Original Articles

Acupuncture and Herbal Medicine

Quantitative assessment of acupuncture's effectiveness in osteoarthritis: a model-based meta-analysis

Aobo Feng[1], Jihan Huang[1,2], Yinghua Lv[1], Ling Xu[1], Yuan Liu[1], Qingshan Zheng[1,*], Luyi Wu[3,*], Lujin Li[1,2,*]

[1]Center for Drug of Clinical Research, Shanghai University of Traditional Chinese Medicine, Shanghai, China; [2]State Key Laboratory of Integration and Innovation of Classic Formula and Modern Chinese Medicine (Shanghai University of Traditional Chinese Medicine), Shanghai, China; [3]Yueyang Hospital of Integrated Traditional Chinese Medicine and Western Medicine, Shanghai University of Traditional Chinese Medicine, Shanghai, China

Abstract

Objective: This study aimed to quantitatively analyze the clinical efficacy of acupuncture for osteoarthritis and its influencing factors. It compared the effects of different types of acupuncture with sham acupuncture, as well as sham acupuncture with oral placebos, to elucidate the efficacy characteristics of acupuncture and provide quantitative data to support its use in the treatment of osteoarthritis.

Methods: Randomized controlled trials on acupuncture for osteoarthritis were retrieved from the PubMed, Cochrane Library, Wanfang, and CNKI databases, including studies published up to October 31, 2025. Using time-course and covariate models, we quantitatively assessed the clinical efficacy and influencing factors of warm acupuncture, traditional acupuncture, electroacupuncture, and sham acupuncture. Subgroup analyses were conducted to explore trends related to potential factors affecting acupuncture efficacy. Furthermore, the effects of sham acupuncture were compared with previously established oral placebo responses to evaluate whether sham acupuncture confers therapeutic benefits.

Results: A total of 72 articles involving 4,919 participants were included. The findings indicated that warm acupuncture significantly outperformed sham acupuncture in Visual Analogue Scale (VAS) pain, Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) pain, WOMAC stiffness, and WOMAC function scores. Electroacupuncture and traditional acupuncture exhibited similar efficacy, with both treatments showing numerical improvements over sham acupuncture; however, these improvements were not statistically significant. Comparisons with oral placebo responses suggested that sham acupuncture was effective, with nonacupoint sham acupuncture slightly outperforming nonpenetrating sham acupuncture. Subgroup analyses indicated that higher baseline values of outcome measures, a lower proportion of female participants, shorter disease duration, and higher acupuncture frequency were associated with greater efficacy of acupuncture for osteoarthritis.

Conclusion: This study establishes a time-course model for acupuncture in the treatment of osteoarthritis, clarifying its efficacy characteristics and clinical advantages while providing essential quantitative data for clinical practice and future trials.

Keywords: Acupuncture, Model-based meta-analysis, Osteoarthritis, Placebo, WOMAC

Graphical Abstract: https://links.lww.com/AHM/A228

Introduction

Osteoarthritis is a prevalent musculoskeletal disease[1–2] characterized by articular cartilage degeneration and subchondral bone remodeling[3–4], affecting approximately 10% of the global population[5–6]. Its symptoms—including pain and functional impairment—significantly affect daily life[7]. Although nonsteroidal anti-inflammatory drugs (NSAIDs) and opioids are frequently used for pain management[8], their long-term use can lead to serious adverse effects, such as gastrointestinal complications and cardiovascular risks[9–10]. Therefore, the development of safer and more effective treatment options is essential.

Acupuncture has demonstrated promise in the treatment of osteoarthritis[7,11]. By stimulating specific acupoints, it activates the nervous system to release endogenous analgesic substances[12], thereby alleviating symptoms. In addition, acupuncture regulates immune responses[13],

Aobo Feng and Jihan Huang are contributed equally to this work.

**Corresponding author. Lujin Li, E-mail: lilujin666@163.com; Luyi Wu, E-mail: wuluyi@shutcm.edu.cn; Qingshan Zheng, E-mail: qingshan.zheng@drugchina.net.*

How to cite this article: Feng AB, Huang JH, Lv YH, Xu L, Liu Y, Zheng QS, Wu LY, Li LJ. Quantitative assessment of acupuncture's effectiveness in osteoarthritis: a model-based meta-analysis. Acupunct Herb Med 2026;6(2):247–257. doi: 10.1097/HM9.0000000000000199

Received 31 July 2025 / Accepted 03 April 2026

Copyright © 2026 Tianjin University of Traditional Chinese Medicine. This is an open-access article distributed under the terms of the Creative Commons Attribution-Non Commercial-No Derivatives License 4.0 (CCBY-NC-ND), where it is permissible to download and share the work provided it is properly cited. The work cannot be changed in any way or used commercially without permission from the journal.

HM9_V6N2.indb 247 25-Jun-26 00:45:18

reduces inflammatory mediators, and improves joint circulation and flexibility. With minimal adverse effects and high tolerability, acupuncture is suitable for the long-term management of osteoarthritis. However, its efficacy is not strongly supported by high-quality evidence, and some researchers have attributed its effects to a substantial placebo response[14]. Furthermore, nonacupoint shallow acupuncture (sham acupuncture) may induce biological effects through cutaneous stimulation, making it a controversial placebo control[15–16].

Previous meta-analyses of acupuncture for osteoarthritis have revealed considerable heterogeneity across trials in terms of acupuncture type, treatment duration, disease severity, and control group design. Traditional meta-analyses have been inadequate in addressing these sources of variability, thereby compromising the reliability of their conclusions. Model-based meta-analysis (MBMA) can address intertrial heterogeneity by constructing efficacy models that describe the time-course relationship of interventions and quantitatively analyze factors influencing efficacy[17]. Our previous work established a model for oral placebo responses in osteoarthritis trials[18]. The present study employed MBMA to evaluate the efficacy of various types of acupuncture, including sham acupuncture, with a focus on comparisons with sham acupuncture and oral placebos[19]. This study clarifies the efficacy characteristics of acupuncture and provides quantitative data to support its use in the treatment of osteoarthritis.

Materials and methods

This study adhered to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines [Supplementary Appendix No. 1, https://links.lww.com/AHM/A229] and was registered with International Prospective Register of Systematic Reviews (PROSPERO) (CRD420251003231). All data were obtained from published literature.

Literature search strategy

We searched the PubMed, Cochrane Library, Wanfang, and CNKI databases for randomized controlled trials on acupuncture for osteoarthritis from database inception through October 31, 2025. The search terms included "osteoarthritis" and "acupuncture," and the detailed search strategies are provided in the Supplementary Appendix No. 2 (https://links.lww.com/AHM/A229). In addition, the reference lists of relevant reviews and meta-analyses were manually screened to ensure comprehensive coverage.

Inclusion and exclusion criteria

The inclusion criteria were as follows: (1) participants diagnosed with osteoarthritis; (2) interventions involving warm acupuncture, traditional acupuncture, electroacupuncture, or sham acupuncture as monotherapy; and (3) reporting baseline and post-treatment changes or change rates in Visual Analogue Scale (VAS) and Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC) scores.

The exclusion criteria were as follows: (1) participants with comorbidities; and (2) interventions combining acupuncture with other active therapies (eg, NSAIDs, physical therapy, or herbal medicine).

Data extraction

Using WPS Excel, we extracted the following data from eligible studies: (1) study characteristics (title, authors, publication year, country, trial registration number, etc); (2) trial design (group allocation, interventions, treatment protocols, sample size, duration, etc); (3) participant characteristics (age, weight, sex distribution, disease duration, baseline values of outcome measures, etc); and (4) outcomes (changes in VAS and WOMAC scores from baseline at each observation time point, as well as dropout rates). For data presented in graphical form, GetData Graph Digitizer (version 2.25.0.32) was used to extract the relevant information. Two researchers independently extracted the data, and the mean of their extracted values was used for analysis. If the discrepancy between the two extractions exceeded 2%, the data were re-extracted.

