loss (mL)			
0 h postpartum	200 (150, 300)	200 (150, 300)	0.515
2 h postpartum	40 (35, 50)	40 (32, 50)	0.998
Apgar scores of newborns			
1 min	10 (10, 10)	10 (10, 10)	0.524
5 min	10 (10, 10)	10 (10, 10)	0.297
<i>n</i>	246	238	
Symptoms of depression based on EPDS	4 (2, 7)	5 (2, 8)	0.010

Data are presented as number (percentage). Statistical analyzed by the Wilcoxon-Mann-Whitney test. EPDS: Edinburgh Postnatal Depression Scale.

Table S3 Subgroup analyses of secondary outcomes

Outcomes	With labor epidural analgesia			Without labor epidural analgesia		
	Intervention group	Control group	<i>P</i> value	Intervention group	Control group	<i>P</i> value
<i>n</i>	200	190		28	19	
Labor duration (min)						
First stage	606 (433, 858)	590 (440, 753)	0.314	415 (310, 624)	445 (330, 650)	0.786
Second stage	73 (43, 122)	72 (43, 106)	0.554	66 (21, 101)	42 (25, 94)	0.811
Third stage	5 (4, 7)	5 (4, 8)	0.572	7 (5, 10)	8 (6, 10)	0.727
Total	698 (508, 935)	663 (523, 844)	0.281	494 (354, 670)	505 (360, 670)	0.983
Blood loss (mL)						
0 h	200 (150, 300)	200 (150, 300)	0.148	200 (150, 250)	250 (200, 300)	0.040
2 h postpartum	40 (35, 50)	40 (32, 50)	0.862	60 (46, 60)	60 (43, 60)	0.897
Symptoms of depression based on EPDS	4 (2, 7)	5 (2, 8)	0.026	4 (1, 7)	3 (6, 9)	0.162
Apgar scores of newborns						
1 min	10 (10, 10)	10 (10, 10)	0.939	10 (10, 10)	10 (10, 10)	0.225
5 min	10 (10, 10)	10 (10, 10)	0.531	10 (10, 10)	10 (10, 10)	1.000

Data are presented as median (25th percentile, 75th percentile). Statistical analyzed by the Wilcoxon-Mann-Whitney test. EPDS: Edinburgh Postnatal Depression Scale.