Blinding Protocols, Treatment Credibility, and Expectancy: Methodologic Issues in Clinical Trials of Osteopathic Manipulative Treatment

John C. Licciardone, DO, MBA
David P. Russo, DO, MPH

Context: In testing an experimental new drug or therapy, the gold standard in biomedical research for determining treatment efficacy is the randomized controlled trial (RCT). In pharmaceutical trials, inert placebos are an easily administered control that facilitates blinded comparisons. In clinical trials that study the effects of manual interventions, researchers must carefully consider their use of treatment control models. Choosing credible controls that will minimize bias in osteopathic manipulative treatment (OMT) clinical trials poses unique challenges to researchers because of heterogeneous OMT methods and practice.

Objective: To compare the treatment credibility of sham manipulative treatment and untreated controls to active OMT.

Methods: Subjects recruited for an OMT clinical trial for chronic low back pain completed a treatment-credibility rating scale comparing two written descriptions of the study interventions offered. The scale was administered to subjects before trial entry and at 6-month follow-up. Scale scores were used to compute credibility ratios for both intervention protocols (ie, OMT vs sham manipulative treatment). Repeated measures analysis of variance was used to assess changes in the credibility ratio over time, including the measurement of study group and time main effects, as well as study group × time interaction effects.

Results: Subjects (N=91) perceived OMT as a more credible therapeutic option than sham manipulative treatment both at trial entry and at 6-month follow-up (P<.05). Among subjects completing the study protocol (n=66), there were no changes in perceived credibility of the study interventions over time. There were no significant differences in the credibility ratio among study groups (P=.64) or over time (P=.79). In addition, there were no significant study group × time interactions (P=.59).

Conclusions: In clinical trials, OMT may be perceived by subjects as a more credible treatment alternative than many control procedures. Treatment credibility can interact with subject expectations and study design in complex ways. When analyzing the treatment effects of OMT, investigators must consider the effects of these two subjective elements when competing interventions are offered and subjects are asked to self-report data. Study design should be optimized to equalize these effects among interventions.

J Am Osteopath Assoc. 2006;106:457–463

In testing an experimental new drug or therapy, the gold standard in biomedical research for determining treatment efficacy is the randomized controlled trial (RCT). In RCTs, researchers and subjects are both blinded (or masked) to comparisons between treatment outcomes, by which one means that subjects and researchers are not able to deduce group assignments (ie, treatment vs control). In pharmaceutical trials, inert placebos are an easily administered control that facilitates double-blinded comparisons.

In clinical trials that study the effects of manual interventions, however, such as osteopathic manipulative treatment (OMT), researchers must carefully consider their use of treatment control models. When the RCT double-blind study model is applied to manual interventions, it quickly becomes apparent that researchers in those fields will not be able to adhere to a strict double-blind biomedical model in which those administering the treatment are unaware of the procedure taking place.1,2 For this reason, RCTs designed to evaluate the efficacy of OMT are often, of necessity, single-blind, with only the subjects unaware of their assignment to the study or control group.3,4 In addition, isolating the specific treatment effects (eg, the “active ingredient” in OMT) from other non-specific treatment effects using single-blind protocols is also a challenge to researchers interested in determining the efficacy of manual techniques.5

The isolation of specific treatment effects is of central importance in osteopathic medical research because osteopathic physicians have long embraced an approach to patient care that optimizes patient-physician rapport and focuses on an array of psychosocial and individual health factors.6 Therefore, studying the effects of OMT when it is removed from its standard context of healthcare delivery raises questions regarding treatment generalizability in clinical practice set-
### Differential dropout rate

In medical trials, subjects sometimes drop out of the study before the completion of the study protocols—or they are lost to follow-up. Because subjects are more likely to discontinue participation if they feel their condition is not improving, investigators need to conduct an analysis of dropout patterns between the treatment and control groups to determine if their results might be influenced by attrition bias (also known as withdrawal bias). The differential dropout rate must be addressed by researchers in their final analysis.