Literature quality assessment

Using the Cochrane Risk of Bias tool, we evaluated study quality across key domains, including random sequence generation, allocation concealment, blinding of participants and personnel, handling of incomplete outcome data, selective reporting, and other sources of bias. Each study was classified as having a low, high, or unclear risk of bias. Two researchers independently assessed study quality, and disagreements were resolved by a third reviewer.

Model construction and evaluation

Exploratory analyses revealed that changes in VAS and WOMAC scores following acupuncture intervention increased gradually over time and eventually reached a plateau. This pattern can be described using the classic pharmacodynamic E_{max} model[20]. In the E_{max} model, two key parameters are defined: E_{max}, representing the maximum effect, and ET_{50}, indicating the time required to achieve 50% of the maximum effect (ie, the onset time). Details of the NONMEM model code are provided in the Supplementary Appendix No. 3 (https://links.lww.com/AHM/A229). Model construction included the development of structural, random-effects, and covariate models. Model diagnostic plots[21], sampling importance resampling (SIR) methods[22], and visual predictive checks[23] were used to evaluate goodness-of-fit, parameter estimation performance, and model stability. A detailed description is provided in the Supplementary Appendix No. 4 (https://links.lww.com/AHM/A229).

Model simulation and subgroup analysis

Monte Carlo simulation[24] was used to randomly sample parameters from the distribution of typical model parameter estimates. The simulation was repeated 1,000 times to calculate typical efficacy values and corresponding 95% confidence intervals (CIs) under different

HM9_V6N2.indb 248 25-Jun-26 00:45:18

covariate levels of acupuncture treatment. Subgroup analyses were conducted for factors not included in the covariate model, including age, disease duration, baseline VAS and WOMAC scores, and sex distribution. Using Bayesian methods, efficacy parameters were adjusted to their median values, and a meta-analysis was performed to pool parameters within each subgroup. Finally, Monte Carlo simulation was applied to calculate typical efficacy values and corresponding 95% CIs for each subgroup.

Sensitivity analysis

Sensitivity analyses[25] were conducted to evaluate the robustness of the model using the following approaches: (1) exclusion of trials with sample sizes ≤30; (2) exclusion of trials with treatment durations ≤3 weeks; and (3) exclusion of trials classified as high risk of bias.

Evidence quality assessment

The certainty of the evidence was assessed from two complementary perspectives. First, considering the methodological characteristics of MBMA, we evaluated the quality of evidence regarding the effects of acupuncture and associated influencing factors based on model reliability, robustness, and the precision of parameter estimation. According to this framework, the evidence was categorized as high, medium, or low.

Second, to align with conventional systematic review standards and enhance transparency, we additionally assessed the certainty of the evidence using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) approach for each outcome (VAS pain, WOMAC pain, WOMAC stiffness, and WOMAC function) across different acupuncture types. The GRADE assessment considered five domains: risk of bias, inconsistency, indirectness, imprecision, and publication bias. Each outcome–intervention pair was rated as high, moderate, low, or very low certainty.

Software

Model construction was performed using NONMEM (version 7.4) with the FOCE-I algorithm[26–27]. Model simulation, data analysis, and graphical visualization were conducted using R (version 4.0.3), while Review Manager (version 5.4) was used for literature quality assessment[28].

Results

Data characteristics

A total of 72 articles involving 4,919 participants were analyzed, as detailed in the Supplementary Appendix No. 5 (https://links.lww.com/AHM/A229). The study screening process is presented in Figure 1. VAS pain was reported in 49 studies (68.1%), WOMAC pain in 31 studies (43.1%), WOMAC stiffness in 26 studies (36.1%), and WOMAC function in 30 studies (41.7%). The included studies were published between 1999 and 2024. Sample sizes ranged from 9 to 191 participants (median, 38), treatment durations ranged from 1.43 to 26 weeks (median, 4 weeks), participant ages ranged from 46.96 to 68.67 years (median, 59.49 years), and the proportion of female participants ranged from 35% to 95% (median, 60%). Baseline scores ranged as follows: WOMAC pain, 14.5 to 46.1 (median, 29.53); WOMAC stiffness, 5.3 to 16.6 (median, 12.0); WOMAC function, 48.4 to 142.3 (median, 78.6); and VAS pain, 3.3 to 8.3 (median, 6.4). Detailed information is provided in the Supplementary Appendix No. 2 (https://links.lww.com/

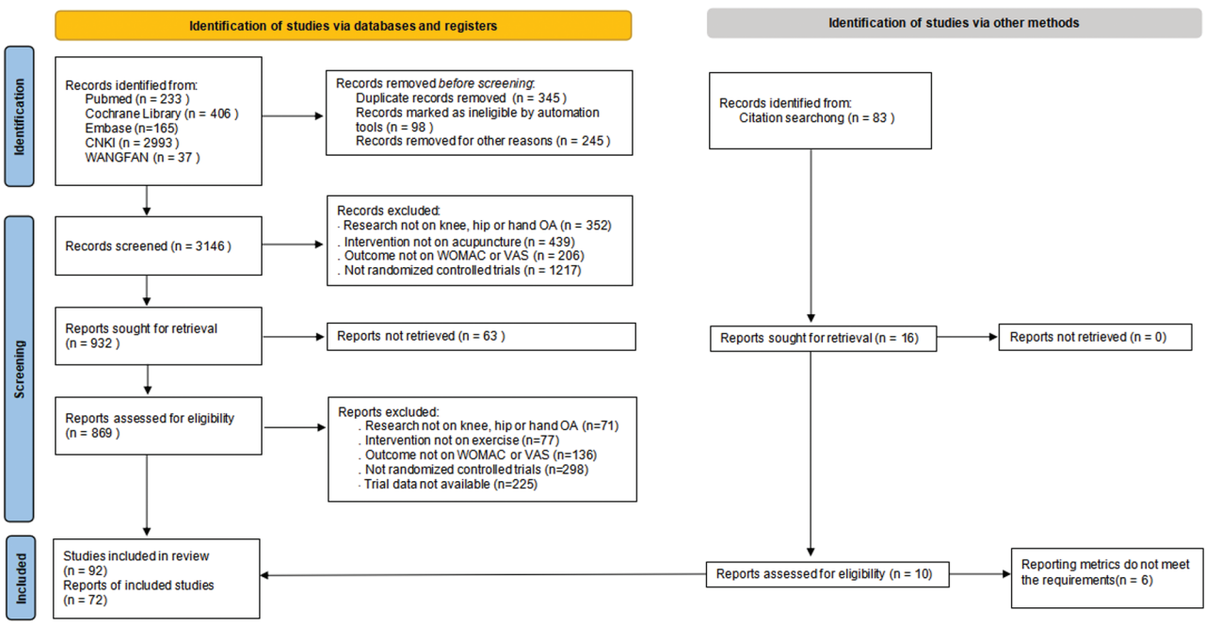


Figure 1. Flow chart of the literature selection.

HM9_V6N2.indb 249 25-Jun-26 00:45:19

AHM/A229). The included studies were generally of high methodological quality, with no evidence of significant overall bias. Quality classification was as follows: 27 studies were rated as high quality (43.5%), 22 as moderate quality (35.5%), and 13 as low quality (21.0%). The detailed quality assessment results are presented in the Supplementary Appendix Nos. 6 and 7 (https://links.lww.com/AHM/A229).

Model construction and evaluation

The E_{max} model was used to describe changes from baseline in WOMAC and VAS scores across different acupuncture groups. The plots of the model parameters and covariate correlations are presented in the Supplementary Appendix No. 8 (https://links.lww.com/AHM/A229). Covariate analysis indicated that baseline WOMAC pain, VAS pain, and WOMAC stiffness scores significantly influenced the E_{max} values of their respective models. However, no significant covariates were identified for the WOMAC function model. The detailed parameter estimates of the final model are presented in Table 1.

$$E_{max,WOMAC_pain} = E_{max,WOMAC_pain,\ typical} + (Baseline_{WOMAC_pain} - 27) \times 0.242 \quad (1)$$

$$E_{max,VAS_pain} = E_{max,VAS_{pain},\ typical} + (Baseline_{VAS_pain} - 6) \times 0.503 \quad (2)$$

$$E_{max,WOMAC_stiffness} = E_{max,WOMAC_{stiffness},\ typical} + (Baseline_{VAWOMAC-stiffness} - 12) \times 0.339 \quad (3)$$

Formulas (1)–(3) indicate that a one-point increase in the baseline WOMAC pain score corresponds to a 0.242-point increase in the E_{max} value for WOMAC pain. Similarly, a one-point increase in the baseline VAS pain score corresponds to a 0.503-point increase in the E_{max} value for VAS pain. In contrast, a one-point increase in the baseline WOMAC stiffness score corresponds to a 0.339-point decrease in the E_{max} value for WOMAC stiffness.

