THE JOURNAL OF ALTERNATIVE AND COMPLEMENTARY MEDICINE
Volume 11, Number 2, 2005, pp. 333–342

PARADIGMS

Bias Control in Trials of Bodywork: A Review of Methodological Issues

WOLF E. MEHLING, M.D., ZELDA DIBLASI, Ph.D., and FREDERICK HECHT, M.D.

ABSTRACT

Objective: To review and summarize the methodological challenges in clinical trials of bodywork or hands-on mind–body therapies such as Feldenkraïs Method, Alexander Technique, Trager Work, Eutony, Body Awareness Therapy, Breath Therapy, and Rolfing, and to discuss ways these challenges can be addressed.

Design: Review and commentary.

Methods: Search of databases PubMed and EMBASE and screening of bibliographies. Published clinical studies were included if they used individual hands-on approaches and a focus on body awareness, and were not based on technical devices.

Results: Of the 53 studies identified, 20 fulfilled inclusion criteria. No studies blinded subject to the treatment being given, but 5 used an alternative treatment and blinded participants to differential investigator expectations of efficacy. No study used a credible placebo intervention. No studies reported measures of patient expectations. Patient expectations have been measured in studies of other modalities but not of hands-on mind–body therapies. Options are presented for minimizing investigator and therapist bias and bias from differential patient expectations, and for maintaining some control for nonspecific treatment effects. Practical issues with recruitment and attrition resulting from volunteer bias are addressed.

Conclusions: Rigorous clinical trials of hands-on complementary and alternative therapy interventions are scarce, needed, and feasible. Difficulties with blinding, placebo, and recruitment can be systematically addressed by various methods that minimize the respective biases. The methods suggested here may enhance the rigor of further explanatory trials.

INTRODUCTION

Bodywork or hands-on mind–body therapies such as Feldenkraïs Method, Alexander Technique, Trager Work, Eutony, Body Awareness Therapy, Breath Therapy, and Rolfing are used with a wide spectrum of diagnoses (Anonymous, 2003; Mehling, 2001). These therapies aim to enhance body awareness using touch, movement, and a mind–body approach. Fuller integration of these therapies into medical care (Frenkel and Borkan, 2003) will depend on better evidence demonstrating that they are effective. Despite many anecdotal reports of their utility, two systematic reviews of trials of Feldenkrais and Alexander Technique revealed insufficient research data for supporting "strong conclusions" for the efficacy of these therapies (Ernst and Canter, 2003; Ives and Sosnoff, 2000).

While quasiexperimental outcomes research (Riley and Berman, 2002; Walach et al., 2002) and other study designs, such as qualitative research, are also needed (Lewith et al., 2002), randomized controlled trials (RCT) remain the gold standard for evaluating the efficacy of complementary therapies (Berman and Straus, 2004; Harlan, 2001; Levin et al.,

Osher Center for Integrative Medicine, University of California San Francisco, San Francisco, CA.

1997; Long, 2002; Miller et al., 2004; Vickers et al., 1997). Randomized, controlled trials of these treatments, however, face particular challenges similar to those encountered in trials of spinal manipulation, physical therapy (PT), and massage (Breen, 2002; Field, 2002). These challenges markedly differ from those of conventional drug therapies and have been previously summarized (Berman and Straus, 2004; Carter, 2003; Cawley, 1997; Ernst, 2003; Hart, 2003; Long, 2002; Mason et al., 2002; Redwood, 2002; Richardson, 2000; Smith, 2004). However, with rare exceptions (Field, 2002), little guidance is available for researchers in this field about how to address these methodological challenges.

The history of trial methodology is the history of controlling bias (Chalmers, 2001; Smith, 2004). Patients' expectations and preferences for complementary methods have a major influence on the benefits of these therapies and can introduce bias into outcome measures (Kalauokalani et al., 2001). If we cannot blind the patient, we lose our ability to equalize patients' expectations across study treatment arms. Randomization and allocation concealment attempt to equalize patient characteristics across study arms. However, it cannot control for bias from differences in expectations toward two different unblinded interventions.

This report identifies the best quantitative studies in the field of bodywork and reviews, illustrates, and discusses three challenging topics: blinding, choice of control group, and recruitment. Within each topic, we briefly reflect on the goals for which the current research guidelines were established, review whether and how these were applied, and discuss possible solutions.

