TECHNICAL COMMENTARY

NeuRA
Discover. Conquer. Cure.
POST-TRAUMATIC STRESS DISORDER LIBRARY

Acupuncture

Introduction

Acupuncture is practiced as an accepted health care model in China, Korea and Japan. Traditionally, it involves the stimulation of specific points (acupoints) by inserting needles into the skin. Electro-acupuncture is similar in that the same points are stimulated during treatment with needles inserted on specific points along the body. Electro-acupuncture uses two needles at time with the needles attached to an electrical device that generates continuous electric pulses that pass from one needle to the other with varying frequency and intensity dictated by the condition. Administration is usually for no more than 30 minutes at a time. Laser acupuncture is essentially the same except that a laser is used instead of needles. Moxibustion is a technique by which either heat from burning a specific herb (*artemisia vulgaris)* or an electric source is used to stimulate specific points or areas of the body. One of the challenges in performing efficacy trials of acupuncture is that it is difficult to provide a control condition. Sham methods that have been used include needling the wrong points or with very superficial technique or using a simulation of laser acupuncture without full stimulation.

Method

We have included only systematic reviews (systematic literature search, detailed methodology with inclusion/exclusion criteria) published in full text, in English, from the year 2010 that report results separately for people with PTSD. Reviews were identified by searching the databases MEDLINE, EMBASE, and PsycINFO. When multiple copies of reviews were found, only the most recent version was included. We prioritised reviews with pooled data for inclusion.

Review reporting assessment was guided by the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist that describes a preferred way to present a meta-analysis[1]. Reviews with less than 50% of items checked have been excluded from the library. Note that early reviews may have been guided by less stringent reporting checklists than the PRISMA, and that some reviews may have been limited by journal guidelines.

Evidence was graded using the Grading of Recommendations Assessment, Development and Evaluation (GRADE) Working Group approach where high quality evidence such as that gained from randomised controlled trials (RCTs) may be downgraded to moderate or low if review and study quality is limited, if there is inconsistency in results, indirect comparisons, imprecise or sparse data and high probability of reporting bias. It may also be downgraded if risks associated with the intervention or other matter under review are high. Conversely, low quality evidence such as that gained from observational studies may be upgraded if effect sizes are large or if there is a dose dependent response. We have also taken into account sample size and whether results are consistent, precise and direct with low associated risks (see end of table for an explanation of these terms)[2]. The resulting table represents an objective summary of the available evidence, although the conclusions are solely the opinion of staff of NeuRA (Neuroscience Research Australia).

Results

We found one systematic review that met our inclusion criteria[3].

- Moderate to low quality evidence found improvements in PTSD symptoms and some improvement in depression symptoms and functioning following acupuncture. There were no improvements in anxiety, sleep, or quality of life. Some participants experienced minor to moderate pain, superficial bleeding, and hematoma at needle insertion sites.

Margarete Ainsworth Building, Barker Street, Randwick NSW 2031. Phone: 02 9399 1000. Email: info@neura.edu.au

Grant S, Colaiaco B, Motala A, Shanman R, Sorbero M, Hempel S

Acupuncture for the Treatment of Adults with Posttraumatic Stress Disorder: A Systematic Review and Meta-Analysis

Journal of Trauma and Dissociation 2018; 19: 39-58

View review abstract online

Comparison	**Effectiveness of needle acupuncture vs. any comparator (treatment as usual, sham acupuncture, wait-list control, cognitive behavioural therapy, or paroxetine antidepressant) for symptoms in people with PTSD.** **Acupuncture ranged from 30 to 60 minutes per session, 2 to 4 sessions per week, and 3 to 12 weeks total in duration.**
Summary of evidence	**Moderate to low quality evidence (mixed sample sizes, some inconsistencies, imprecise, indirect) found improvements in PTSD symptoms and some improvement in depression symptoms and functioning following acupuncture. There were no improvements in anxiety, sleep, or quality of life. Some participants experienced minor to moderate pain, superficial bleeding, and hematoma at needle insertion sites.**

Symptoms

Large improvement in PTSD symptoms with acupuncture at post-intervention and follow-up;

PTSD symptoms post-intervention: 6 RCTs, N = 508, SMD = -0.80, 95%CI -1.59 to -0.01, $p < 0.05$, $I^2 = 90\%$