The model parameter estimates are listed in Table 1. The goodness-of-fit plots of the final model demonstrated adequate agreement between the model predictions and the observed data [Supplementary Appendix, No. 9, https://links.lww.com/AHM/A229]. The visual predictive check results showed that the 95% CIs of the simulated responses for each group covered most of the observed data [Supplementary Appendix, No. 10, https://links.lww.com/AHM/A229], suggesting good predictive performance. The medians of the model parameters obtained using the SIR method were close to the original estimates, indicating model stability (Table 1).

Comparison of typical efficacy across different acupuncture types

Using the final model, the time-course changes in WOMAC subscales and VAS pain scores following acupuncture were simulated. At 26 weeks, warm acupuncture, electroacupuncture, traditional acupuncture, and sham acupuncture resulted in reductions in WOMAC pain of 15.99 (95% CI, 13.77–18.23), 11.46 (95% CI, 8.99–13.96), 10.48 (95% CI, 8.23–12.78), and 8.49 points (95% CI, 6.51–10.48), respectively (Figure 2). Warm acupuncture demonstrated significantly greater efficacy than sham acupuncture. Similar trends were observed for WOMAC stiffness, WOMAC function, and VAS pain scores.

Time-course characteristics of acupuncture efficacy

Due to limited data at individual time points, a common ET_{50} value was applied across all acupuncture types. The results indicated that VAS pain, WOMAC pain, WOMAC stiffness, and WOMAC function reached 80% of E_{max} at 3, 2, 8, and 13.7 weeks, respectively. For example, with warm acupuncture, VAS pain decreased by 3.8, 4.1, and 4.2 points at 4, 8, and 12 weeks, respectively. Corresponding reductions in WOMAC pain were 14.7, 15.5, and 15.8 points; in WOMAC stiffness, 5.9, 7.1, and 7.6 points; and in WOMAC function, 38.9, 50.5, and 56.1 points (Figure 2).

Acupuncture efficacy under different covariate levels

For warm acupuncture (Figure 3), at baseline WOMAC pain scores of 15, 25, and 35, the reductions in WOMAC pain at 26 weeks were 12.5, 15.4, and 18.3 points, respectively. Similarly, at baseline VAS pain scores of 4,

Table 1

Parameter estimates and sampling importance resampling results of the final model

	VAS pain		WOMAC pain		WOMAC stiffness		WOMAC function	
Parameter	Estimate (RSE %)	SIR median (95% CI)	Estimate (RSE %)	SIR median (95% CI)	Estimate (RSE %)	SIR median (95% CI)	Estimate (RSE %)	SIR median (95% CI)
$E_{max\text{-}SA}$	1.8 (28.8)	1.6 (30.0)	8.7 (11.7)	8.0 (12.9)	4.9 (18.1)	4.5 (19.0)	31.8 (11.2)	29.9 (12.2)
$E_{max\text{-}TA}$	3.0 (10.8)	2.9 (15.7)	10.7 (11.1)	10.1 (12.4)	5.5 (13.7)	5.0 (18.6)	40.8 (13.0)	39.8 (12.9)
$E_{max\text{-}EA}$	3.3 (7.7)	3.2 (7.0)	11.7 (11.1)	11.1 (15.4)	5.8 (10.2)	5.0 (10.5)	43.0 (18.1)	47.9 (20.5)
$E_{max\text{-}WA}$	4.6 (6.4)	4.6 (7.7)	16.3 (7.0)	16.1 (8.8)	8.8 (8.4)	8.8 (10.1)	68.8 (10.3)	70.9 (10.0)
ET_{50}	0.752 (fixed)	0.752 (–)	0.5 (fixed)	0.5 (–)	2 (fixed)	2 (–)	3.4 (16.1)	3.5 (15.7)
η_{Emax} (%)	0.5 (23.5)	0.5 (27.7)	0.2 (38.6)	0.3 (42.4)	0.3 (25.5)	0.3 (29.8)	0.3 (11.1)	0.3 (14.0)
ε	0.3 (17.9)	0.3 (29.3)	0.2 (19.9)	0.3 (62.0)	0.3 (15.6)	0.3 (21.0)	23.0 (23.1)	24.7 (23.1)

η_{Emax}: Intertrial variability of the E_{max} parameter; ε: Residual error; CI: Confidence interval; $E_{max\text{-}EA}$: Maximum efficacy parameter of electroacupuncture; $E_{max\text{-}SA}$: Maximum efficacy parameter of sham acupuncture; $E_{max\text{-}TA}$: Maximum efficacy parameter of traditional acupuncture; $E_{max\text{-}WA}$: Maximum efficacy parameter of warm acupuncture; ET_{50}: Time required to achieve 50% of the maximum effect; RSE: Relative standard error; SIR: Sampling importance resampling.

