Randomized, Blinded, Sham-Controlled Trial of Acupuncture for the Management of Aromatase Inhibitor–Associated Joint Symptoms in Women With Early-Stage Breast Cancer

Katherine D. Crew, Jillian L. Capodice, Heather Greenlee, Lois Brafman, Deborah Fuentes, Danielle Awad, Wei Yann Tsai, and Dawn L. Hershman

ABSTRACT

Purpose
Women with breast cancer (BC) treated with aromatase inhibitors (AIs) may experience joint symptoms that can lead to discontinuation of effective therapy. We examined whether acupuncture improves AI-induced arthralgias in women with early-stage BC.

Methods
We conducted a randomized, controlled, blinded study comparing true acupuncture (TA) versus sham acupuncture (SA) twice weekly for 6 weeks in postmenopausal women with BC who had self-reported musculoskeletal pain related to AIs. TA included full body/auricular acupuncture and joint-specific point prescriptions, whereas SA involved superficial needle insertion at nonacupoint locations. Outcome measures included the Brief Pain Inventory–Short Form (BPI-SF), Western Ontario and McMaster Universities Osteoarthritis Index (WOMAC), and Modified Score for the Assessment of Chronic Rheumatoid Affections of the Hands (M-SACRAH) obtained at baseline and at 3 and 6 weeks.

Results
Of 51 women enrolled, 43 women were randomly assigned and 38 were evaluable. Baseline characteristics were comparable between the two groups. Our primary end point was the difference in mean BPI-SF worst pain scores at 6 weeks, which was lower for TA compared with SA (3.0 v 5.5; P = .003). We also found differences between TA and SA in pain severity (2.6 v 4.5; P = .003) and pain-related interference (2.5 v 4.5; P = .002) at 6 weeks. Similar findings were seen for the WOMAC and M-SACRAH scores. The acupuncture intervention was well-tolerated.

Conclusion
Women with AI-induced arthralgias treated with TA had significant improvement of joint pain and stiffness, which was not seen with SA. Acupuncture is an effective and well-tolerated strategy for managing this common treatment-related side effect.

J Clin Oncol 28:1154-1160. © 2010 by American Society of Clinical Oncology

INTRODUCTION

As a result of early detection and improved treatments, women with breast cancer are living longer. The increase in breast cancer survival is largely due to the benefits of hormonal therapy, such as tamoxifen and aromatase inhibitors (AIs), for the treatment of hormone-sensitive breast cancer. Recent clinical trials have demonstrated that AIs are more effective than tamoxifen at reducing breast cancer recurrences.1-6 In these trials, AI therapy was also associated with an incidence of musculoskeletal disorders of 5% to 35%, with nearly 5% of patients discontinuing therapy because of side effects.1-6 Observational studies, however, have shown that AI-related arthralgias are more prevalent than originally reported.7,8 In a cross-sectional survey of 200 consecutive women receiving adjuvant AI therapy for breast cancer, 94 (47%) reported AI-related joint pain and 88 (44%) reported joint stiffness.7 This musculoskeletal disorder may lead to nonadherence to this life-saving treatment and may reduce patients’ quality of life.

Acupuncture is a popular nonpharmacologic modality used for treating a variety of conditions, including musculoskeletal pain. The pain-relieving properties of acupuncture may be mediated by the release of neurotransmitters.9,10 Acupuncture has been shown to have short-term analgesic effects for pain,11,12 and clinical trials have found that patients...
Acupuncture for the Aromatase Inhibitor-Associated Joint Symptoms

with knee osteoarthritis have less pain when acupuncture is used as an adjunct to conventional pain-relieving treatments like nonsteroidal anti-inflammatory drugs. Because research in this area is limited, we conducted a small pilot study evaluating the use of acupuncture to relieve symptoms of AI-associated arthralgias.

In this randomized, cross-over study of 21 women treated with a 6-week course of acupuncture, improvements were reported in pain severity, pain-related functional outcomes, and physical well-being. This study was limited because of its small sample size and lack of an adequate control group.