PTSD symptoms at follow-up (1-6 months): 4 RCTs, N = 387, SMD = -0.46, 95%CI -0.85 to -0.06, $p < 0.05$, $I^2 = 7\%$

Medium-sized improvement in depression symptoms with acupuncture at follow up only;

Depression symptoms post-intervention: 6 RCTs, N = 508, SMD = -0.58, 95%CI -1.18 to 0.01, $p > 0.05$, $I^2 = 82\%$

Depression symptoms at follow-up: 4 RCTs, N = 387, SMD = -0.56, 95%CI -0.88 to -0.23, $p < 0.05$, $I^2 = 0\%$

Large improvement in functioning with acupuncture at post-intervention and follow up;

Functioning post-intervention: 1 study, N = 55, SMD = -0.83, 95%CI -1.38 to -0.29, $p < 0.05$

Functioning at follow-up: 1 study, N = 55, SMD = -0.97, 95%CI -1.53 to -0.42, $p < 0.05$

There were no differences in;

Anxiety symptoms post-intervention: 4 RCTs, N = 424, SMD = -0.82, 95%CI -2.16 to 0.53, $p > 0.05$, $I^2 = 94\%$

Anxiety symptoms at follow-up: 3 RCTs, N = 332, SMD = -0.35, 95%CI -1.17 to 0.47, $p > 0.05$, $I^2 =$

Acupuncture

58% Sleep quality: 2 RCTs, N = 53, SMD = -0.46, 95%CI -3.95 to 3.03, $p > 0.05$, $I^2 = 0\%$ Physical quality of life: 1 study, N = 55, SMD = -0.47, 95%CI -1.01 to 0.07, $p > 0.05$ Mental health-related quality of life: 1 study, N = 55, SMD = -0.33, 95%CI -0.87 to 0.21 $p > 0.05$ There were no moderating effects of type of needle (traditional Chinese medicine or auricular), whether acupuncture was monotherapy or adjunctive, or whether the control condition was passive or active.	
Risks	There were few serious adverse events, some participants experienced minor to moderate pain, superficial bleeding, and hematoma at needle insertion sites.
Consistency in results‡	Some inconsistencies
Precision in results§	Imprecise
Directness of results‖	Indirect (mixed control conditions)

Explanation of acronyms

CI = confidence interval, I^2 = the percentage of the variability in effect estimates that is due to heterogeneity rather than sampling error (chance), N = number of participants, p = statistical probability of obtaining that result, RCT = randomised controlled trial, SMD = standardised mean difference, vs. = versus

Margarete Ainsworth Building, Barker Street, Randwick NSW 2031. Phone: 02 9399 1000. Email: info@neura.edu.au

Explanation of technical terms

* Bias has the potential to affect reviews of both RCT and observational studies. Forms of bias include; reporting bias – selective reporting of results; publication bias - trials that are not formally published tend to show less effect than published trials, further if there are statistically significant differences between groups in a trial, these trial results tend to get published before those of trials without significant differences; language bias – only including English language reports; funding bias - source of funding for the primary research with selective reporting of results within primary studies; outcome variable selection bias; database bias - including reports from some databases and not others; citation bias - preferential citation of authors. Trials can also be subject to bias when evaluators are not blind to treatment condition and selection bias of participants if trial samples are small[4].

† Different effect measures are reported by different reviews.

Prevalence refers to how many existing cases there are at a particular point in time. Incidence refers to how many new cases there are per population in a specified time period. Incidence is usually reported as the number of new cases per 100,000 people per year. Alternatively some studies present the number of new cases that have accumulated over several years against a person-years denominator. This denominator is the sum of individual units of time that the persons in the population are at risk of becoming a case. It takes into account the size of the underlying population sample and its age structure over the duration of observation.

Reliability and validity refers to how accurate the instrument is. Sensitivity is the proportion of actual positives that are correctly identified (100% sensitivity = correct identification of all actual positives) and specificity is the proportion of negatives that are correctly identified (100% specificity = not identifying anyone as positive if they are truly not).

Weighted mean difference scores refer to mean differences between treatment and comparison groups after treatment (or occasionally pre to post treatment) and in a randomised trial there is an assumption that both groups are comparable on this measure prior to treatment. Standardised mean differences are divided by the pooled standard deviation (or the standard deviation of one group when groups are homogenous) that allows results from different scales to be combined and compared. Each study's mean difference is then given a weighting depending on the size of the sample and the variability in the data. Less than 0.4 represents a small effect, around 0.5 a medium effect, and over 0.8 represents a large effect[4].