HM9_V6N2.indb 250 25-Jun-26 00:45:19

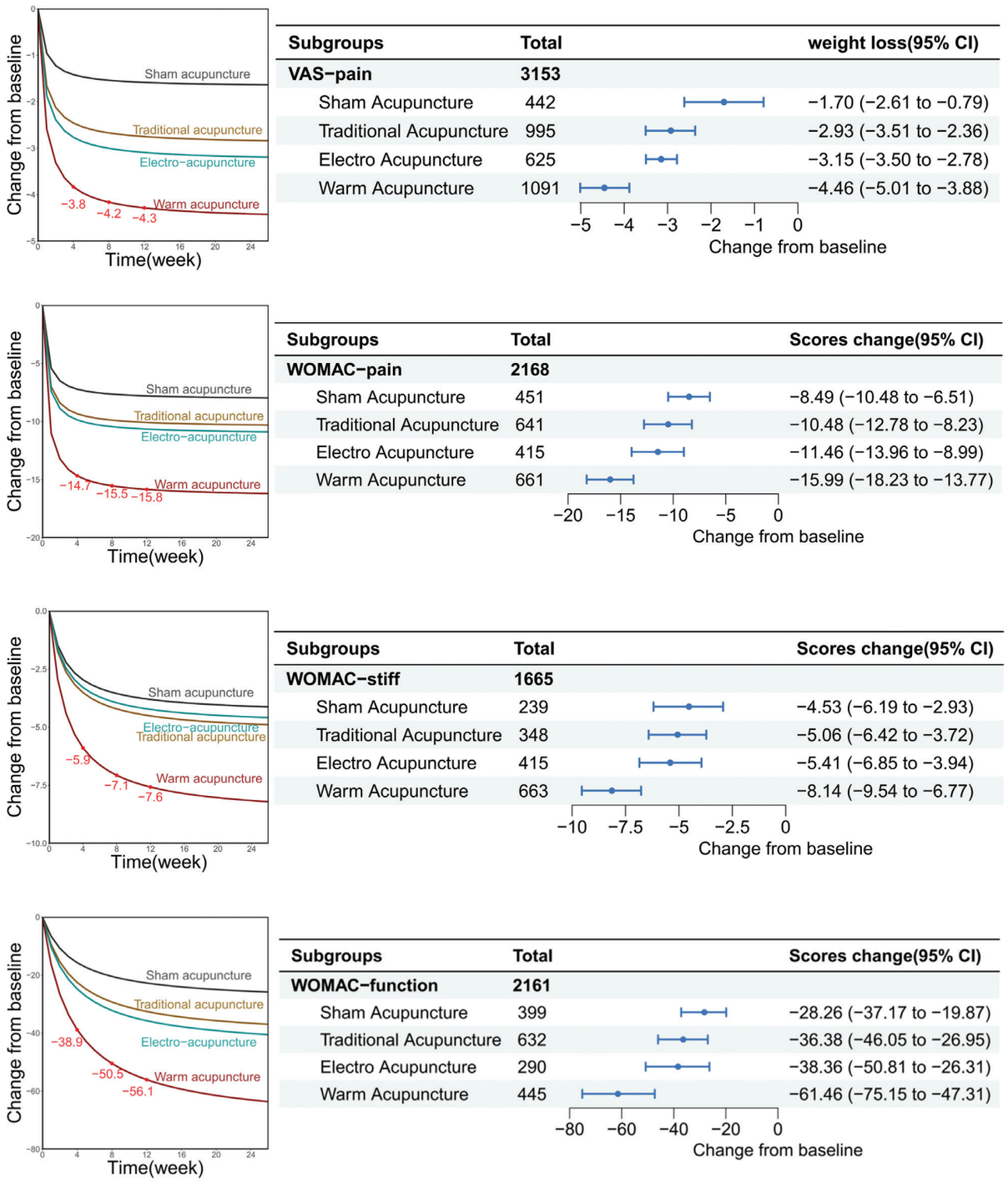


Figure 2. Change from baseline scores at 26 weeks for different types of acupuncture on WOMAC subscales and the VAS scale. Each point and bar represent the typical values along with their 95% CI. CI: Confidence interval; VAS: Visual Analogue Scale; WOMAC: Western Ontario and McMaster Universities Osteoarthritis Index.

6, and 8, the corresponding reductions in VAS pain at 26 weeks were 3.5, 4.5, and 5.4 points, respectively.

Subgroup analysis of acupuncture efficacy across different factors

To explore potential factors influencing acupuncture efficacy, subgroup analyses were conducted across all acupuncture groups. The results (Figure 4) indicated that disease duration, acupuncture frequency, and the proportion of female participants significantly affected acupuncture efficacy. Specifically, shorter disease duration, higher acupuncture frequency, and a lower proportion of female participants were associated with better outcomes.

HM9_V6N2.indb 251 25-Jun-26 00:45:21

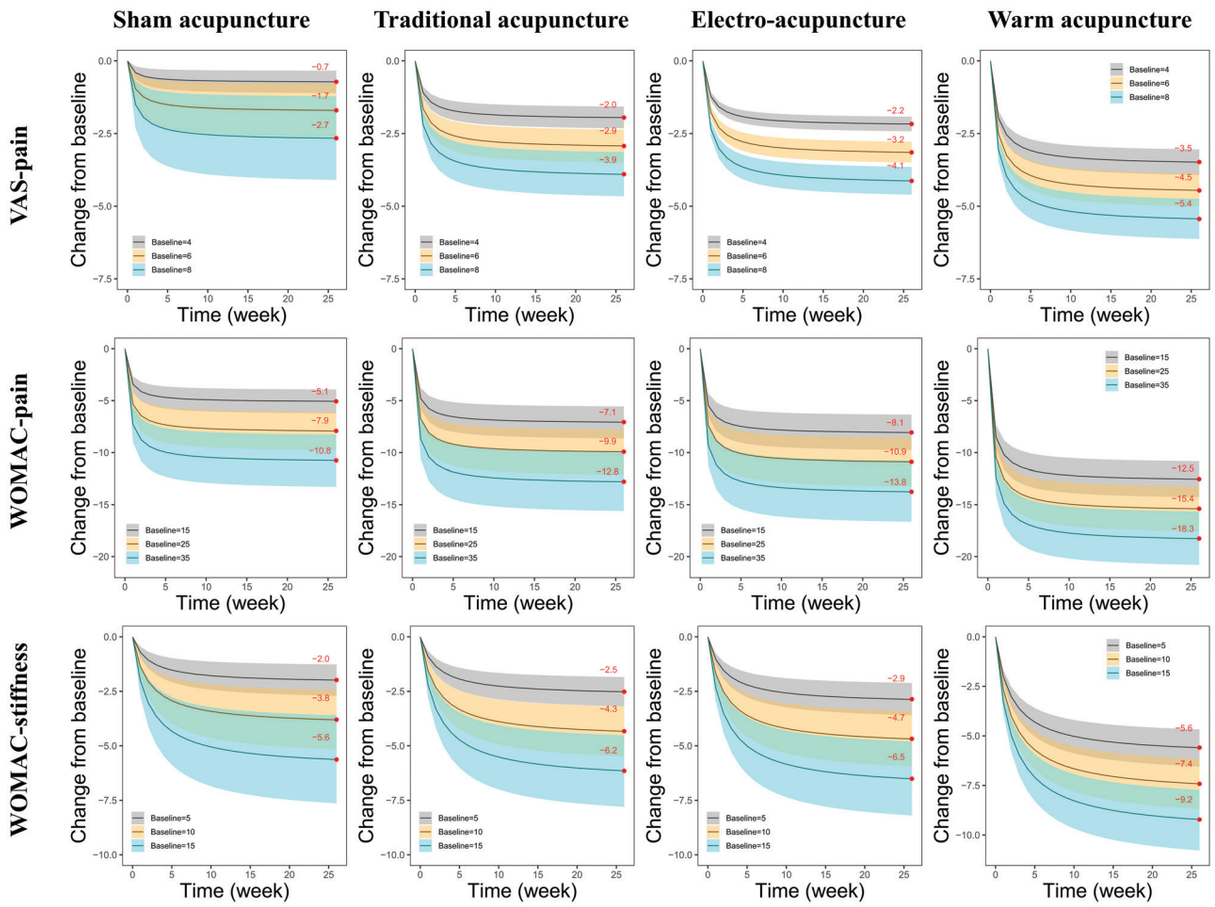


Figure 3. Typical efficacy simulation of different types of acupuncture at varying baseline levels. The solid line in the plot represents the median of the typical values, while the shaded area indicates the 95% CI. CI: Confidence interval; VAS: Visual Analogue Scale; WOMAC: Western Ontario and McMaster Universities Osteoarthritis Index.

Comparison of sham acupuncture and oral placebo responses

Our previously established model of oral placebo responses in osteoarthritis trials was used for this analysis. After adjustment of baseline values to the median levels observed in the sham acupuncture trials, sham acupuncture demonstrated greater efficacy than oral placebos (Figure 5). At 26 weeks, oral placebos reduced WOMAC pain, WOMAC stiffness, and WOMAC function by 4.70 (95% CI, 3.72–5.21), 1.70 (95% CI, 1.15–1.89), and 13.2 (95% CI, 9.14–15.0) points, respectively. In contrast, sham acupuncture resulted in reductions of 8.5 (95% CI, 5.91–9.38), 4.5 (95% CI, 2.30–4.84), and 28.2 (95% CI, 18.2–30.6) points, respectively. Overall, sham acupuncture demonstrated significantly greater efficacy than oral placebos.

Efficacy comparison of different sham acupuncture types

Subgroup analysis (Figure 6) indicated that the efficacy of nonpenetrating sham acupuncture on WOMAC subscales and VAS pain scores was nearly equivalent to that of nonacupoint sham acupuncture. However, the CIs for the efficacy of nonacupoint sham acupuncture across different outcomes were narrower, and the variability in efficacy was lower.

Sensitivity analysis

Sensitivity analyses excluding trials with sample sizes ≤30, treatment durations ≤3 weeks, and those classified as high risk of bias yielded results consistent with the primary findings, thereby confirming the robustness of the conclusions. Detailed assessment results are presented in the Supplementary Appendix No. 11 (https://links.lww.com/AHM/A229).