METHODS

We searched the PubMed and EMBASE databases up to September 2004 using the following search terms: Feldenkrais, Alexander technique, Trager, Eutony, Breath Therapy, Breathing Exercises, Rolfing. Bibliographies and the Feldenkrais Research Archive (Psychology Department of the University of Utah, 2004) were screened for additional studies. Because RCTs are scarce, we included studies that met the following criteria for methodological quality and study characteristics:

1. Clinical study on patients with established medical diagnoses;
2. Quantitative data preintervention and postintervention provided;
3. Intervention included individual hands-on approach (i.e., not based on technical devices or verbal guidance only) and focused awareness (going beyond plain massage or PT);
4. Outcome measures had to be understandable within the frame of conventional medical science;
5. Publication in peer-reviewed journals.

Studies of Reiki, *qigong*, or other energy manipulations were not included in this review because these therapies are conceptually different, do not require direct touch, or are done in groups only.

RESULTS

We identified 40 studies in PubMed and EMBASE and 13 additional studies from bibliographies. Of those 53 studies, a total of 20 fulfilled our criteria. Table 1 presents an overview of the methods used in these studies.

Blinding

In controlled trials, the purpose of blinding is to reduce ascertainment or observer biases by keeping the various parties involved in the study blind to the participants' group assignment (Schulz et al., 2002). This bias is present when the assessment of outcomes is systematically influenced by knowledge of which intervention a participant receives. This knowledge is associated with expectations that might differ for the allocated treatments (Crow et al., 1999). Blinding applies to one or several of the following: study participant, therapist, outcome evaluator, investigator, and data analyst.

Blinding the patient. The purpose of blinding the patient is to control for differential patient expectations and bias in self-report outcome rating (Pocock, 1983). Patient blinding can occur to different degrees with rather different effects:

1. Masking the interventions the participants undergo will equalize bias from patient expectations.
2. Blinding the participant to which intervention the investigator expects to work better may reduce bias introduced by investigator's influence on patient expectations, but may not equalize participant expectations between unmasked interventions.

The first type of blinding probably is impossible in studies of bodywork because of their sensory nature (Deyo, 1988) and was never attempted in the reviewed studies.

The second type of blinding depends in part on the way information about the study interventions is framed during informed consent, when participants' knowledge and expectations around the interventions are largely shaped (Bergmann, et al., 1994). An example of this is consenting to behavioral therapy versus Feldenkrais for premenstrual syndrome without revealing that Feldenkrais is not expected to help this condition (Kirkby, 1994). If patients can successfully be blinded to which treatment the investigator sees as potentially effective, allocation concealment has the purpose of equalizing patient expectations without deceiving the participants. This is not possible in hands-on studies with a no-intervention control arm or a control intervention that

Table 1. Design Details of Clinical Trials[a] of Hands-on Bodywork Therapy

Therapy methods and authors	*Number of studies and conditions treated*	*Randomized*	*Controlled*	*Arms/crossover (C)*	*Sham/placebo*	*Patient blind*	*No-intervention arm*	*Control groups*
Feldenkrais	Eight studies							
Kirkby (1994)	Premenstrual syndrome	−	+	3	−	−	+	1) CBT, 2) NT (waitlist)
Laumer et al. (1997)	Eating disorder	−	+	2	−	−	−	same multimodal program, but no FM
Joynson et al. (1999)	Multiple sclerosis	−	+	2C	+	−	−	sham bodywork, same provider in both groups
Lundblad et al. (1999)	Neck-shoulder problems	+	+	3	−	−	+	1) PT, 2) NT (waitlist)
Smith et al. (2001)	Chronic LBP	+	+	2	+	−	−	audiotape narrataive ('story')
Stephens et al. (2001)	Multiple sclerosis	+	+	2	−	−	−	educational classes
Loewe et al. (2002)	Acute MI	−	+	3	−	−	+	1) progressive muscle relaxation, 2) NT
Malmgren-Olsson et al.[b] (2002)	Chron. nonspecific musculoskeletal disorders	−	+	3	−	−	−	1) BAT, 2) PT
Alexander technique[c]	Three studies							
Stallibrass (1997)	Parkinson's disease	−	−	1	−	−	−	none
Stallibrass et al. (2002)	Parkinson's disease	+	+	3	−	−	+	1) massage, 2) NT
Elkayam et al. (1996) [45]		−	−	1	−	−	−	none (limited historic control)
Breath therapy[d]	Six studies							
Loew et al. (1996a)	Asthma	+	+	3C	+	+	−	1) placebo relaxation, 2) MDI: patients are their own controls on 2 following days.
Loew et al. (1996b)	Asthma	−	+	2C	−	−	−	MDI: patients are their own control on the following day
van Dixhoorn et al. (1999)	Rehabilitation after MI	+	+	2	−	−	−	same exercise training, but no BT
Loew et al. (2000)	Asthma	+	+	2	+	+	−	placebo relaxation
Loew et al. (2001)	Asthma	+	+	3C	+	−	−	1) placebo relaxation, 2) MDI: patients are their own controls on 2 following days.
Manocha et al. (2002)	Asthma	+	+	2	−	−	−	relaxation technique including auto-suggestions, visualization, progressive relaxation, and group discussions
Body Awareness Therapy	Four studies							
Engel et al. (2000)	Chron. toxic encephalopathy	−	+	2	−	−	−	none
Grahn et al. (2000)	Chron. musculoskeletal pain	−	+	2	−	−	−	primary care, standard care as outpatients
Haugli et al. (2001)	Chron. musculoskeletal pain	+	+	3	−	−	−	standard care only
Malmgren-Olsson and Branholm, (2002)[b]	Chron. nonspecific musculoskeletal disorders	−	+	3	−	−	−	1) FM, 2) PT
Eutony	None							
Trager	None							
Rolfing[e]	None							