Unless dropout controls are retained in the data analysis, attrition bias can seriously affect study outcomes. However, it may be appropriate in some studies to exclude dropout controls from analysis if those subjects discontinued participation after the selection process but before the study began (eg, nonblinded studies where subjects are not randomized to the study group they prefer [ie, treatment vs control]).

### Investigator bias

Subject assignments to study groups should not be left to the discretion of investigators. Investigators should not, for example, be aware of the next group assignment when they are screening patients for study eligibility. In addition, matters of scheduling for certain procedures or diagnostic tests could be affected by this kind of bias in investigations studying certain somatic dysfunctions or disease states.

### Investigator-subject relationship

In the course of treating musculoskeletal complaints, physicians often provide general patient education and offer exercise and diet advice to patients. They may also evaluate the impact of concomitant medical issues or treatments, making any necessary adjustments. When such routine aspects of routine medical care are introduced to the clinical trial setting, they may improve outcomes though they are not considered by researchers to be part of the treatment under investigation. Thus, the “bedside manner” of the physician is an important part of the nonspecific treatment effects inherent in the patient-physician relationship and should be monitored in the clinical trial setting.

In addition to nonspecific treatment effects, the quality and strength of the relationship between an individual physician and his or her patient—particularly with regard to the patient’s perceptions of the physician’s status—can influence treatment outcomes (similar to attention bias or the Hawthorne effect). In clinical trials where the same physician meets with the subject at all follow-up visits, nonspecific treatment effects may be stronger compared with trials where a research coordinator (or another member of the research team) meets with subjects. However, study personnel should be careful not to appear cold or uncaring toward study participants as these behaviors can negatively impact protocol adherence.

In reviewing protocol adherence rates in phase 3 clinical trials, sites where subjects perceived study staff as more empathetic, as instilling a sense of purpose in subjects, and promoting less formal interpersonal relationships with them had high rates of adherence to treatment protocols.*

In order to diffuse nonspecific treatment effects away from any individual member of the research team, investigators should have multiple clinicians administer both active and control interventions, ensuring that contact time between clinicians and subjects in both arms of the study is identical.

---


---

*Discussing which aspect of the RCT experience each term describes (Figure).

It is widely accepted in a variety of research fields that an individual’s beliefs and expectations about treatment efficacy may have a significant effect on commonly measured outcomes, including pain, function, and quality of life. Other areas of nonpharmacologic research using RCT methods have developed valid and reliable measures of treatment credibility that may be helpful to OMT clinical investigators. Treatment credibility, one component of patient belief, refers to the degree to which subjects believe that the treatments offered in competing arms of an RCT are believable, convincing, and a rational
solution to the chief complaint. A separate component, expectancy, refers to the degree to which subjects believe that improvements will occur as a result of the treatment provided. Nonspecific treatment effects is a term that describes a broad category of idiosyncratic factors common to many medical experiences such as suggestion, persuasion, patient-physician (or investigator-subject) rapport, and the subjects’ frequently unconscious desire to please the investigator or be “acceptable” to study personnel.

Poor study design or methodologic flaws may result in subjects or investigators becoming unblinded to random assignment, thereby rendering interpretations of the data vulnerable to numerous sources of bias, including investigator bias and withdrawal bias. Although randomized blinded assignment protocols are well-accepted methodologic techniques used to conceal treatment assignments and minimize the impact of nonspecific treatment effects among study groups, treatment credibility and expectancy are less understood subject-specific psychosocial variables that should also be carefully measured in OMT clinical trials. By assessing potential study subjects for the following five variables, osteopathic medical researchers can strengthen the inference from observed outcomes that measured improvements are the result of treatment effects and not attributable to a source of bias:

- treatment history or knowledge of OMT
- treatment history or knowledge of other manual therapies
- perceptions of credibility for all competing interventions
- expectations for treatment
- perceptions of patient-physician (or investigator-subject) rapport

Measures of subjects’ perceptions of treatment credibility have been used to address issues involving placebo-controlled

