Credibility of a comparative sham control intervention for Craniosacral Therapy in patients with chronic neck pain

Heidemarie Haller\textsuperscript{a,}*, Thomas Ostermann\textsuperscript{b}, Romy Lauche\textsuperscript{a}, Holger Cramer\textsuperscript{a}, Gustav Dobos\textsuperscript{a}

\textsuperscript{a} Department of Internal and Integrative Medicine, Kliniken Essen-Mitte, Faculty of Medicine, University of Duisburg-Essen, Am Deimelsberg 34a, 45276 Essen, Germany
\textsuperscript{b} Institute for Integrative Medicine, Department of Health, Witten/Herdecke University, Gerhard-Kienle-Weg 4, 58313 Herdecke, Germany

Available online 6 October 2014

Summary

Objectives: Determining efficacy in complementary medicine research requires valid placebo/sham control groups that are credible to patients and ensure successful blinding. Within the scope of this study, a light touch sham-control intervention for trials of Craniosacral Therapy (CST) was developed and tested for its credibility.

Methods: Patients of a randomized controlled trial on chronic non-specific neck pain (NCT01526447) obtained the Credibility/Expectancy Questionnaire and the Helping Alliance/Satisfaction Questionnaire. Treatment and sham group respectively received 8 weekly sessions of CST or light touch. Data without (N = 50) and with multiple imputation (N = 54) were analyzed separately using logistic regression models. Adjusted odds ratios (AOR) and 95% confidence intervals (CI) were calculated to assess whether group outcome could be predicted from patients’ credibility ratings. An additional t-test for analysis of the overall compliance/attendance was conducted.

Results: Patients’ ratings of treatment expectancy, credibility and therapeutic alliance were not found to have significant power for classifying patients into CST or sham group (p > .05). Only satisfaction with treatment revealed a significant impact (AOR: 6.83; 95% CI: [1.54;30.24]; p = .011) in the non-imputed analysis, but not in the multiple imputation analysis (AOR: 4.09; 95% CI: [0.94;17.76]; p = .060). Compliance of both groups was not significantly different (p > .05) as were reasons for non-attendance. No serious adverse events were reported.

Conclusions: Patients’ expectancy, credibility and therapeutic alliance did not appear to affect study outcomes, blinding patients to group allocation was possible, and sham intervention was tolerable and safe. The design can therefore be recommended as control for non-specific treatment effects in future CST clinical trials.

© 2014 Elsevier Ltd. All rights reserved.
Introduction

The question about the extent of specific effects, through which therapeutic interventions achieve symptom alleviation, is currently best answered by controlling for placebo effects in randomized controlled trials. Aside from effects due to specific treatment mechanisms, there are several of non-specific and placebo effects due to psycho-physiological mechanisms associated with the treatment procedure; including contextual effects (setting, location), practitioner characteristics (gender, qualification, illness and treatment beliefs), the patient’s response to observation, assessment and interaction with the practitioner (attention, empathy) as well as the patient’s expectations and prior experiences of the treatment.\(^1\)–\(^3\) Placebo response is mediated by unconscious and conscious mechanisms of classical and operant conditioning, and by secretion of endogenous opiates, endocannabinoids, dopamine or serotonin.\(^4\)–\(^6\)

A valid placebo/sham control group should therefore mimic the intervention of interest best possible in terms of treatment duration, frequency, procedure, and therapist attention without producing specific treatment effects. Moreover, interventions being compared should generate similar levels of expectancy, i.e. the patients’ preconceptions about their clinical improvement from treatment; and credibility, i.e. the patient’s beliefs that the treatment is rational and logical for managing their condition.\(^6\) In pain research, both variables have been shown to be predictors of the study outcome, in the way that higher expectations for recovery and a higher experienced credibility are correlated with increased functional improvement.\(^5\)–\(^8\) The impact of the alliance between patient and therapist is also not negligible. Study results could even demonstrate that patients’ confidence in the treatment procedure admittedly correlated with the treatment outcome; the overall effect however depended considerably more on the quality of the therapeutic alliance.\(^9\)