[a]If a single trial generated multiple reports, only one report is quoted.

[b]This study fits into two fields as one method is control for the other.

[c]We know of at least one unpublished 3-arm randomized controlled trial (RCT) on low-back pain patients (Little, Great Britain).

[d]The term "breath therapy" or "breath exercises" is relatively broad and includes a wide variety of Western and Eastern approaches from diverse conceptual backgrounds. We found numerous studies assessing the health implications of influencing breathing patterns by using technical devices (biofeedback, video instructions, music, mouthpiece) thus providing well standardized interventions and valuable data on the clinical importance of various breathing patterns and their manipulation by physiologic elements of breath therapy. However, for the purpose of this review, these did not add to the discussion of methodological issues in hands-on bodywork research and did not meet our inclusion criteria.

[e]We found 2 RCTs on healthy volunteers only.

FM, Feldenkrais; BAT, Body Awareness Therapy; BT, Breath Therapy; PT, physical therapy; CBT, cognitive–behavioral therapy; NT, no-treatment; MDI, metered-dose inhaler; LBP, low-back pain; MI, myocardial infarction.

was explained as being not effective during informed consent.

Four studies reported single blinding for study participants. In these studies, at least one control group received an intervention that was presented as being an alternative effective method providing comparable benefits, although it was clearly seen as placebo by the investigator: placebo–relaxation versus Breath Therapy for asthma (Loew et al., 1996a, 2001), true relaxation versus Breath Therapy for asthma (Manocha et al., 2002), and Feldenkrais versus cognitive–behavioral therapy for premenstrual syndrome (Kirkby et al., 1994). Participants were kept uninformed about the other interventions and the investigators' hypotheses were not disclosed.

In a fifth study (Loew et al., 1996b) reporting patient blinding, participants received two different relaxation instructions, Breath Therapy or a placebo relaxation instruction, crossing over on consecutive days. Again, participants expected two active interventions and were deceived about one intervention that was used as sham. However, in this study each participant experienced both interventions and could reach conclusions rendering blinding ineffective.

Although the sensory characteristics of hands-on interventions likely invalidate blinding in a crossover design (Deyo, 1988), this was used in four studies (Loew et al., 1996a, 1996b, 2001; Johnson et al., 1999). No data were provided as to whether participants or therapists had similar expectations of the two treatments.

Blinding the therapist. None of the reviewed studies blinded the therapists. The purpose of blinding the therapist is to control for conscious or unconscious influences on the participant, which could modify the intervention's effect (Schulz et al., 2002). It is obvious that in interventions that are hands-on and provide guidance for awareness or movement exercises the therapists cannot be blind to the treatment they are delivering. Therapist blinding has been attempted in massage research by providing massage with versus without pressure. Presumably, the therapists were not informed that pressure-free massage was a sham (Diego et al., 2004; Field, 2002). Apart from this being a questionable assumption (Cassileth and Vickers, 2004), therapists applying pressure-free massage are almost certain to believe they are delivering suboptimal treatment. Even this limited degree of blinding is not possible in studies involving a guided body-awareness focus.

