Preliminary communication

Acupuncture: a promising treatment for depression during pregnancy

Rachel Manbera,*, Rosa N. Schnyerb, John J.B. Allenb, A. John Rushc, Christine M. Blaseya

a Department of Psychiatry and Behavioral Sciences, Stanford University, 401, Quarry Rd., Stanford, CA 94305, United States
b University of Arizona, AZ, United States
c Southwestern Medical Center, University of Texas, TX, United States

Received 17 March 2004; accepted 26 May 2004

Abstract

Background: Few medically acceptable treatments for depression during pregnancy are available. The aim of this randomized controlled pilot study was to determine whether acupuncture holds promise as a treatment for depression during pregnancy.

Methods: Sixty-one pregnant women with major depressive disorder and a 17-item Hamilton Rating Scale for Depression (HRSD17) score ≥14 were randomly assigned to one of three treatments, delivered over 8 weeks: an active acupuncture (SPEC, N=20), an active control acupuncture (NSPEC, N=21), and massage (MSSG, N=20). Acupuncture treatments were standardized, but individually tailored, and were provided in a double-blind fashion. Responders to acute phase treatment (HRSD17 score ≤14 and ≥50% reduction from baseline) continued the treatment they were initially randomized to until 10 weeks postpartum.

Results: Response rates at the end of the acute phase were statistically significantly higher for SPEC (69%) than for MSSG (32%), with an intermediate NSPEC response rate (47%). The SPEC group also exhibited a significantly higher average rate of reduction in BDI scores from baseline to the end of the first month of treatment than the MSSG group. Responders to the acute phase of all treatments combined had significantly lower depression scores at 10 weeks postpartum than nonresponders.

Limitations: Generalizability is limited by the small sample and its relative homogeneity.

Conclusion: Acupuncture holds promise for the treatment of depression during pregnancy.

© 2004 Elsevier B.V. All rights reserved.

Keywords: Depression; Pregnancy; Postpartum; Acupuncture; Massage

1. Introduction

Major depression during pregnancy is common, with an estimated prevalence of 3.5% (Cutrona, 1983) to 11% (Gotlib et al., 1991; Holcomb et al., 1996;
O’Hara et al., 1984; Pajulo et al., 2001). Three recent longitudinal studies documented equivalent or slightly higher depressive scores during the last 2 months of pregnancy than at 6–8 weeks postpartum (Evans et al., 2001; Hayes et al., 2001; Josefsson et al., 2001).

Despite the substantial prevalence, many cases are undetected and thus untreated, resulting in deleterious consequences to the mother and infant (such as, increased risk for postpartum depression (Beck, 2001; Chaudron et al., 2001; Cutrona, 1994; Lum, 1990; O’Hara et al., 1984; Pfost et al., 1990) and negative pregnancy outcome (Chung et al., 2001; Hedegaard et al., 1993; Lindgren, 2001; Steer et al., 1992)). The treatment of major depression during pregnancy has therefore been identified as a priority area for improvement in the clinical management of depression by both the Committee on Research on Psychiatric Treatments of the American Psychiatric Association and, independently, by the Summit on Women and Depression, organized by the American Psychological Association and the National Institute of Mental Health.

Current clinical guidelines for the pharmacological treatment of depression during pregnancy recommend carefully weighing the risks to the woman and the fetus associated with no treatment relative to the risks of treatment (American Psychiatric Association, 1993; Robert, 1996). Manufacturers of antidepressant medications advise that they be avoided during pregnancy, and many pregnant women are reluctant to undergo pharmacological treatment for their depression. Psychotherapy is a safe treatment option during pregnancy, but only interpersonal psychotherapy (IPT) has been evaluated in pregnant women (Spinelli and Endicott, 2003). Most empirically supported psychotherapies are not readily available or affordable. Consequently, there is a need for safe, effective, and affordable alternative treatments for depression during pregnancy. The little available empirical research suggests that acupuncture may hold promise for the treatment of depression (Allen et al., 1998; Roschke et al., 2000). The present randomized controlled pilot study evaluated the efficacy and safety of acupuncture as a treatment for depression during pregnancy. We hypothesized that acupuncture treatment tailored for depression symptoms would be more efficacious than either of two comparison treatments, control acupuncture and massage.