Evidence quality assessment

The certainty of the evidence was assessed from two complementary perspectives. From the MBMA perspective, the reliability and precision of the model parameter estimates supported warm acupuncture as high-certainty evidence for VAS pain, WOMAC pain, and WOMAC stiffness outcomes, and electroacupuncture as high-certainty evidence for VAS pain and WOMAC stiffness. In addition, the certainty of the evidence was evaluated using the GRADE approach for each outcome across different acupuncture types. According to the GRADE assessment, warm acupuncture demonstrated moderate-certainty evidence for improvements in VAS pain, WOMAC pain, WOMAC stiffness, and WOMAC function, with downgrading primarily due to indirectness (approximately 90%

HM9_V6N2.indb 252 25-Jun-26 00:45:22

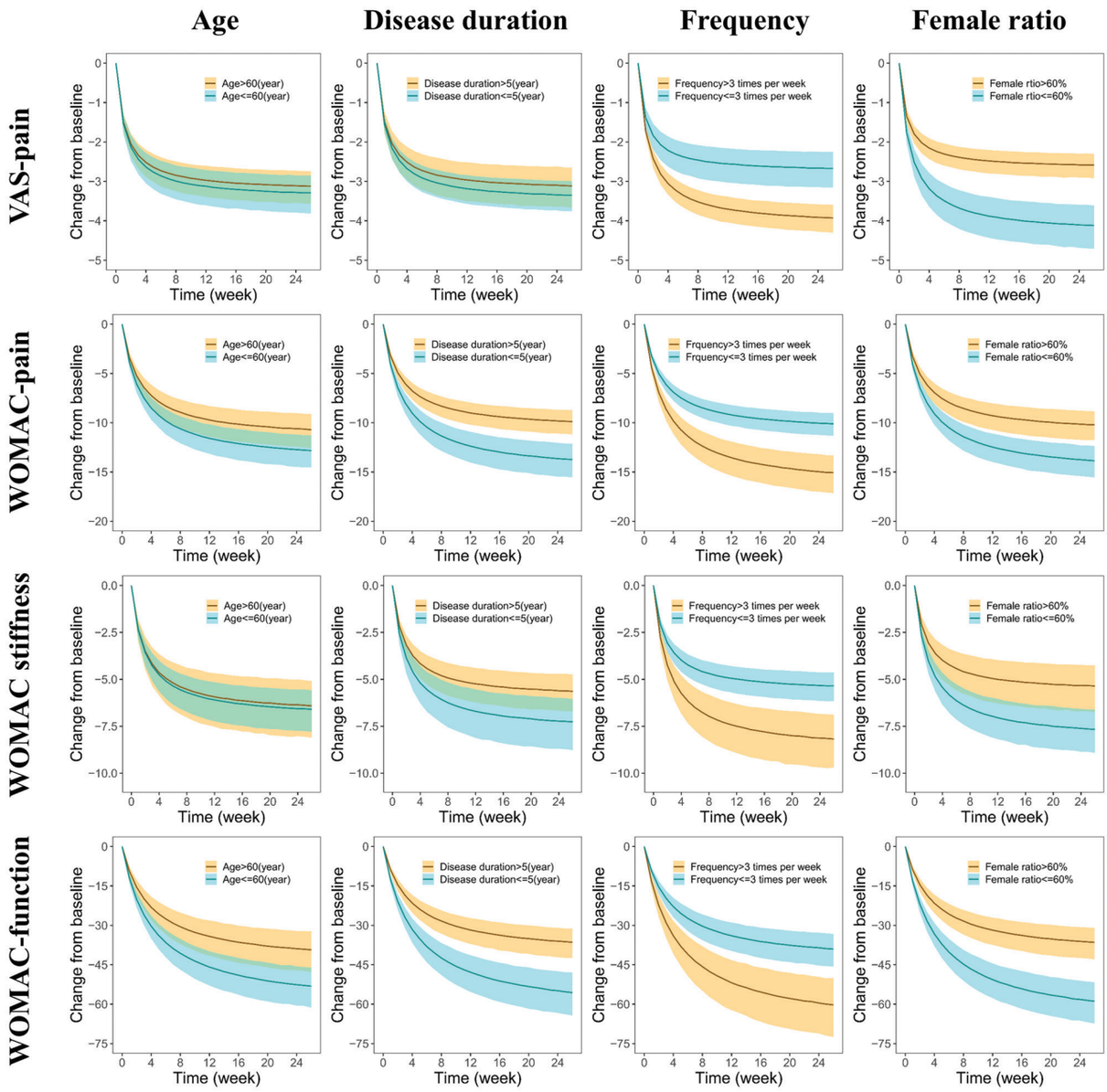


Figure 4. Subgroup analyses of changes in WOMAC scale and VAS pain scale scores from baseline across different factors. The solid line represents typical efficacy, while the shaded area indicates the 95% CI. CI: Confidence interval; VAS: Visual Analogue Scale; WOMAC: Western Ontario and McMaster Universities Osteoarthritis Index.

of participants were from Chinese populations). Electroacupuncture and traditional acupuncture generally provided low- to very low-certainty evidence across most outcomes, with downgrading due to indirectness, imprecision (CIs overlapping with sham acupuncture), and inconsistency (for WOMAC function and WOMAC stiffness). For factors influencing acupuncture efficacy (eg, disease duration, treatment frequency, and sex distribution), the certainty of the evidence ranged from moderate to low, with no high-certainty evidence identified from either assessment framework. Detailed GRADE assessment results, including the summary of findings table and reasons for downgrading each outcome–intervention pair, are provided in the Supplementary Appendix Nos. 12 and 13 (https://links.lww.com/AHM/A229).

Discussion

This study employed MBMA to comprehensively evaluate the clinical efficacy of acupuncture for osteoarthritis by distinguishing among different acupuncture types, constructing time-course models, and exploring potential influencing factors. The results indicated that warm acupuncture significantly outperformed sham acupuncture across all four outcomes. At 26 weeks, warm acupuncture achieved reductions of 4.46 (95% CI, 3.88–5.01) in VAS pain, 15.99 (95% CI, 13.77–18.23) in WOMAC pain, 8.14 (95% CI, 6.77–9.54) in WOMAC stiffness, and 61.46 (95% CI, 47.31–75.15) in WOMAC function, compared with reductions of 1.70 (95% CI, 0.79–2.61), 8.49 (95% CI, 6.51–10.48), 4.53 (95% CI, 2.93–6.19), and 28.26 (95% CI, 19.87–37.17), respectively, in the sham acupuncture group. In contrast,

HM9_V6N2.indb 253 25-Jun-26 00:45:23

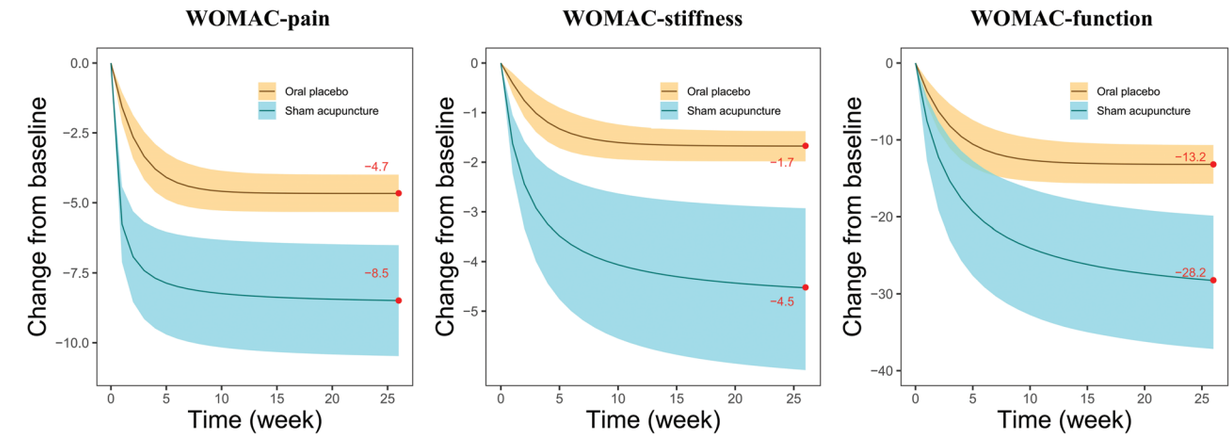


Figure 5. Comparison of efficacy between sham acupuncture and oral placebo. The solid line in the plot represents the median of the typical values, while the shaded areas indicate the 95% CI. CI: Confidence interval; WOMAC: Western Ontario and McMaster Universities Osteoarthritis Index.