Odds ratio (OR) or relative risk (RR) refers to the probability of a reduction (< 1) or an increase (> 1) in a particular outcome in a treatment group, or a group exposed to a risk factor, relative to the comparison group. For example, a RR of 0.75 translates to a reduction in risk of an outcome of 25% relative to those not receiving the treatment or not exposed to the risk factor. Conversely, a RR of 1.25 translates to an increased risk of 25% relative to those not receiving treatment or not having been exposed to a risk factor. A RR or OR of 1.00 means there is no difference between groups. A medium effect is considered if RR > 2 or < 0.5 and a large effect if RR > 5 or < 0.2[5]. lnOR stands for logarithmic OR where a lnOR of 0 shows no difference between groups. Hazard ratios measure the effect of an explanatory variable on the hazard or risk of an event.

Correlation coefficients (eg, r) indicate the strength of association or relationship

Margarete Ainsworth Building, Barker Street, Randwick NSW 2031. Phone: 02 9399 1000. Email: info@neura.edu.au

between variables. They can provide an indirect indication of prediction, but do not confirm causality due to possible and often unforseen confounding variables. An r of 0.10 represents a weak association, 0.25 a medium association and 0.40 and over represents a strong association. Unstandardised (*b*) regression coefficients indicate the average change in the dependent variable associated with a 1 unit change in the independent variable, statistically controlling for the other independent variables. Standardised regression coefficients represent the change being in units of standard deviations to allow comparison across different scales.

‡ Inconsistency refers to differing estimates of effect across studies (i.e. heterogeneity or variability in results) that is not explained by subgroup analyses and therefore reduces confidence in the effect estimate. I^2 is the percentage of the variability in effect estimates that is due to heterogeneity rather than sampling error (chance) - 0% to 40%: heterogeneity might not be important, 30% to 60%: may represent moderate heterogeneity, 50% to 90%: may represent considerable heterogeneity and over this is considerable heterogeneity. I^2 can be calculated from Q (chi-square) for the test of heterogeneity with the following formula[4];

$$I^2 = \left(\frac{Q - df}{Q}\right) \times 100\%$$

§ Imprecision refers to wide confidence intervals indicating a lack of confidence in the effect estimate. Based on GRADE recommendations, a result for continuous data (standardised mean differences, not weighted mean differences) is considered imprecise if the upper or lower confidence limit crosses an effect size of 0.5 in either direction, and for binary and correlation data, an effect size of 0.25. GRADE also recommends downgrading the evidence when sample size is smaller than 300 (for binary data) and 400 (for continuous data), although for some topics, these criteria should be relaxed[6].

‖ Indirectness of comparison occurs when a comparison of intervention A versus B is not available but A was compared with C and B was compared with C that allows indirect comparisons of the magnitude of effect of A versus B. Indirectness of population, comparator and/or outcome can also occur when the available evidence regarding a particular population, intervention, comparator, or outcome is not available and is therefore inferred from available evidence. These inferred treatment effect sizes are of lower quality than those gained from head-to-head comparisons of A and B.

Margarete Ainsworth Building, Barker Street, Randwick NSW 2031. Phone: 02 9399 1000. Email: info@neura.edu.au

References

1. Moher D, Liberati A, Tetzlaff J, Altman DG, PRISMAGroup (2009): Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *British Medical Journal* 151: 264-9.
2. GRADEWorkingGroup (2004): Grading quality of evidence and strength of recommendations. *British Medical Journal* 328: 1490.
3. Grant S, Colaiaco B, Motala A, Shanman R, Sorbero M, Hempel S (2018): Acupuncture for the Treatment of Adults with Posttraumatic Stress Disorder: A Systematic Review and Meta-Analysis. *Journal of Trauma and Dissociation* 19: 39-58.
4. CochraneCollaboration (2008): Cochrane Handbook for Systematic Reviews of Interventions. Accessed 24/06/2011.
5. Rosenthal JA (1996): Qualitative Descriptors of Strength of Association and Effect Size. *Journal of Social Service Research* 21: 37-59.
6. GRADEpro (2008): [Computer program]. Jan Brozek, Andrew Oxman, Holger Schünemann. *Version 32 for Windows*