Since non-pharmacological or multimodal treatments generally induce greater placebo responses than single drug trials,\(^10\) the determination of efficacy in complementary and alternative medicine (CAM) remains a methodological challenge. Studies on physical therapies that attempted to control for non-specific effects used several forms of manual placebos such as gentle manipulation, effleurage, soft tissue massage, or mild acupressure – with more or less success of generating credible control groups.\(^11\)–\(^15\) In current trials of Craniosacral Therapy (CST) study quality was quite low, mostly using observational designs or randomized controlled trials with weak controls such as waiting lists or ineffective physical treatments.\(^16\)–\(^18\)

Therefore, a two-arm randomized controlled design comparing CST to a light touch sham attention-control procedure was developed and tested for its credibility and the success of blinding participants to group allocation.

Methods

Design

This was a secondary analysis of a randomized controlled trial (NCT01526447) in chronic non-specific neck pain patients with a two-arm parallel group design of 8 weeks of treatment. The trial protocol was reviewed and approved by a national ethics committee prior to recruitment. Written informed consent was obtained from all patients. After baseline assessment and asking about their treatment expectations, patients were randomized 1:1 to groups of either CST or sham treatment using non-stratified block randomization with random varying block lengths and sealed and opaque envelopes sorted in ascending order of randomization. Credibility measures were assessed post treatment at week 8. All data were collected by blinded outcome assessors.

Setting

The study took place at the Department of Internal and Integrative Medicine at Klinikum Essen-Mitte, University of Duisburg-Essen, Germany. All observations, assessments and treatments were realized between February 2012 and May 2013. During treatment patients lay in a supine position and were asked to close their eyes; they were fully clothed, except for shoes, belts or other restrictive items.

Intervention development

According to Upledger’s conception of CST\(^19\) a standardized CST treatment and sham protocol for chronic neck pain was developed. In craniosacral theory, causes of neurophysiological and psychosomatic symptoms are founded in traumas of fascial body tissues due to accidents, injuries, postural or emotional stressors. Adhesions in cranial sutures, cerebral/spinal membranes, and related connective tissues lead to abnormal cerebrospinal fluid rhythms, which can be assessed and treated by gentle manual palpation techniques. Releasing tissue restrictions in the whole body and eliminating stasis within the dynamic system between cranium and sacrum seek to balance sensory, motor, cognitive and emotional aspects of patients’ central and autonomic nervous system.

For both groups, treatment protocols comprised 8 individual sessions once a week lasting for 45 min each. Number and duration of the offered sessions based on previous research.\(^15\) At each session, subjects of both groups received initial structural CST examination including assessment of recent symptoms, general and local listenings for connective tissue restrictions, and evaluation of the cerebrospinal fluid/craniosacral rhythm. Possible adverse events of the last session were also enquired. The CST group was then treated using standardized application of gentle fascial traction, release and unwinding techniques\(^19\)–\(^20\) on the basis of located restrictions (Table 1). Dialog techniques for reinforcing body awareness as part of the somato-emotional release were also used.\(^21\) Thereby therapists followed a treatment algorithm, in which each further step of the therapy path depended on the previous outcome, because a strict equal procedure for all patients is neither feasible nor expedient.\(^9\)–\(^12\) During the sham sessions, which were designed to be believable but not specifically effective, the same anatomic regions as in the CST group were treated, but only lightly touching them without therapeutic intention for 2 min each time.\(^13\)–\(^22\) Subjects of the sham group were
also asked to listen into their body for pleasant or unpleasant physical reactions. At the end of each session, global CST assessment was repeated for both groups and treatment steps were recorded by the therapist using a structured log.