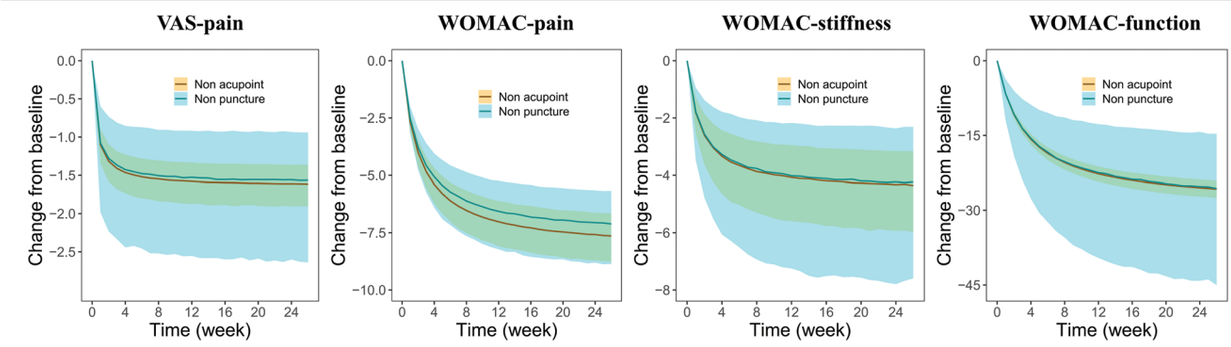


Figure 6. Comparison of efficacy between non-puncture sham acupuncture and non-acupoint sham acupuncture. The solid line in the plot represents the median of the typical values, while the shaded areas indicate the 95% CIs. Non-puncture = Non-puncture sham acupuncture (sham acupuncture that does not cause actual puncture); Non-acupoint = Non-acupoint sham acupuncture (sham acupuncture that deviates from the effective treatment of acupoints). CI: Confidence interval; WOMAC: Western Ontario and McMaster Universities Osteoarthritis Index.

electroacupuncture and traditional acupuncture exhibited comparable efficacy, with both demonstrating numerically greater improvements than sham acupuncture, although these differences were not statistically significant[29]. At 26 weeks, electroacupuncture reduced VAS pain by 3.15 (95% CI, 2.78–3.50) and WOMAC pain by 11.46 (95% CI, 8.99–13.96), compared with reductions of 2.93 (95% CI, 2.36–3.51) and 10.48 (95% CI, 8.23–12.78) for traditional acupuncture, and 1.70 (95% CI, 0.79–2.61) and 8.49 (95% CI, 6.51–10.48) for sham acupuncture. This finding contrasts with that of a recent meta-analysis reporting no significant difference between acupuncture and sham acupuncture, which may be attributable to our more detailed differentiation of acupuncture types and our explicit consideration of treatment duration and baseline heterogeneity.

Warm acupuncture[30] demonstrated significantly greater efficacy than electroacupuncture and traditional acupuncture. By combining needling with moxibustion, warm acupuncture provides thermal stimulation at acupoints, which promotes blood circulation, enhances metabolism, and modulates immune function[31]. To ensure that the model accurately assessed the specific efficacy of acupuncture, we excluded trials that combined acupuncture with other active treatments (eg, NSAIDs or physical therapy). This decision was methodologically essential. Our review identified substantial heterogeneity in the definition of "usual care" or "other treatments" across trials, ranging from basic health education to potent analgesics, which were often insufficiently specified. Current MBMA methodologies have limited capacity to adequately address the complex confounding introduced by such concomitant interventions. Based on the available published aggregate data, we were unable to reliably isolate the independent contribution of acupuncture. Therefore, although acupuncture is frequently used alongside other therapies in real-world clinical practice, this exclusion was implemented to ensure methodological rigor. We consider this issue an important direction for future research and anticipate that analyses based on individual patient data will allow a more comprehensive evaluation of the role of acupuncture within multimodal treatment strategies.

Although electroacupuncture[32] and traditional acupuncture also demonstrated advantages over sham acupuncture in pain management, these differences were not statistically significant. This finding suggests that both methods may provide some benefit in pain control. Our observation of comparable efficacy between electroacupuncture and traditional acupuncture contrasts with previous studies favoring electroacupuncture. Electroacupuncture, which combines traditional

HM9_V6N2.indb 254 25-Jun-26 00:45:24

acupuncture with electrical stimulation, represents a convenient, scalable, and relatively safe treatment modality. However, it does not involve the same degree of manual manipulation as traditional acupuncture, which may influence therapeutic outcomes. Traditional acupuncture emphasizes individualized treatment and manual techniques[33–34], requiring skilled and experienced practitioners. Thus, although electroacupuncture and traditional acupuncture each have distinct advantages and limitations, they demonstrated similar efficacy in this analysis.

We calculated the time required to reach the efficacy plateau for VAS pain, WOMAC pain, WOMAC stiffness, and WOMAC function, which were 3, 2, 8, and 13.7 weeks, respectively. These findings indicate that acupuncture alleviates WOMAC pain most rapidly, within 2 weeks, whereas improvement in WOMAC function occurs more gradually, requiring at least 13.7 weeks. Most of the included trials had treatment durations ranging from 1.43 to 26 weeks, with a median of 4 weeks, which may be insufficient to observe the maximum effects for WOMAC stiffness and WOMAC function. At 4 weeks, the efficacy reached only 50.0% and 28.7% of the maximum effect for these two outcomes, respectively. Therefore, we recommend that future acupuncture trials extend the treatment duration to at least 14 weeks.

The included trials reported baseline WOMAC pain scores ranging from 14.5 to 46.1, VAS pain scores from 3.3 to 8.3, and WOMAC stiffness scores from 5.3 to 16.6. Higher baseline WOMAC pain, VAS pain, and WOMAC stiffness scores were associated with greater treatment effects. Therefore, baseline heterogeneity should be considered in meta-analyses and cross-study comparisons. Subgroup analyses revealed that patients receiving acupuncture more than three times per week achieved significantly better outcomes than those treated less frequently. High-frequency acupuncture may provide more sustained stimulation, thereby more effectively alleviating symptoms. Moreover, subgroup analyses indicated that a higher proportion of female participants was associated with smaller treatment effects, which may be related to physiological and psychological differences between sexes[35]. Postmenopausal women experience declines in estrogen levels, which can weaken soft tissues and joint function and potentially reduce treatment responsiveness. In addition, women tend to exhibit greater sensitivity to disease symptoms and emotional responses[36], which may influence their subjective perception of treatment effects. Consequently, sex distribution should be considered during randomization in osteoarthritis trials. Furthermore, shorter disease duration was associated with greater acupuncture efficacy. A shorter disease duration may reflect milder joint damage and inflammation, thereby resulting in more pronounced treatment effects[37]. Thus, disease duration represents a critical factor in the management of osteoarthritis. Subgroup analysis by age showed no significant differences in efficacy, suggesting that age has a minimal impact on the effectiveness of acupuncture for osteoarthritis.

This study found no significant effect of baseline values on the WOMAC function model. The WOMAC function assessment included 30 trials, with baseline WOMAC function scores ranging from 48.4 to 142.3 (median, 78.6). Given the number of trials and the wide distribution of baseline values, we were unable to identify a significant covariate relationship between baseline values and WOMAC function. These findings suggest that the influence of baseline severity on WOMAC function outcomes is limited. Our previous methodological research[38] demonstrated that when the number of trials exceeds 10 and the distribution range of covariates is greater than 66.6% of the median, the accuracy of correctly identifying relevant covariates is relatively high. Therefore, this negative result also has clinical relevance, suggesting that improvement in WOMAC function following acupuncture is weakly correlated with baseline severity.