<table>
<thead>
<tr>
<th>Source of bias in clinical trials</th>
<th>Measures</th>
</tr>
</thead>
<tbody>
<tr>
<td>Regression to the mean</td>
<td>Chronic disorders such as back pain tend to vary in severity over time. Participants are more likely to seek treatment, and to be enrolled in clinical trials, when their symptoms are severe. During the study period, symptom severity may decrease simply as a result of the passage of time. In other words, improvements to symptom severity are sometimes the result of the natural resolution of the dysfunction rather than an accurate measure of treatment response. Also called regression toward the mean.</td>
</tr>
<tr>
<td>Self-selection bias</td>
<td>If subjects are given a choice between study groups, many will try to increase their chances of getting assigned to the treatment group. Subjects who believe that they have been assigned to the control group often have less motivation to remain compliant with study protocols as well as lower expectations of improvement. Some types of self-selection bias are less obvious and, therefore, more difficult to control. See also temporal change in the subject pool, below. Participants who seek care from clinics or physicians identified with certain treatment types may be less motivated to comply with study protocols if they believe that they have been assigned to an alternative treatment group or the control group. Also called subject bias or volunteer bias.</td>
</tr>
<tr>
<td>Subject expectation</td>
<td>Subjects’ preconceptions about the effectiveness of the study protocols in their assigned groups can profoundly influence trial outcomes. For example, subjects may be inclined to evaluate their outcomes more positively if they believe that they have been assigned to the treatment group instead of the control group. In addition, they may be more motivated to remain compliant with study protocols (eg, home stretching program) if they believe the treatment will be effective.</td>
</tr>
<tr>
<td>Temporal change in the subject pool</td>
<td>Some clinical trials may extend over several years. During this time, the patient population at a given clinic may change, especially if there is publicity about the trial or if new treatments become available. In addition, clinics or individual investigators may become well known for providing treatment with a specific treatment modality. This particular kind of self-selection bias (see above) is known as centripetal bias or referral bias. To minimize the impact of changes in the pool of potential subjects over time, patients are usually randomized in relatively small blocks to ensure that comparable numbers of subjects are assigned to each study arm at each stage of the investigation.</td>
</tr>
<tr>
<td>Treatment credibility</td>
<td>A cumulative rating of subjects’ assessments of the credibility of the competing treatment options available within a study. Subjects may evaluate their outcomes more negatively if they believe that a competing treatment is more appropriate or logical for their condition.</td>
</tr>
</tbody>
</table>

Figure (continued). Sources of bias in clinical trials. More information on good clinical practice guidelines for clinical investigators can be found at the US Food and Drug Administration Web site (http://www.fda.gov/oc/gcp/default.htm).
comparison in other areas of nonpharmacological research.\textsuperscript{11–13} Several years ago, the authors of the present study collaborated with other researchers to conduct an RCT on the use of OMT for chronic low back pain, using two control groups in the study design: one that received sham manipulative treatment and a second that received no intervention.\textsuperscript{15} This study design allowed a direct comparison of the efficacy of OMT relative to both control groups as well as an opportunity to further explore how treatment credibility may affect trial completion and clinical outcomes.

\section*{Methods}

Methods and results for the chronic low back pain RCT have been reported in detail elsewhere.\textsuperscript{15} All trial procedures were approved by the institutional review board of the University of North Texas Health Science Center at Fort Worth and verbal and written informed consent were received from study participants. As noted, the RCT was designed with two control groups (sham manipulative treatment and no intervention) for comparison with a single OMT group. Subjects, who were masked to group assignments, were randomized evenly among the control and treatment groups in a 1:1:2 ratio. All subjects were allowed to continue their standard medical care for low back pain (eg, over-the-counter medications, physical or massage therapy, acupuncture, herbal therapies). However, study exclusion criteria prohibited subjects from receiving chiropractic spinal manipulation or additional sessions of OMT. In addition to gathering basic demographic data, investigators at regular intervals (baseline and follow-up at 1-, 3-, and 6-months postintervention) recorded data on 14 primary outcomes. Predoctoral osteopathic manipulative medicine fellows performed clinical assessments and osteopathic structural examinations for all subjects and provided individualized OMT to the study subjects in the OMT group and sham manipulative treatment to subjects randomized to that control group. The study’s final outcomes consisted of the five domains of patient-based outcomes recommended by Bombardier\textsuperscript{16} as essential for the evaluation of spinal disorders: general health status, pain, back-specific function, work disability, and back-specific patient satisfaction.