Practitioner characteristics

In contrast to other studies investigating CST,18,24 more than one therapist took part in treating patients, although weak inter-rater reliability has been reported for craniosacral assessment techniques.25–27 The multi-therapist design was chosen with respect to higher generalizability intending not to evaluate the therapist effect but the therapy effect. The 4 practitioners were female and licensed physical therapists with advanced CST qualification and on average 6 years of clinical practice. Indeed, moving between therapists was not envisaged, because of possible dilution of study effects.23 From day one, therapists were involved in the development of treatment protocols and received a further training to standardize CST and sham techniques.

Patient characteristics

The study population included patients with chronic non-specific neck pain for at least 3 months with at least moderate pain intensity of ≥45 mm on a 100 mm visual analog scale (VAS), aged 18–65 years. Whereas patients were blinded to treatment allocation, only naive subjects were included to prevent systematic bias due to de-blinding. Study patients were also not aware of the fact that one group will receive a sham treatment; instead they were told that two different CST techniques would be tested. This approach was recommended for generating valid sham procedures in manual therapy trials.28 So the description of the study interventions was conceived to objectively inform patients about the procedures, without overtly labeling them.29 Exclusion criteria for study participation were specific neck pain due to degenerative, inflammatory or neurological diseases, physical trauma or neoplasms of the spine. Patients with severe comorbid somatic or psychiatric disorders, current pregnancy, intake of corticosteroids, opiates, muscle relaxants or antidepressants, or recently starting another medication or invasive/manipulative treatment of the spine were also excluded.

Assessment of credibility measures

The credibility of the sham protocol was evaluated by two self-report measures. Firstly, the Credibility/Expectancy Questionnaire (CEQ), which has shown high internal consistency and test–retest reliability,6,30 was used in an adapted form.31 Three of originally six items were selected to assess treatment expectations, the logical consistency of the treatment and the willingness for recommending the treatment to a friend. A principal component analysis confirmed the original two-factor solution, in which the first item had higher loadings on the factor expectancy (0.99) and the second and third items on the factor credibility (0.96 for both items). The three questions with a rating scale from 1 (not at all) to 9 (very much) read as follows:

1. How confident are you that Craniosacral Therapy will be successful in reducing your neck pain symptoms?
2. How logical does Craniosacral Therapy seem to you for treating your limitations due to neck pain?
3. How confident would you be in recommending Craniosacral Therapy to a friend with the same problems?

Secondly, patients obtained the Helping Alliance Questionnaire (HAQ)32 to assess their interaction with the practitioner and control for attention effects; also a valid and reliable adaption of altogether 11 items that were answered on a 6-point rating scale from strongly disagree to strongly agree.33 The values were then summarized to the two subscales: quality of the therapeutic relation and satisfaction with the treatment. As satisfaction always might influence credibility and alliance ratings, this variable was considered, too.

To evaluate patients’ compliance with their allocated treatment option, the number of attended sessions was assessed for both groups, as well as reasons for non-attendance.

Data analysis

For the analyses of baseline data, drop-out characteristics and patients’ compliance chi-square tests and independent samples t-tests were computed.

To analyze the impact of credibility of the sham protocol a logistic regression model was applied on the treatment group variable with ”treatment expectancy”, “logical
consistency of treatment”, “recommendation of treatment”, “relation to therapist” and “satisfaction with treatment” as independent continuous predictors. To assess whether treatment group could be predicted from these variables, adjusted odds ratios (AOR) and 95% confidence intervals (CI) were calculated for all included variables. The Hosmer–Lemeshow goodness-of-fit test was used to assess how well the chosen model fits the data. Logistic regression was carried out on both the original dataset (all cases with complete data) and a pooled dataset (intention-to-treat population) using multiple imputation techniques based on a multivariate regression model of baseline values and socio-demographic parameters. Missing predictor values were imputed 20 times. The resulting data sets were then analyzed separately and combined into a final point estimate of the AORs. To compare the results of both analyses AORs and 95% CI were displayed graphically.

All analyses were performed using SPSS software (version 20.0, IBM).