Sham acupuncture is commonly employed as a control intervention in acupuncture trials[39]. However, its implementation varies and may involve shallow needling at nontherapeutic acupoints or the use of nonpenetrating blunt needles. Some researchers have argued that sham acupuncture does not constitute a true inert control, as even needling at nontherapeutic acupoints or the application of nonpenetrating needles may elicit physiological responses. Functional magnetic resonance imaging studies[40] have demonstrated that sham acupuncture can activate brain regions such as the anterior cingulate cortex and thalamus, potentially inducing the release of bioactive substances. In our comparison of sham acupuncture with oral placebo responses, sham acupuncture demonstrated greater efficacy. At 26 weeks, sham acupuncture reduced WOMAC pain, WOMAC stiffness, and WOMAC function by 8.0, 4.1, and 25.9 points, respectively, whereas oral placebos resulted in reductions of 4.7, 1.7, and 13.2 points. Subgroup analysis showed that the therapeutic effect of nonpenetrating sham acupuncture on WOMAC subscales and VAS pain scores was nearly equivalent to that of nonacupoint sham acupuncture. In addition, nonpenetrating sham acupuncture exhibited greater between-trial variability, possibly due to limited standardization and an increased risk of compromised blinding[41].

Liu et al.[42] recently published a meta-analysis on acupuncture for the treatment of knee osteoarthritis, concluding that acupuncture is effective; however, the certainty of the evidence was rated as low. In contrast to traditional meta-analyses, our study developed a time-course model of acupuncture treatment, which enabled estimation of both the maximum effect size of acupuncture on various symptoms and the corresponding onset times. Additionally, we quantitatively evaluated potential influencing factors, thereby providing information that cannot be derived from conventional meta-analytic approaches. This expanded quantitative framework contributes to a more robust evidence base for clinical decision-making.

This study has several limitations that should be considered when interpreting the findings. First, the generally small sample sizes of the included trials may introduce potential bias and limit the generalizability of the results. Although sensitivity analyses confirmed the robustness of the primary conclusions, the limited data available for certain patient subgroups result in less precise model estimates for these populations. Consequently, caution is warranted when applying

HM9_V6N2.indb 255 25-Jun-26 00:45:24

these findings to specific subpopulations. Second, substantial heterogeneity was observed in acupoint selection protocols across the trials, reflecting the diversity of practices in real-world clinical settings. To facilitate a clearer understanding of this heterogeneity and enhance study reproducibility, we added a section entitled "Summary of Acupuncture Treatment Protocols and Deqi Response" to the Supplementary Appendix No. 14 (https://links.lww.com/AHM/A229). This section details the acupoint selections, needle retention time, treatment frequency, and deqi sensation reported in each included study. Therefore, the present results represent the average effect across different manipulation protocols, and the efficacy of specific acupoint combinations may vary. Furthermore, a key advantage of the MBMA approach is its ability to incorporate heterogeneity in study characteristics—such as follow-up duration—as informative data for modeling the time-course of treatment response, rather than treating it as a source of error to be eliminated. For example, the model's ET_{50} parameter is directly informed by studies with varying treatment durations. In addition, we directly examined the impact of heterogeneity in acupuncture frequency on efficacy through subgroup analysis (Figure 4). However, this model-based approach has limitations when applied to traditional sensitivity analyses (eg, excluding studies solely on the basis of the I^2 statistic), as such exclusions could compromise accurate estimation of time–effect parameters. To enhance transparency, we explicitly reported the range of follow-up durations and acupuncture frequencies of the included studies in the Supplementary Appendix Nos. 15 and 14 (https://links.lww.com/AHM/A229) and indirectly demonstrated the robustness of the results through sensitivity analyses that excluded studies with small sample sizes or a high risk of bias. Third, because the analysis relied on aggregated literature data, we were unable to examine many additional potential influencing factors. Future research using individual participant data is recommended to validate the present findings and explore these additional variables. Fourth, assessment of publication bias using conventional quantitative methods, such as funnel plots or the Egger test, is methodologically challenging in the context of MBMA. These approaches are designed for pairwise meta-analyses and rely on assumptions that may not hold when multiple treatment arms and time-course data are modeled simultaneously. To mitigate potential publication bias, we conducted comprehensive literature searches across multiple databases and manually screened reference lists; however, publication bias cannot be fully excluded, particularly for outcomes with fewer than 10 included studies. Finally, the generalizability of the findings may be constrained by the geographical and ethnic composition of the evidence base, as approximately 90% of the included trials were conducted in Chinese populations. Factors such as cultural differences, health care systems, and patient beliefs regarding acupuncture—all of which vary across regions—may influence treatment response. Therefore, direct extrapolation of these quantitative results to Western or other non-Chinese patient populations should be undertaken with caution.

Conclusion

This study represents the largest MBMA to date evaluating acupuncture for osteoarthritis and distinguishes the efficacy of various acupuncture types. Warm acupuncture significantly outperformed sham acupuncture, whereas traditional acupuncture and electroacupuncture exhibited comparable efficacy. Both demonstrated a numerical trend toward superiority over sham acupuncture; however, these differences were not statistically significant. Additionally, sham acupuncture demonstrated significant effects compared with oral placebos, with nonacupoint sham acupuncture performing nearly equivalently to nonpenetrating sham acupuncture. We also identified several factors influencing acupuncture efficacy, including baseline values of outcome measures, sex distribution, disease duration, acupuncture frequency, and age. These findings provide quantitative evidence to inform clinical practice and future research. To enhance generalizability, future studies should involve multicenter, multiethnic trials to validate the efficacy of acupuncture across diverse populations.

Conflict of interest statement

The authors declare no conflict of interest.

Funding

This work was supported by Shanghai Municipal Health Commission Smart Healthcare Special Project (2025ZHYL037), the National Key R&D Program of China (SQ2024YFC3500039), and the Open Research Program of the State Key Laboratory of Integration and Innovation of Classic Formula and Modern Chinese Medicine (LSLSKL20240203). The funders had no role in the study design, data collection, analysis, interpretation, or the decision to submit for publication.

Author contributions

All authors contributed to the study conception and design. Lujin Li, Yinghua Lv, and Qingshan Zheng conceived and designed the study. Aobo Feng, Jihan Huang, Yinghua Lv, Ling Xu, and Yuan Liu collected the data. Aobo Feng and Jihan Huang drafted the manuscript. Aobo Feng, Jihan Huang, Yinghua Lv, Ling Xu, Yuan Liu, Lujin Li, Luyi Wu, and Qingshan Zheng critically revised the manuscript for important intellectual content. Lujin Li, Luyi Wu, and Qingshan Zheng serve as guarantors. All authors reviewed the manuscript for intellectual content and approved the final version for submission. All authors had full access to all study data and accept responsibility for submitting the manuscript for publication.

Ethical approval of studies and informed consent

Not applicable.

Acknowledgments

None.

HM9_V6N2.indb 256 25-Jun-26 00:45:24

Declaration of generative AI in scientific writing

The conception and writing of this article were carried out collaboratively by all authors. Artificial intelligence was utilized exclusively for refining sentence coherence and amending punctuation. All citations in this manuscript were manually searched for and incorporated. After utilizing the aforementioned tools/services, the authors thoroughly reviewed and revised the content and take full responsibility for the entirety of this publication.

References

[1] Jang S, Lee K, Ju JH. Recent updates of diagnosis, pathophysiology, and treatment on osteoarthritis of the knee. *Int J Mol Sci* 2021;22(5):2619.

[2] Buckwalter J, Martin J. Osteoarthritis. *Adv Drug Deliv Rev* 2006;58(2):150–167.

[3] Gao K, Ma WZ, Huck S, et al. Association between sarcopenia and depressive symptoms in Chinese older adults: evidence from the China Health and Retirement Longitudinal Study. *Front Med (Lausanne)* 2021;8:755705.

[4] Wang Y, Li D, Lv Z, et al. Efficacy and safety of Gutong Patch compared with NSAIDs for knee osteoarthritis: a real-world multicenter, prospective cohort study in China. *Pharmacol Res* 2023;197:106954.