The present study focuses on subjects’ perceptions of treatment credibility with regard to the OMT and sham manipulative treatment protocols used in that January 2000 to February 2001 trial. All subjects were asked to complete a brief treatment credibility scale prior to randomized blinded assignment and then again at 6-month follow-up. Previous versions of the measure used for the present study have been used in other nonpharmacologic RCTs, including psychotherapy,\textsuperscript{11,17} acupuncture,\textsuperscript{18} and biofeedback therapy.\textsuperscript{19,20} Subjects were asked to rate their level of agreement with statements about OMT and sham manipulative treatment. The statements were designed to describe OMT and sham manipulative treatment objectively, without explicitly labeling them. For this purpose, the interventions were named \textit{Treatment 1} and \textit{Treatment 2}, respectively. Furthermore, the descriptions of these treatments were worded to minimize the likelihood that subjects would subsequently discern their treatment allocation (ie, be at risk of unblinding). The description of OMT was provided to all study subjects as follows:

\textbf{Treatment 1} uses a practitioner’s hands to find and treat painful areas of the body and back. A practitioner will look for areas of your body that hurt you and then apply manually guided forces in order to decrease pain and restore proper alignment between bones, muscles, and connective tissue that may have been altered by injury.

If you got assigned to this treatment, a practitioner would meet with you for approximately 15–30 minutes and treat areas on your body and back that hurt.

Sham manipulative treatment was described to all study subjects as follows:

\textbf{Treatment 2} involves light pressure applied to certain painful areas of the body and back. A practitioner will look for areas of your body that hurt you, lay his hands on them, and apply light pressure in order to decrease pain, improve communication between your body’s nervous and endocrine systems that may have been altered by injury, and help you relax.

If you got assigned to this treatment, a practitioner would meet with you for approximately 15–30 minutes and treat areas on your body and back that hurt.

The pertinent questionnaire item stem used to summarize treatment credibility relative to each statement was, “I am confident that this treatment can alleviate my complaint.” The available Likert weighted-scale responses for both treatment descriptions were as follows: 5, strongly agree; 4, agree; 3, undecided; 2, disagree; and 1, strongly disagree. The primary outcome measure was the credibility ratio (ie, the ratio of credibility in OMT relative to sham manipulative treatment for each subject, as computed by the relative weights for their responses to the two treatment descriptions). A 95% confidence interval was computed for each credibility ratio to assess statistical significance.

Repeated measures analysis of variance was used to assess the credibility ratio over time, including the measurement of study group and time main effects, as well as study group and time interaction effects. All data analyses were performed with the SYSTAT software package (Version 7; SYSTAT Software Inc, Richmond, Calif), using a .05 level of significance.

\section*{Results}

A total of 91 (46\%) of the 199 potential subjects who responded to recruitment procedures were eligible for study participation after screening and were randomly assigned to one of the three study groups. The three groups were comparable with regard to all baseline characteristics, including age, sex, race, pain duration and intensity, self-reported functional assessments (ie, MOS [Medical Outcomes Study] 36-Item Short-Form Health Survey,\textsuperscript{21} Roland-Morris Low Back Pain and Disability Questionnaire\textsuperscript{22}), and the number of work or school days lost in the past 4 weeks as a result of back pain.
Attrition rates during the study protocols have been reported elsewhere (1 mo, 9 [10%]; 3 mo, 20 [22%]). For the subjects who completed the study (66 [73%]), there were no significant differences in attrition over time among the three study groups ($P=.36$). The present study analyzes treatment credibility data gathered at trial entry and exit, as available from the 91 subjects who entered the RCT and the 66 subjects who completed the 6-month study.