Given the potential placebo effect in an unblinded study, we conducted a randomized, sham-controlled trial to assess the benefits of acupuncture in improving pain, stiffness, and functional ability in women with AI-induced arthralgias.

**Methods**

**Participants**

Women who were postmenopausal, as determined by cessation of menopause for at least 1 year or follicle-stimulating hormone more than 20 mIU/mL (if the patient had a hysterectomy and the ovaries were intact), with a history of stage I to III hormone receptor–positive breast cancer, and currently taking a aromatase inhibitor (AI) and who had a baseline worst pain score over the past week on the Brief Pain Inventory–Short Form (BPI-SF). The M-SACRAH consists of three domains assessing pain, stiffness, and functional status in patients suffering from hand osteoarthritis and rheumatoid arthritis, answered on 100-mm visual analog scales.

**Interventions**

Patients were randomly assigned to 12 true acupuncture (TA) or sham acupuncture (SA) sessions over 6 weeks. A randomization list was prepared using random permuted blocks, and consecutive assignments were placed into separate numbered, sealed envelopes. During this time, patients were permitted to take nonsteroid and nonsteroidal anti-inflammatory medications as needed. The acupuncture protocol and procedures used were devised with adherence to the Standards for Reporting of Controlled Trials in Acupuncture recommendations.

Our acupuncture protocol was selected based on a standard Traditional Chinese Medicine point prescription to treat musculoskeletal pain and the National Acupuncture Detoxification Association protocol applied to one ear to relieve pain and decrease stress. The protocol consisted of a standardized set of acupuncture points given for 30 minutes twice weekly over 6 weeks, which included full body and auricular acupuncture in alternating ears. In addition, each session included a specific point prescription tailored to up to three of the patient’s most painful joint areas. The sham intervention, which was a control for acupoint specificity, used superficial needle insertion at body points not recognized as true acupoints. The TA and SA points were chosen based on the concept of treatment manualization to facilitate the systematic delivery of standardized acupuncture treatments.

Acupuncture needles were single-use, sterile, and disposable. Full-body acupuncture needles were 25 mm or 40 mm and 34-gauge (Cloud & Dragon Medical Device, Wujing City, China) and auricular needles were 15 mm and 38-gauge (Seirin-America, Weymouth, MA). The needling protocol consisted of first swabbing all points with alcohol and needling auricular points, then needling full-body points. Needles were inserted to the proper needling depth as determined by standard point locations, and a de qi sensation was obtained at all standardized full-body acupuncture points. The needles remained in situ for 20 to 25 minutes, during which time the acupuncturist returned to stimulate the needles once, using even needle technique to re-elicit the de qi sensation. One United States–trained and New York State licensed acupuncturist performed all treatments.

The acupuncturist was the only person on the study team not blinded to the treatment assignment. Participants were blinded to treatment group. To assess the adequacy of blinding, participants gave an opinion about their treatment assignment at the end of the intervention.

**Outcome Measures**

Because there are no well-validated measures specifically for AI-induced arthralgias, we selected scales that captured joint pain, stiffness, and functional status in the knees (Western Ontario and McMaster Universities Osteoarthritis Index [WOMAC]) and hands (Modified Score for the Assessment and Quantification of Chronic Rheumatoid Affections of the Hands [M-SACRAH]) and a general pain scale used in cancer patients (BPI-SF). The BPI-SF is a 14-item questionnaire that asks patients to rate pain over the prior week and the degree to which it interferes with activities on a 0 to 10 scale. Severity is measured as worst pain, least pain, average pain, and pain right now. The severity composite score was calculated as the arithmetic mean of the four severity items, whereas the pain-related interference score was the average of seven interference items. Numeric rating scales such as the BPI are among the most common, valid, and reliable measures used to assess cancer pain.