Results

Patients

A total of 54 patients (81.5% female) with an average age of 44.6 ± 10.0 years were randomized. At baseline, no significant differences between study groups were found regarding age, gender, educational level, neck pain duration, or pain intensity (p > .05). Fifty patients of the initial sample completed measurement at week 8; data from two CST patients and two sham patients were missing due to reasons of scheduling or loss of interest and had to be imputed. The four drop-outs did not significantly differ from completers in terms of baseline variables described above (p > .05) (Table 2).

Credibility of protocols

The results of the logistic regression model are displayed in Fig. 1. The AORs of “treatment expectancy”, “logical consistency of treatment”, “recommendation of treatment” and “relation to therapist” were not found to have significant predictive power for classifying patients into CST or sham group in both datasets (Table 3). Only the predictor “satisfaction with treatment” revealed a significant AOR of 6.83 (95% CI: [1.54;30.24]; p = .011) in the complete data analysis. However this AOR was not significant for the

Table 2 Sample characteristics at baseline (mean ± standard deviation).

<table>
<thead>
<tr>
<th></th>
<th>Craniosacral Therapy (n = 27)</th>
<th>Sham (n = 27)</th>
<th>p value</th>
<th>Completers (n = 50)</th>
<th>Drop-outs (n = 4)</th>
<th>p value</th>
<th>Total (n = 54)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (years)</td>
<td>44.2 ± 9.7</td>
<td>45.0 ± 10.5</td>
<td>.769</td>
<td>44.6 ± 10.0</td>
<td>45.0 ± 11.5</td>
<td>.934</td>
<td>44.6 ± 10.0</td>
</tr>
<tr>
<td>Gender (f/m) (%)</td>
<td>70.4/29.6</td>
<td>92.6/7.4</td>
<td>.076</td>
<td>80.0/20.0</td>
<td>100/0</td>
<td>.322</td>
<td>81.5/18.5</td>
</tr>
<tr>
<td>Education (&lt;hs/hs/ud) (%)</td>
<td>33.3/29.6/37.1</td>
<td>34.6/42.3/23.1</td>
<td>.153</td>
<td>34.0/34.0/32.0</td>
<td>33.3/66.7/0</td>
<td>.553</td>
<td>34.0/35.8/30.2</td>
</tr>
<tr>
<td>Duration of pain (years)</td>
<td>9.9 ± 9.2</td>
<td>9.3 ± 8.8</td>
<td>.809</td>
<td>9.9 ± 9.0</td>
<td>4.7 ± 4.6</td>
<td>.328</td>
<td>9.6 ± 8.9</td>
</tr>
<tr>
<td>Pain intensity (VAS)</td>
<td>64.1 ± 12.8</td>
<td>64.4 ± 13.3</td>
<td>.942</td>
<td>6.4 ± 1.3</td>
<td>6.6 ± 3.1</td>
<td>.785</td>
<td>64.3 ± 12.9</td>
</tr>
</tbody>
</table>

F = female; hs = high school; m = male; n = sample size; ud = university degree.

Figure 1 Adjusted odds ratios and 95% confidence intervals of the logistic regression model.

pooled analysis of the imputed dataset (AOR: 4.09; 95% CI: [0.94;17.76]; p = .060).

Compliance with and safety of protocols

For both groups similar levels of compliance were found. Patients of the CST group attended on average 7.7 ± 0.9 sessions (96.25%), patients of the sham group 6.7 ± 2.4 sessions (83.75%). Although this group difference was not statistically significant (p = .054) twice as many patients of the sham group as of the CST group did not participate in all treatment sessions. In the sham group following reasons for intermittent attendance were reported: scheduling or personal problems by four patients, infectivity of the treatment by three patients, and symptom worsening by one patient. In comparison, three patients of the CST group stated scheduling problems and one patient symptom worsening as the reasons for not attending the whole number of sessions provided. However, none of the patients reported serious adverse events neither through CST nor through sham treatment.