For all subjects, the credibility ratio at baseline was 1.10 (1.03–1.16), indicating a small, but statistically significant, credibility differential favoring OMT over sham manipulative treatment. Similarly, at 6-month follow-up, the credibility ratio was 1.15 (1.06–1.23) in favor of OMT. Changes in the credibility ratio over time were not significant ($P=.36$), however.

When subjects completing the trial are compared with those who did not, the credibility ratios at baseline were 1.12 (1.03–1.21) and 1.03 (0.98–1.08), respectively ($P=.18$). The credibility ratios over time according to trial completion status and study group are presented in the Table.

The results of the repeated measures analysis of variance—which included only those subjects who provided credibility data at baseline and at 6-month follow-up—indicated that there were no significant differences in the credibility ratio among study groups ($P=.64$) or over time ($P=.79$). In addition, there were no significant study group $\times$ time interaction effects ($P=.59$).

Comment

The present study is the first to empirically explore issues related to treatment credibility, sham manipulative treatment, and methodology in OMT research. Our results indicate that subjects generally reported slightly elevated credibility ratios, suggesting a small treatment credibility differential favoring OMT over sham manipulative treatment. Although subjects appeared to perceive the description of OMT at baseline and at 6-month follow-up as a more credible therapeutic option than sham manipulative treatment, this measurable difference did not appear to affect primary outcomes. It is also noteworthy that there were no significant differences in the credibility ratio among the three study groups, over time, nor when these two factors were combined. Furthermore, drop-out rates could not be attributed to differences in perceived credibility of OMT and sham manipulative treatment, suggesting that other factors were related to attrition. Thus, within the context of this study, we can feel confident that the small, measurable differences in treatment credibility were not a significant threat to the study's internal validity.

Randomized controlled trials of OMT may combine the use of that treatment modality with standard medical care for the condition being studied. As illustrated by recent OMT trials studying management of low back pain, control groups may either include those who receive conventional medical care and sham manipulative treatment or those who receive standard care alone. There are at least two important problems with designing a study that uses controls who receive no manual intervention. First, the subjects may deduce that they have been randomly assigned to the control group (ie, unmasked), making them more likely to drop out of the trial (ie, self-selection). Second, no-intervention control subjects may receive less clinical care, either real or perceived, for their condition—violating an ethical obligation to provide care. In either case, monitoring and documenting subjects' individual beliefs about their treatment assignments as part of standard study protocol would provide investigators with significant insights into these issues and aid them in drawing valid conclusions about the cause-effect nature of their interventions.

Sham manipulative treatment is intended to overcome the problems associated with using control subjects who receive no manual intervention. In the process of providing sham treatment, however, researchers must acknowledge that

<table>
<thead>
<tr>
<th>Treatment Credibility Measured</th>
<th>Osteopathic Manipulative Treatment Subjects</th>
<th>Control Subjects</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline</td>
<td>48 (53) 1.09 (0.99–1.19)</td>
<td>23 (25) 1.14 (0.99–1.29)</td>
</tr>
<tr>
<td>Participants who discontinued</td>
<td>16 (64) 1.03 (0.95–1.10)</td>
<td>4 (16) 1.08 (0.82–1.35)</td>
</tr>
<tr>
<td>treatment*</td>
<td></td>
<td>5 (20) 1.00 (1.00–1.00)*</td>
</tr>
<tr>
<td>Participants who completed</td>
<td>32 (49) 1.12 (0.97–1.26)</td>
<td>19 (29) 1.15 (0.97–1.34)</td>
</tr>
<tr>
<td>treatment†</td>
<td></td>
<td>15 (23) 1.09 (1.00–1.17)</td>
</tr>
<tr>
<td>6-month follow-up‡</td>
<td>1.18 (1.01–1.34)</td>
<td>1.17 (1.07–1.27)</td>
</tr>
</tbody>
</table>

* Among the five control subjects who received no intervention and also discontinued “treatment,” all five reported that osteopathic manipulative treatment and sham manipulative treatment were equally credible therapeutic options.