The WOMAC version 3.1 is a validated measure for assessing osteoarthritis of the knees or hips and consists of 24 questions related to three subscales: pain (0 to 500), stiffness (0 to 200), and physical function (0 to 100). The M-SACRAH consists of three domains assessing pain, stiffness, and functional status in patients suffering from hand osteoarthritis and rheumatoid arthritis, answered on 100-mm visual analog scales.

The Functional Assessment of Cancer Therapy–General (FACT-G) measures physical, social/family, emotional, and functional well-being. The FACT scale has five response levels (commonly referred to as “very much”), where higher scores reflect better well-being and fewer symptoms. The WOMAC, BPI-SF, and FACT-G have been validated in Spanish.

Participants were asked to complete a baseline questionnaire covering demographic information and reproductive history. Participants were also surveyed on their use of complementary and alternative medicine (CAM). At baseline and at 3 and 6 weeks, self-administered questionnaires including the BPI-SF, WOMAC, M-SACRAH, and FACT-G were completed. All questionnaires were translated into Spanish. If the participant was unable to read the questionnaire because of poor vision or literacy, a bilingual study coordinator administered the survey. Follow-up assessments were conducted just before the 3-week and 6-week acupuncture sessions. Additional efficacy end points included change in analgesic use and the frequency and severity of adverse events according to the National Cancer Institute Common Terminology Criteria version 3.0.

**Statistical Analysis**

We described demographic, clinical, and outcome variables by using means and standard deviations for continuous variables and percentages for categoric variables by treatment group.

The primary objective was to compare the mean BPI-SF-worst pain score at 6 weeks between groups. All other analyses were secondary objectives. Two-sample t tests and repeated-measure analyses of variance with a time-interaction term were used to compare between-group differences and paired t tests for within-group means for the TA and SA groups for each of the outcome measures using a continuous scale. A generalized estimating equation was used to test the interaction between treatment and visit. We also measured the mean percent change from baseline for each group to account for differences in baseline measures. We evaluated the success of blinding by using χ² tests to compare the percentage of participants in each group who believed that they had received TA.

A sample size of 44 participants (22 per arm) was calculated to yield 80% power to detect a reduction of ≥ 2 points on the BPI-SF worst pain item at a 5% significance level, assuming a standard deviation of difference in score between treatment groups of 2.3 points. A reduction of ≥ 2 points on the
BPI-SF worst pain item was considered to be a clinically meaningful decrease based on the literature. For a sample size of 38 evaluable patients, we had more than 80% power to detect a 2.1-point reduction on the BPI-SF. All statistical analyses were two-sided and performed using SAS version 9.1 (SAS Institute, Cary, NC).

**RESULTS**

Between May 2006 and September 2008, 51 participants were enrolled and completed baseline questionnaires, 43 were randomly assigned, and 38 were evaluable at 6 weeks (Fig 1). The baseline characteristics of the unevaluable participants were not significantly different from the evaluable participants. The main reason for not initiating or discontinuing the intervention was due to scheduling difficulties.

Baseline demographic and clinical characteristics were comparable between the two groups (Table 1). The median age of the women enrolled was 58 years (range, 37 to 77 years). More than half of the participants were Hispanic, 30% had less than a high school education, and 68% were currently not employed. Approximately twice as many women in the TA group received adjuvant chemotherapy compared with the sham arm ($P = .02$).

The mean pain score at baseline, as measured by the BPI-SF worst pain item on a scale of 0 to 10, was 6.7 and 5.6 for the TA and SA groups, respectively (Table 2). At baseline, 14% of the entire sample reported mild joint pain (BPI worst pain of 3 to 4), 58% reported moderate pain (5 to 7), and 28% reported severe pain (8 to 10). Patients reported symptoms in the following joints: knees (87%), fingers (66%), lumbar area (11%), and other (19%). At 6 weeks, the mean BPI-SF worst pain score was 3.0 and 5.5 ($P = .002$) for the TA and SA groups, respectively (Table 2), corresponding to a 50% improvement in scores compared with baseline for the TA group (Fig 2). Similar 6-week differences were seen for pain severity ($2.59 \pm 4.53$; $P < .001$) and pain-related interference ($2.48 \pm 4.54$; $P = .002$). Eighty percent of participants in the TA group reported at least a 2-point improvement in the BPI-SF worst pain score as compared with 22% with SA ($P < .001$). A generalized estimating equation was performed and found the interaction between treatment groups and visit significant for the BPI worst pain score ($P = .0001$), as well as other secondary outcome measures (Table 2).