2. Methods

Sixty-one pregnant women with nonpsychotic major depressive disorder were randomly assigned to one of three treatments: an active acupuncture that specifically addressed depression symptoms (SPEC, N=20); a valid control acupuncture that did not specifically address depression symptoms (NSPEC, N=21); massage (MSSG, N=20), which provided control for attention, physical contact, relaxation, and respite from daily stress.

Acute phase treatments lasted 8 weeks and included 12 sessions (25–30 min each). Responders (HRSD17 score <14 and ≥50% reduction from baseline) entered a continuation phase and received the same treatment to which they were initially randomized (biweekly until delivery and weekly for 8 weeks post-delivery). Because women entered at different points in the pregnancy (11–28 weeks), treatment was not begun at the same week with respect to conception for each woman. Consequently, the continuation phase was of variable length and consisted of 8–17 sessions (of which 8 were postpartum). To allow for a meaningful intent-to-treat analysis of participants’ postpartum status, participants who did not receive continuation treatment were nevertheless invited to a 10-week postpartum assessment.

Acupuncture treatments were provided in a double-blind fashion (Manber et al., 2002). The blinding of acupuncturists was achieved by separating the provision of treatment from the assessment and treatment design. An assessing acupuncturist designed a set of treatments at baseline that was reassessed and updated monthly during the acute phase. A different acupuncturist, the treating acupuncturist, needled the points that were prescribed by the assessing acupuncturist. Each SPEC and NSPEC treatment consisted of the same number of acupuncture points distributed across the same general areas of the body. The treating acupuncturists and the participants were told that some treatments addressed symptoms of depression directly and others did not.

2.1. Participants

Participants were recruited from obstetric clinics and from advertisements in local parent and baby
magazines. To qualify, participants had to be 18 years or older, gestation age between 11 and 28 weeks at screening, be receiving prenatal care in the community, satisfy DSM-IV criteria for current nonpsychotic Major Depressive Episode (MDE), and score at least 14 on the 17-item Hamilton Rating Scale for Depression (HRSD$_{17}$; Hamilton, 1967). Participants were excluded for an index MDE lasting 2 years or more, psychotic features or a seasonal pattern, current active suicidal potential, cluster B Axis II disorder or other Axis I disorders in the past 2 months, except for simple phobia, social phobia, or generalized anxiety disorder [determined by the SCID-IV (First et al., 1994) and the SCID-II (First et al., 1997)]; abnormal thyroid panel; an uncontrolled medical condition, a condition that may be a medical basis for depression; current use of any medication that impacts mood; confounding treatments for depression; conditions that necessitate bed rest. Participants were not allowed to receive massage or acupuncture, other than what was provided by the study. After complete description of the study, participants provided written informed consent, approved by the Human Subjects Committee at Stanford University. A total of 88 participants underwent in-person screening interviews and 61 were randomized.

2.2. Treatments

Acupuncture treatments did not consist of a fixed set of points. Instead, treatments were individually tailored following the principles of traditional Chinese medicine. The assessment, treatment design, needle insertion, and needle stimulation were all standardized (Schnyer et al., 2001). Following conservative safety measures for acupuncture during pregnancy (Flaws, 1993), the following points were not needled during pregnancy: LI 4, Sp 1, Sp 6, GB 21, UB 60, UB 67, Ren 3, 4, 5, and 6, St 36, St 45, UB 23, UB 32, Kd 4, and GB 44, all ear points, and Ren 12 during the latter part of pregnancy. Massage was also provided in a standardized fashion. Acupuncture and massage sessions were provided at the same duration and frequency and included minimal verbal contact. All sessions were audiotaped and randomly inspected to ensure that these instructions were being followed.

2.3. Measures

The main outcome measure was response status at the end of the acute treatment phase. Response was defined jointly by (a) failure to meet full criteria for MDD; (b) at least 50% reduction from baseline of HRSD$_{17}$ score; and (c) HRSD$_{17}$$\leq$14. The HRSD$_{17}$ was administered at baseline, midpoint, end of acute treatment, and 10 weeks postpartum by interviewers who were masked to treatment assignment. The MDE portion of the SCID was administered following each HRSD interview by the same interviewer. The Beck Depression Inventory (Beck et al., 1996) was completed weekly during the acute treatment phase. Providers' and patients' expectations regarding the efficacy of treatment (Shoham-Salomon et al., 1989) were assessed after the first and third treatments.