[5] Wei FL, Zhou CP, Gao QY, et al. Decompression alone or decompression and fusion in degenerative lumbar spondylolisthesis. *EClinicalMedicine* 2022;51:101559.

[6] Gao QY, Wei FL, Li T, et al. Oblique lateral interbody fusion vs. minimally invasive transforaminal lumbar interbody fusion for lumbar spinal stenosis: a retrospective cohort study. *Front Med (Lausanne)* 2022;9:829426.

[7] Wang H, Li M, Wang X, et al. The Notch signaling pathway in regulating bone and cartilage homeostasis: novel insights into the pathogenesis and therapeutics of osteoarthritis. *Cell Commun Signal* 2025;24(1):35.

[8] Hughes RW, Seth B, De Gray LE. Drugs used to treat joint and muscle disease. *Anaesth Intensive Care Med* 2023;24(12):806–812.

[9] Macečková Z, Pergner J, Krbec M, et al. Application of amniotic membrane in osteoarthritis management. *J Cartilage Joint Preserv* 2025;5(1):100174.

[10] Da Costa BR, Pereira TV, Saadat P, et al. Effectiveness and safety of non-steroidal anti-inflammatory drugs and opioid treatment for knee and hip osteoarthritis: network meta-analysis. *BMJ* 2021;375:n2321.

[11] Lin JG, Chen WL. Review: acupuncture analgesia in clinical trials. *Am J Chin Med* 2009;37(1):1–18.

[12] Zhang M, Shi L, Deng S, et al. Effective oriental magic for analgesia: acupuncture. *Evid Based Complement Altern Med* 2022;2022:1–14.

[13] Lou H, Yao J, Zhang Y, et al. Potential effect of acupuncture on mitochondrial biogenesis, energy metabolism and oxidation stress in MCAO rat *via* PGC-1α/NRF1/TFAM pathway. *J Stroke Cerebrovasc Dis* 2024;33(11):107636.

[14] Sun C, Xiong Z, Sun C, et al. Placebo response in sham-acupuncture–controlled trials for migraine: a systematic review and meta-analysis. *Complement Ther Clin Pract* 2023;53:101800.

[15] Ots T, Kandirian A, Szilagyi I, et al. The selection of dermatomes for sham (placebo) acupuncture points is relevant for the outcome of acupuncture studies: a systematic review of sham (placebo)-controlled randomized acupuncture trials. *Acupunct Med* 2020;38(4):211–226.

[16] Ye C, Zhou J, Wang M, et al. The effectiveness comparison of different acupuncture-related therapies on knee osteoarthritis: a meta-analysis. *Evid Based Complement Altern Med* 2022;2022:1–6.

[17] Bipat S, Zwinderman AH. Multivariate fixed- and random-effects models for summarizing ordinal data in meta-analysis of diagnostic staging studies. *Res Synth Methods* 2010;1(2):136–148.

[18] Wen X, Luo J, Mai Y, et al. Placebo response to oral administration in osteoarthritis clinical trials and its associated factors: a model-based meta-analysis. *JAMA Netw Open* 2022;5(10):e2235060.

[19] Lee B, Kwon C-Y, Lee HW, et al. The effect of sham acupuncture can differ depending on the points needled in knee osteoarthritis: a systematic review and network meta-analysis. *Heliyon* 2024;10(4):e25650.

[20] Rui M, Jiang L, Pan J-J, et al. Effects of tacrolimus on proteinuria in Chinese and Indian patients with idiopathic membranous nephropathy: the results of machine learning study. *Int Urol Nephrol* 2024;56(9):3047–3055.

[21] Oh S, Carriere KC, Park T. Model diagnostic plots for repeated measures data using the generalized estimating equations approach. *Computat Stat Data Anal* 2008;53(1):222–232.

[22] Dosne A-G, Bergstrand M, Karlsson MO. An automated sampling importance resampling procedure for estimating parameter uncertainty. *J Pharmacokinet Pharmacodyn* 2017;44(6):509–520.

[23] Ren Y, Deng C, Wang X, et al. Comparison study of model evaluation methods: normalized prediction distribution errors vs. visual predictive check. *Acta Pharmaceutica Sinica* 2011;46:1123–1131.

[24] Uyeno D. Monte Carlo simulation on microcomputers. *Simulation* 1992;58(6):418–423.

[25] Alexander ER. Sensitivity analysis in complex decision models. *J Am Planning Assoc* 1989;55(3):323–333.

[26] Sukarnjanaset W, Wattanavijitkul T, Jarurattanasirikul S. Evaluation of FOCEI and SAEM estimation methods in population pharmacokinetic analysis using NONMEM® across rich, medium, and sparse sampling data. *Eur J Drug Metab Pharmacokinet* 2018;43(6):729–736.

[27] Cha HJ, Choe K, Shin E, et al. Leveraging large language models in pharmacometrics: evaluation of NONMEM output interpretation and simulation capabilities. *J Pharmacokinet Pharmacodyn* 2025;52(3):34.

[28] Yang L, Zhang H, Shen H, et al. Quality assessment in systematic literature reviews: a software engineering perspective. *Inf Softw Technol* 2021;130:106397.

[29] Liu K, Jiao R, Ji J, et al. Placebo response in sham acupuncture therapy trials for simple obesity: a systematic review and meta-analysis. *J Integr Med* 23:264–273.

[30] Kim YJ. Clinical observation on warm acupuncture therapy for rheumatoid arthritis. *J Tradit Med Clin Naturopathy* 2017;6(3):1–3.

[31] Cui L, Bai Z. Effect of moxibustion instrument combined with intermediate frequency electrostatic therapy on pain and joint function in elderly patients with cold-dampness knee arthritis. *Comput Math Methods Med* 2022;2022:1–8.

[32] Chen Y, Xing C, Wan Q, et al. Electroacupuncture superiority in knee osteoarthritis: a meta-analysis of four acupuncture techniques. *Front Med* 2025;12:1563715.

[33] Tsui P, Leung MCP. Comparison of the effectiveness between manual acupuncture and electro-acupuncture on patients with tennis elbow. *Acupunct Electrother Res* 2002;27(2):107–117.

[34] Goh YL, Ho CE, Zhao B. Acupuncture and depth: future direction for acupuncture research. *Evid Based Complement Altern Med* 2014;2014(1):871217.

[35] Jäckle K, Kolb JP, Schilling AF, et al. Analysis of low-dose estrogen on callus BMD as measured by pQCT in postmenopausal women. *BMC Musculoskelet Disord* 2020;21(1):693.

[36] Boden MT. Prevalence of mental disorders and related functioning and treatment engagement among people with diabetes. *J Psychosom Res* 2018;106:62–69.

[37] Lillegraven S, Van der Heijde D, Uhlig T, et al. What is the clinical relevance of erosions and joint space narrowing in RA? *Nat Rev Rheumatol* 2012;8(2):117–120.

[38] Yu J, Li T, Luo J, et al. Examining the reliability of model-based meta-analysis (MBMA) outcomes: a simulation study. *CPT Pharmacometrics Syst Pharmacol* 2025;15:e70053.

[39] Chae Y, Lee J-H, Lee MS. Using the right design controls for acupuncture trials: methodological rigor and validity in research. *Integr Med Res* 2024;13(1):101026.

[40] Makary MM, Lee J, Lee E, et al. Phantom acupuncture induces placebo credibility and vicarious sensations: a parallel fMRI study of low back pain patients. *Sci Rep* 2018;8(1):930.

[41] Assefi NP, Sherman KJ, Jacobsen C, et al. A randomized clinical trial of acupuncture compared with sham acupuncture in fibromyalgia. *Ann Intern Med* 2005;143(1):10–19.

[42] Liu CY, Duan YS, Zhou H, et al. Clinical effect and contributing factors of acupuncture for knee osteoarthritis: a systematic review and pairwise and exploratory network meta-analysis. *BMJ Evid Based Med* 2024;29(6):374–384.

HM9_V6N2.indb 257 25-Jun-26 00:45:25