† Some percentages do not total 100 because of rounding.

‡ For 6-month follow-up, No. (%) data for each study group are provided in the preceding row.

Abbreviations: CI, confidence interval; CR, credibility ratio.
some therapeutic benefit may occur, thereby reducing the comparative efficacy of OMT. Indeed, a 2001 review of placebo effects found that, for conditions involving continuous measures of pain, there is a small, but statistically significant, beneficial effect attributed to placebos when compared with no treatment.24 This placebo effect has been estimated to be the equivalent of one third of the effect of nonsteroidal anti-inflammatory drugs.25

The present study raises questions about OMT research and the nature of treatment credibility and the placebo effect. For some osteopathic medical investigators, the possibility of natural (i.e., regression to the mean) or placebo-related improvement in patients’ musculoskeletal conditions and pain levels dictates the use of rigorous control and masking procedures in OMT study protocols. When RCTs focus on strict mechanistically oriented questions,26–28 only sham-controlled, double-blind trials can determine if beneficial outcomes are inherently attributable to the manual techniques used. When sham-controlled trials are not feasible, careful measures should be taken to ensure that subject concealment, credibility, and expectations are equivalent between treatment arms at baseline and throughout the study period. Monitoring individual subject’s beliefs about treatment and using independent evaluators to assess outcomes establishes the cornerstone of “dual-blinding”—a promising modification of the traditional double-blind RCT methodology that may be more appropriate for evidence-based OMT research.29

If a study’s hypotheses about the effects of OMT are presumed to transcend a purely biomechanistic explanation, or if resources are limited, other pragmatic study designs may be considered. Under such conditions, OMT trials involving non–placebo-controlled interventions can provide useful clinical data—provided that subjects believe that all “treatments” offered are of comparable credibility. And yet, without directly addressing treatment mechanisms, investigators are able to offer only limited conclusions about treatment efficacy. At most, researchers can conclude that the hands-on treatment under study contained some as-yet-unknown active ingredient(s) that caused some degree of change (i.e., beyond that caused by factors common to all forms of human touch,30 physician-directed attention,31 placebo conditions, or chance). Such conclusions are likely to have only limited influence in persuading other clinical investigators, patients, policymakers, and health insurance providers about the effectiveness of OMT. A recent large, well-designed study of chiropractic manipulation, physical therapy, and standard medical care demonstrated that subjects’ initial confidence in the assigned treatment correlated directly with outcome, but that the strength of the effect depended on the quality of interaction between the subject and the operator.32

Thus, along with subjective outcome measures, it is important for OMT investigators to assess objective measures such as changes in physiologic parameters or biomechanics because these data are less susceptible to expectation bias and clearly demonstrate specific treatment effects. Objective measures of physical functioning and hormonal or neurologic change may be appropriate biomarkers for some studies of OMT and it is relatively easy to protect blinded assessment with this methodologic design. The soundness, or validity, of future studies would benefit from improved objective outcome measures and blinded assessment.

When blinded assessment is used, it is essential that investigators regularly test the integrity of the study protocols by assessing subjects and study personnel for their level of knowledge regarding treatment allocation. When operator blinding is not possible, the study design should control for investigator bias (if only partially) using independent evaluators to provide masked assessment. When subject blinding is not possible, the study design should control for expectation bias (if only partially) by regularly assessing the integrity of blinding protocols, evaluating any changes in treatment credibility and subject expectations, and ensuring adequate randomization.

The complex interaction of treatment credibility, subject expectations, and the influence of nonspecific treatment effects on clinical outcomes form the bridge that the osteopathic medical profession must cross as we seek to further evaluate the efficacy of OMT in an evidence-based paradigm. Careful consideration of these three factors and their influence on clinical outcomes are one of the hallmarks of a well-designed RCT in osteopathic medicine.

References