3. Results

Among the total of 61 pregnant eligible women, the average age (33.3±4.7 years), gestation week at randomization (20.0±5.6 weeks), and baseline HRSD$_{17}$ score (21.0±4.2) did not differ significantly across groups ($p>0.5$). Caucasians constituted 75% of the sample. The level of education was high (93% had at least some college education) as was the income level (67% with a family income above $70,000). Of the 2/3 of participants who had previously been pregnant, 55% experienced depression during a previous pregnancy and 37% during postpartum. Overall, 85.5% reported a previous depressive episode, 14% stopped antidepressant medications while planning to conceive, and 19% after conception. A modified intent-to-treat sample, consisting of 54 participants with at least one post-randomization evaluation, was analyzed$^1$. Mean baseline HRSD$_{17}$ score (23.1), age (32.1 years), and gestation week (20.5) for these seven participants were within 1 standard deviation of the sample mean.

$^1$ Three participants dropped out before the first treatment (2 NSPEC and 1 SPEC). The other 4 participants dropped out due to dissatisfaction with treatment (1, MSSG) or due to pregnancy-related events (3, SPEC) that included initiation of anti-emetic treatment that impacts mood, a miscarriage, and bed-rest. There were no instances of discontinuation from treatment due to side effects.
3.1. Acute treatment response

Response status was determined using the last observation carried forward. A priori hypotheses testing revealed that a significantly larger proportion of participants responded to SPEC acupuncture treatment (68.8%) than to MSSG [31.6%, Fisher exact test, \(p=0.031\); number to treat (NTT) effect size=2.7; 95% confidence interval 1.5–16). Response rate in the NSPEC group was 47.4% and the Fisher's exact test for SPEC versus NSPEC was not statistically significant (\(p=0.18\); NTT effect size=4.7).

The proportion of participants who responded to treatment and the proportion of participants who continued to meet DSM-IV criteria for MDD at the end of the acute phase are depicted in Table 1. Analyses of variance revealed no statistically significant differences in participants’ and providers’ expectations of benefit as a function of treatment group (\(p>0.24\)).

Repeated measures analysis of variance, with treatment as a between-subjects factor and time (baseline, week 4, week 8) as a within-subjects factor, indicated a significant reduction in HRSD\(_{17}\) scores (\(F=54.9, df=2, p<0.0001\), Greenhouse–Geisser \(\varepsilon=0.94\))\(^3\). However, there was no significant treatment by time interaction (\(p>0.22\)). Fig. 1a depicts the HRSD\(_{17}\) data at the 3 time points. A random regression analysis strategy was used to evaluate differential rate of change in self-rating of symptom severity (BDI), taking advantage of the availability of BDI data at 10 time points (Fig. 1b). Given the small sample size, a simulated random regression model was used (Rogosa, 1994). Individual regression lines were fit to every subject’s data, yielding a slope to represent individual change. T-tests were then used to compare the groups on the rate of change in the weekly BDI. This analysis was performed separately for the first and second month of treatment because previous data indicate that most of the change in depression severity occurs during the first half of the acute treatment (Allen et al., 1998; Keller et al., 2000).

The SPEC group had a significantly higher average rate of improvement relative to MSSG during the first month (\(t=1.72, p=0.047\); Cohen’s \(d=0.70\))\(^3\) but not during the second month of treatment (\(p=0.083\); Cohen’s \(d=0.54\)). There were no significant differences between SPEC and NSPEC (first month: \(p=0.092\), Cohen’s \(d=0.55\); second month: \(p=0.115\), Cohen’s \(d=0.36\)).