Discussion

To facilitate efficacy research it is imperative to test CST against a valid and credible sham intervention with successful patient-blinding; even if previous studies queried the use of manual sham procedures for clinical trials of CST.31,34 The results of the present study illustrate that the
The implementation of a valid sham treatment was feasible and tolerable for patients. None of the expectancy, credibility or alliance measures was found to be a significant predictor of group affiliation. Only satisfaction with the allocated treatment seemed to have certain influence, but was not stable for the intention-to-treat analysis, which sought to adjust possible systematic bias due to drop-outs. The good compliance of both CST and sham patients corroborates the tolerability of the used sham protocol. Even if the attendance was 12.5% less than in the CST group, this difference did not reach the level of significance. Reasons for canceled appointments were also similar between groups. Serious adverse events were not reported. In general, results suggest that rather other factors than lacking credibility of the sham protocol were related to patients’ overall compliance. Only for the three withdrawals, who reasoned ineffectivity of the sham treatment, blinding might not succeeded. The common objection that the implementation of a sham protocol would be inherently difficult for therapists could not be observed. Thus, qualified craniosacral therapists may be equivalent to trained actors, who only pose as a practitioner while doing sham.21

However, study results should be read in the light of some limitations. First of all, measurement tools were partly used in a non-validated form. As in previous trials,13 the credibility instrument (CEQ) was adapted, but not revalidated before application to study patients. The performed component analysis admittedly confirmed the original factor structure, but was only executed post hoc. In addition, conclusions about blinding success could only be derived indirectly from collected data, because patients were not asked directly about their expected group assignment. Second, the comparability with similar study designs might be limited in consequence of inconsistent timing of credibility measurement. Recommendations varied from early-stage assessments before randomization, administration after exposure of the first treatment unit and end-of-trial tests.15 In this study, patients’ expectations about their clinical improvement were obtained before the first treatment experience and before randomization to separate them from credibility beliefs and possible treatment effects. Credibility and alliance/satisfaction measures were obtained post treatment taking into account that unblinding may occur during the whole study period.15 But although satisfaction was not statistically predictive in the intention-to-treat analysis, it might be closely related to treatment success16 and thereby also to credibility, which argues against the use of end-of-trial assessment for this variable. Third, generalizability of the results to other patient populations including more men or being not naive to CST might be reduced, although there are examples of successful blinding non-naive patients to sham, where light touch was used to control for non-specific effects of osteopathic manipulation.13

While other control procedures used in CST clinical trials have been provably inappropriate21 or not yet tested23,24 the examined sham protocol is the first one that complies with the CST protocol in terms of duration, frequency, procedure and therapist attention; and that is safe and credible to patients as well as enables to interpret study results independently of the alliance between patient and therapist. As no significant differences between light touch sham manipulation and a no treatment group were reported in prior studies,23 the two-arm sham design can be applied on future CST efficacy trials. Besides, researchers should always attempt to blind patients to group allocation as this was a core component of a low risk of bias appraisal37 and should also not miss to assess patients’ expectancy and credibility of the used treatment groups as well as the quality of the therapeutic alliance.

### Conclusions

Study results indicate that patients’ expectancy, credibility and alliance to the therapist did not appear to systematically affect study outcomes, blinding patients to treatment allocation was possible and sham manipulation was tolerable and safe. Therefore, the design can be regarded as a credible means to control for placebo and non-specific treatment effects and considered in future clinical trials of Craniosacral Therapy.
Conflict of interest statement

None declared.

Acknowledgments

The authors would like to thank Petra Klose, Ph.D., Department of Internal and Integrative Medicine, University of Duisburg-Essen, Germany, for her assistance in study organization and literature search; and Florian Gass, M.D., Department of Internal and Integrative Medicine, University of Duisburg-Essen, Germany, for his assistance in patient enrollment and assessment.

References

34. Elden H, Ostgaard HC, Gliant A, Marciniak P, Linner AC, Olsen MF. Effects of craniosacral therapy as adjunct to standard treatment for pelvic girdle pain in pregnant women: a multicenter,