Similar findings were seen comparing the 6-week mean scores for the WOMAC and M-SACRAH subscales (Table 2). The TA group had up to a 70% decrease in the WOMAC and M-SACRAH scores compared with baseline (Fig 3). Physical well-being being measured by the FACT-G showed a significant improvement for TA compared with SA (19.8 v 15.4; $P = .03$) (Table 2). No significant differences were observed for the FACT-G social/family, emotional, and functional well-being subscales (Table 2).

Of 25 participants (68%) who reported taking acetaminophen or nonsteroidal anti-inflammatory drugs for pain relief, five participants (20%) reported discontinuing after 6 weeks (TA, $n = 3$; SA, $n = 2$). Initiation of narcotics or new CAM therapies was not permitted on the trial. In the post-treatment questionnaire, only three participants thought that acupuncture was moderately painful (TA, $n = 2$; SA, $n = 1$). No other adverse events were reported.

---

**Table 1. Baseline Demographic and Clinical Characteristics in the True and Sham Acupuncture Groups**

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>True Acupuncture (n = 20)</th>
<th>Sham Acupuncture (n = 18)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age, years</td>
<td>Median 58</td>
<td>57</td>
</tr>
<tr>
<td></td>
<td>Range 44-77</td>
<td>37-77</td>
</tr>
<tr>
<td>Race</td>
<td></td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>8 (40)</td>
<td>7 (39)</td>
</tr>
<tr>
<td>Hispanic</td>
<td>11 (55)</td>
<td>10 (56)</td>
</tr>
<tr>
<td>Black</td>
<td>1 (5)</td>
<td>—</td>
</tr>
<tr>
<td>Asian</td>
<td>—</td>
<td>1 (6)</td>
</tr>
<tr>
<td>Highest educational level</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Less than high school</td>
<td>4 (21)</td>
<td>7 (39)</td>
</tr>
<tr>
<td>High school graduate/some college</td>
<td>8 (42)</td>
<td>2 (11)</td>
</tr>
<tr>
<td>College or postgraduate degree</td>
<td>7 (37)</td>
<td>9 (50)</td>
</tr>
<tr>
<td>Current employment</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Employed, full-time or part-time</td>
<td>6 (30)</td>
<td>6 (33)</td>
</tr>
<tr>
<td>Not employed (unemployed/retired/homemaker)</td>
<td>14 (70)</td>
<td>12 (67)</td>
</tr>
<tr>
<td>Body mass index, kg/m²</td>
<td>Median 29</td>
<td>28</td>
</tr>
<tr>
<td></td>
<td>Range 18-45</td>
<td>23-41</td>
</tr>
<tr>
<td>Menopause</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Natural</td>
<td>11 (55)</td>
<td>8 (44)</td>
</tr>
<tr>
<td>Surgically induced</td>
<td>3 (15)</td>
<td>1 (6)</td>
</tr>
<tr>
<td>Chemically induced</td>
<td>6 (30)</td>
<td>9 (50)</td>
</tr>
<tr>
<td>Years since menopause</td>
<td>Median 7</td>
<td>6</td>
</tr>
<tr>
<td></td>
<td>Range 1-38</td>
<td>1-24</td>
</tr>
<tr>
<td>Stage of breast cancer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I</td>
<td>12 (60)</td>
<td>12 (67)</td>
</tr>
<tr>
<td>II</td>
<td>8 (40)</td>
<td>5 (28)</td>
</tr>
<tr>
<td>III</td>
<td>—</td>
<td>1 (6)</td>
</tr>
<tr>
<td>Adjuvant chemotherapy</td>
<td>16 (80)</td>
<td>7 (44)</td>
</tr>
<tr>
<td>Adjuvant taxane</td>
<td>8 (40)</td>
<td>4 (22)</td>
</tr>
<tr>
<td>Prior tamoxifen</td>
<td>8 (40)</td>
<td>5 (29)</td>
</tr>
<tr>
<td>Aromatase inhibitor therapy</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anastrozole</td>
<td>14 (70)</td>
<td>12 (67)</td>
</tr>
<tr>
<td>Letrozole</td>
<td>6 (30)</td>
<td>4 (22)</td>
</tr>
<tr>
<td>Exemestane</td>
<td>—</td>
<td>2 (11)</td>
</tr>
<tr>
<td>Duration of aromatase inhibitor therapy, months</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median</td>
<td>7 (12)</td>
<td></td>
</tr>
<tr>
<td>Range</td>
<td>3-32</td>
<td>3-22</td>
</tr>
</tbody>
</table>