### Table 1

Clinical status at the end of the acute phase of treatment and at 10 weeks postpartum

<table>
<thead>
<tr>
<th></th>
<th>MSSG</th>
<th>NSPEC</th>
<th>SPEC</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>End of the acute phase</strong> N=19</td>
<td>N=19</td>
<td>N=16</td>
<td></td>
</tr>
</tbody>
</table>
| Responders (n/N) | 31.6% (6/19) | 47.4% (9/19) | 68.8% (11/16)
| Meeting criteria for MDD (n/N) | 21.1% (4/19) | 26.3% (5/19) | 12.5% (2/16)
| HRSD\(_{17}\) score (SD) | 10.3 (5.6) | 12.6 (7.5) | 9.6 (7.8)
| BDI (SD) | 10.0 (4.0) | 12.2 (5.4) | 9.2 (6.1)
| **10 weeks postpartum** N=15 | N=18 | N=14 |
| HRSD\(_{17}\) score (SD) | 9.3 (6.4) | 9.5 (7.4) | 8.6 (6.5)
| BDI (SD) | 10.2 (6.6) | 10.8 (9.8) | 6.9 (7.7)
| Meeting criteria for MDD (n/N) | 21.4% (3/14) | 9.5% (2/18) | 7.1% (1/14)
| In full remission (n/N) | 66.7% (10/15) | 50.0% (9/18) | 85.7% (12/14)

MSSG=massage, NSPEC=nonspecific acupuncture, SPEC=specific acupuncture. A priori hypotheses compared SPEC to each of the two control groups.

\(^{a}\) For SPEC versus MSSG Fisher exact test, \(p=0.031\); number to treat (NTT) effect size=2.7; 95% confidence interval 1.5–16).

\(^{b}\) For SPEC versus NSPEC Fisher’s exact test, \(p=0.18\); NTT effect size=4.7.

\(^{c}\) All the participants who did not meet criteria for MDD at 10 weeks postpartum also had HRSD\(_{17}\) scores <14 and therefore no longer met study entry criteria—a measure of clinical significance (Jacobson et al., 1999).

\(^{d}\) For SPEC versus NSPEC Fisher exact test, \(p=0.039\); number to treat (NTT) effect size=2.8; 95% confidence interval 1.5–16).

\(^{2}\) See Laupacis et al. (1988). The larger the NTT, the less effective is the treatment relative to the control.

\(^{3}\) Missing data were handled using last observation carried forward (LOCF).
3.2. Postpartum depression

Repeated measures analyses of variance, with the HRSD17 (and separately with the BDI) at study baseline and at 10 weeks postpartum as the dependent variable, revealed that all three groups had a significant reduction in symptoms after delivery (HRSD17: \( F=146, df=1.44, p<0.0001 \); BDI: \( F=111, df=1.41, p<0.0001 \)). However, there was no treatment group by time interaction (BDI: \( F=1.6, df=2, 41, p=0.21 \); HRSD17: \( F=2.2, df=2, 41, p=0.12 \)). The Cohen’s \( d \) effect sizes comparing SPEC with each of the two control treatments were 0.52 for the comparison with MSSG and 0.60 for the comparison with NSPEC. Table 1 depicts clinical status at 10 weeks postpartum (HRSD17 and BDI scores) and the proportions of participants who met DSM-IV criteria for MDD and for full remission (HRSD17≤8). The latter two index the clinical significance of the results. The proportions of remitted participants in the SPEC at 10 weeks postpartum were significantly greater than in NSPEC (Fisher’s exact test, \( p=0.04, NTT=2.8 \)).

A 2×3 ANCOVA model was used to explore the role of response status at the end of acute treatment on postpartum depression severity. The model included the HRSD17 scores at 10 weeks postpartum as the dependent variable, response status at the end of the acute phase and treatment group as between-subjects factors, and baseline HRSD17 as a covariate. There was a statistically significant main effect for response status at the end of the acute phase of treatment (\( F=4.3, df=1.47, p=0.045 \)). Other effects in the model were not statistically significant (\( F=0.82, df=2.47, p=0.45 \) for the effect of treatment; \( F=0.55, df=2.47, p=0.58 \) for treatment by response status interaction; and \( F=1.9, df=1.47, p=0.18 \) for baseline HRSD17).

Bonferroni-corrected paired \( t \)-tests revealed no statistically significant group differences in delivery data (available for 49 participants), including gestation age at birth, infants’ weights (after controlling for gestation age at birth), and infants’ Apgar scores. There was also no statistically significant group difference in the proportion of participants who had vaginal delivery.

4. Discussion

This randomized controlled pilot study found a 69% response rate to acupuncture specific for depression. This rate is comparable to the response rates in clinical trials of standard treatments for depression, typically 50–70% (Elkin et al., 1989). In comparison, response rates were statistically significantly lower for massage (32%) and meaningfully lower for the control acupuncture (47%). Typically, the magnitude of response to placebo in 8-week placebo-controlled antidepressant trials is estimated at 40% (Kim and Holloway, 2003). The average rate of improvement in symptom severity (BDI) after 1 month of treatment was also significantly higher for SPEC than for MSSG.

The magnitude of symptom reduction (BDI and HRSD17) observed in the present study are comparable to those observed following 8 weeks of inter-
personal psychotherapy for depression during pregnancy (Spinelli and Endicott, 2003)\(^4\) and postpartum (O’Hara et al., 2000). Both studies demonstrated further reduction of symptoms with additional 8 weeks of treatment. Similarly, the reduction in HRSD\(_{17}\) scores following 8 weeks of specific acupuncture (from 21.5 to 9.6) are similar to those observed following 8 weeks of treatment with antidepressants (from 20.3 to 14.8) or cognitive therapy (from 20.6 to 15.7) in a mixed gender sample (Blackburn and Moore, 1997). Thus, although this pilot study was underpowered, it suggests that SPEC acupuncture produces a sizable and clinically meaningful improvement in depression during pregnancy.

The collection of postpartum data, even from participants who did not continue to receive treatment, has allowed for meaningful intent-to-treat analyses. These analyses revealed that although SPEC and NSPEC did not differ significantly at the end of the acute phase of treatment, SPEC had a longer-term advantage in that a significantly larger proportion of participants in SPEC were in full remission at 10 weeks postpartum. Most importantly, these analyses revealed that, regardless of treatment modality, early successful treatment of depression during pregnancy incurs protection against postpartum depression in this “at risk” group. This finding highlights the importance of treating depression during pregnancy not only because it can alleviate suffering during pregnancy, but also because it may prevent postpartum depression in this high-risk group.

The results of the present study cannot be interpreted as evidence that massage is not effective for depression during pregnancy because the duration of the massage treatments in the study was set to equal the length of acupuncture sessions (20 mins) and was, therefore, significantly shorter than the duration of standard massage sessions. Instead, the results demonstrate that SPEC acupuncture produced better clinical outcome than a credible control.

The most significant limitation to the generalizability of the results is the small initial sample size, which was further reduced by 7 early dropouts, yielding a modified intent-to-treat sample of 54 participants. Any pilot study, including this one, can, at best, provide initial indication that a large randomized control trial is warranted. It is, however, reassuring that even the worst-case analysis that included the 7 early dropouts found a strong trend favoring SPEC over MSSG (\(p=0.075\)). (In this worst-case analysis, all but one participant whose BDI score dropped from 35 to 9 at the last observation were classified as non-responders.) Another limitation is the relatively homogenous sample, which consisted mostly of educated Caucasian women with relatively high income. A final limitation is the ecological validity of the treatments. Typically, a single provider assesses, designs, and delivers each treatment. In contrast, in this study, the assessment and treatment were separated and reassessments were not done at each session. These two distinguishing factors have likely compromised the efficacy of SPEC. During pregnancy, when energetic changes are relatively rapid, it might be particularly beneficial to evaluate and modify the treatments more frequently than once a month.

5. Conclusion

Despite limitations, this randomized controlled pilot study indicates that acupuncture holds promise as a safe, effective, and acceptable treatment of depression during pregnancy, and that a larger clinical trial is warranted. This study also indicates that any successful treatment of depression during pregnancy incurs protection from postpartum depression.

Acknowledgements

This research was supported by AHRQ (Agency of Health Research and Quality) grant # HS09988. Special thanks are given to Robin Martin-Okada, Elizabeth Arnow, and Mary I. Huang.

References


\(^4\) The sample of Spinelli et al. had very similar baseline characteristics in terms of depression severity gestation age.