*$P = .02$, according to a $\chi^2$ test comparing the true and sham acupuncture groups.*
Approximately two thirds of participants found acupuncture to be moderately relaxing and enjoyable; 74% wanted to continue treatments, 91% would recommend acupuncture to a friend, and 59% were willing to pay for acupuncture. These results did not differ between acupuncture treatment groups. Previous CAM use was high: 47% had seen at least one CAM practitioner in the past, and 83% were currently taking vitamins or supplements. Ninety percent of participants in the TA group thought that they had received TA as compared with 57% in the sham arm (P = .08). All seven patients with prior acupuncture in the past for other conditions thought they were in the TA group; however, four patients were randomly assigned to the sham arm.

## DISCUSSION

This study confirms that postmenopausal women with AI-induced arthralgias demonstrate an improvement in joint pain, stiffness, functional ability, and physical well-being after 6 weeks of acupuncture therapy and represents the first randomized controlled trial of any CAM intervention for the management of this syndrome. No significant benefits were observed in women randomly assigned to SA. In addition, the acupuncture intervention was well-tolerated, with minimal side effects.

There are several reports that acupuncture has efficacy in the treatment of noncancer-related musculoskeletal pain. Randomized trials have concluded that acupuncture is more effective than SA in the treatment of knee osteoarthritis and chronic low back pain. Our results support substantial prior evidence that acupuncture is a safe and generally well-tolerated procedure.
increasing amount of evidence suggests that acupuncture has a role in the supportive care of patients with cancer for treating chemotherapy-induced nausea and vomiting, serostomia, leukopenia, and other chemoradiotherapy-induced symptoms.

Acupuncture is often not used as a result of the concern regarding out-of-pocket expenditure associated with its use and uncertainty regarding insurance coverage; however, many third-party payers are beginning to reimburse for this therapy on the basis of evidence-based clinical trials. A cross-sectional analysis of 600,000 insurance enrollees found that only 1.3% submitted claims for acupuncture, despite coverage. In a study surveying Medicaid reimbursement specialists, 78% of Medicaid programs provide coverage for at least one alternative therapy. Factors determining whether insurers offer coverage for therapies include the potential cost-effectiveness, consumer interest, demonstrable clinical efficacy, and state mandates.

Ideally, relief from AI-associated arthralgia should be through nonhormonal mechanisms, so as not to interfere with the efficacy of the drug. Although the exact mechanisms of the analgesic effects of acupuncture are not known, studies have postulated that polymodal receptors may play a role in the mechanisms of acupuncture efficacy. Discharges of polymodal receptors in deep tissue seem to increase the intensity of de qi sensation of acupuncture. Studies have also demonstrated that acupuncture causes an increase in opioid peptides in the circulation, which can be prevented by the opioid receptor antagonist, naloxone. In addition, improvement in blood flow and other mechanotransduction-based responses to acupuncture needleling have been suggested as possible mechanisms.

The optimal control group for studies of acupuncture is controversial. SA methods include acupuncture for an unrelated condition, needle insertion at nonacupoint locations, noninsertive simulated acupuncture, or use of retractable needles. We found that 22% of subjects who received SA reported a 2-point decrease in the BPI, consistent with the literature that reports a 25% to 55% placebo effect for nonacupoint and superficial insertion sham methods. However, most acupuncture controls are not completely inert. Potential benefits may be conferred from the physiologic effect of needleling, even when not performed according to established principles. Detrimental effects from sham needleling are also possible, and other studies have found no benefit in the sham group. Given the large benefit seen in the TA group, this did not interfere with the ability of our trial to detect an effect, and although there may have been some benefit in the sham group, that would bias our results toward the null. It is possible that patients can determine their treatment assignment, which may influence their response to treatment. In our study, 57% of patients in the sham group believed they were receiving TA, whereas 90% in the TA group reported their treatment assignment correctly. It is unclear whether the patient’s perception of the treatment group may have biased our results or whether the efficacy of TA influenced their responses.

Limitations of our single-institution study include the relatively small sample size, larger than anticipated drop-out, and possible selection bias introduced by those who agree to participate in this type of study given the time requirements. In addition, the study relied on patient-reported outcomes (PROs) for the primary end point. Traditionally, the US Food and Drug Administration requires substantial evidence for reaching a conclusion that a drug or intervention will have an effect on a given condition. Evidence that shows not only a change in symptoms, but also how that change is translated into the patients’ ability to perform activities, is sometimes required by the US Food and Drug Administration for approval. Although the PRO of functional improvement was strong and consistent between the BPI-SF, WOMAC, and M-SACRAH, our study may have benefited from the use of objective measures, such as magnetic resonance imaging of joints to assess for structural changes or functional assessments, including grip strength. In addition, although improvement in the secondary outcome measures is promising, these results will need to be confirmed in larger, multicenter prospective studies.

Studies assessing AI-induced arthralgias have shown a correlation between PROs and objective findings. Morales et al
demonstrated that in a 6-month period, women taking AIs were more likely than those receiving tamoxifen to have an increase in tenosynovial changes as seen on magnetic resonance imaging, a decrease in grip strength as measured by a sphygmomanometer, as well as increased pain and stiffness as measured by self-administered questionnaire.

To our knowledge, this report is the first randomized, placebo-controlled trial establishing the use of an intervention to control AI-related joint symptoms, which should be confirmed in a larger randomized trial. With the increasing use of long-term AI therapy in the adjuvant setting, AI-induced arthralgia is becoming a major issue for breast cancer survivors. However, given the profound benefits associated with these medications, the majority of women who experience arthralgias continue with their treatment and suffer from musculoskeletal side effects. Our study suggests that acupuncture is a promising nonpharmacologic modality for relieving AI-related joint pain and stiffness. Its widespread use, however, is limited due to lack of insurance coverage for this therapy. Efforts should be aimed at improving access to this promising treatment.

AUTHORS’ DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

The author(s) indicated no potential conflicts of interest.

AUTHOR CONTRIBUTIONS

Conception and design: Katherine D. Crew, Jillian L. Capodice, Heather Greenlee, Dawn L. Hershman

Financial support: Katherine D. Crew, Dawn L. Hershman

Administrative support: Dawn L. Hershman

Provision of study materials or patients: Katherine D. Crew, Lois Brafman, Deborah Fuentes, Dawn L. Hershman

Collection and assembly of data: Katherine D. Crew, Jillian L. Capodice, Danielle Awad

Data analysis and interpretation: Katherine D. Crew, Jillian L. Capodice, Heather Greenlee, Wei Yann Tsai, Dawn L. Hershman

Manuscript writing: Katherine D. Crew, Dawn L. Hershman

Final approval of manuscript: Katherine D. Crew, Jillian L. Capodice, Heather Greenlee, Lois Brafman, Deborah Fuentes, Danielle Awad, Wei Yann Tsai, Dawn L. Hershman

REFERENCES