Blinding of outcome assessment. The purpose of blinding study personnel who are evaluating outcome measures is to avoid conscious or unconscious biasing of the outcomes' assessment. This is particularly important if the outcome measures are vulnerable to unconscious manipulation by unblinded assessors (Schulz et al., 2002). An independent person whose only role is to perform outcome measures can be kept blinded to the group assignment. Thirteen (13) of the reviewed studies used outcome measures, that might be vulnerable to bias when assessed by unblinded observers, such as assessment of physical function, range of movement, flexibility (Elkayam et al., 1996; Grahn et al., 1998; Johnson et al., 1999; Lundblad et al., 1999; Stephens, 2001), or pulmonary function (Loew et al., 1996a, 1996b, 2001; Manocha et al., 2002). Only three of these studies reported an independent outcome assessment (Johnson et al., 1999; Stallibrass et al., 2002; van Dixhoorn and Duivenvoorden, 1999). Blinding research personnel entering questionnaires' data was reported in one study (Stallibrass et al., 2002).

Blinding: conclusions and recommendations. Unlike drug trials with identical placebo pills, in studies involving touch and verbal guidance it is difficult to blind either therapist or patient. Our review did not yield any convincing methods of blinding patients or therapists. Ernst and Canter (2003), researchers who strongly advocate for blinding and RCTs, concede in their systematic review of studies of Alexander Technique that "patient blinding seems impossible." However, double blinding is not a *sine qua non* of RCTs (Schulz et al., 2002), and there are several approaches that future research might use to diminish bias that may result from lack of blinding:

1. Objective outcome measures: Along with subjective outcome measures, it may be important to assess objective outcome measures such as changes in laboratory tests. Many researchers regard these as less susceptible to bias from expectations. Objective measures of physical functioning may be appropriate endpoints for studies of bodywork. It is relatively easy to blind an observer making these assessments. The internal validity of future studies may benefit from including objective outcome measures and blinding study staff that makes the assessments.
2. Assessing the success of blinding: If blinding is attempted, patients can be queried about whether or not they believe they received the study intervention (Schulz et al., 2002). The results of the patient query can be used as a control variable in data analysis, but this was not attempted in any of the reviewed studies. Observers making outcome assessments can be queried as well.
3. Assessing patient expectations: Patient expectations can be assessed for therapies that cannot be blinded. In a study comparing acupuncture to massage or an education booklet in patients with low-back pain, questions on expectations for each treatment were included in the outcome instruments (Kalauokalani et al., 2001). Overall results showed a moderate advantage of massage over acupuncture, which was similar to placebo. However, outcome in any therapy was superior to the other according to expectations. The analysis of study outcome data can incorporate a measure of expectations as a covariate in a regression analysis or compare data stratified according

to expectations. This allows an unblinded study to maintain some ability to control for patient expectations.

4. Choice of control group: How much patient expectations differ between groups depends in part on the control group. In an open study, a no-intervention control (i.e., waitlist) creates negligible patient expectations. Placebo or active controls create expectations that can be measured and compared with those for the study intervention, thus improving control over bias. This issue is discussed in more detail in the section on control group.
5. Preconsent randomization: Participants may be disappointed when assigned to what might be perceived as an inferior treatment. By randomizing participants before consent and describing only the treatment (or observation procedures) to which they are randomized, patients can be blinded to the study's purpose and disappointment or negative expectations can be reduced. However, treatment expectations may still differ between groups with this approach. Preconsent randomization was performed in studies by Cherkin et al. (1996) and Williams et al. (2003), where participants in the control group (placebo or usual care) were not informed about the main study intervention to avoid disappointment. An independent ethical review board needs to decide whether the benefits of conducting a more rigorous study and of avoiding disappointment for the participants outweigh the ethical concerns of a design that does not fully inform participants.

Control group

Control group options are placebo, active control, or no intervention. The purpose of using a placebo control group is to control for numerous nonspecific or placebo effects of the intervention on outcome measures (Ernst, 2001; Gotzsche, 1994). Difficulties in identifying a suitable placebo intervention as the control in a RCT are related to the blinding issue, as an unmasked placebo will not control for patient expectations (Crow et al., 1999). Ideally, placebo controls mimic the study intervention as closely as possible (i.e., function as a "sham" intervention) in order to blind at least the patient. Unfortunately, the provider who cannot be blinded might bring a very different degree of passion and intention to heal to a sham intervention as compared to the real intervention, thus producing nonspecific or placebo effects of rather different degrees between groups (Ernst, 2001; Freund et al., 1972; Gracely et al., 1985; Gryll and Katahn, 1978; Smith, 1989). Hands-on interventions are particularly difficult to convert to a sham intervention because nonverbal human touch communicates motivation, empathy, and mental presence (Latey, 2001). They cannot effectively control for bias from nonspecific effects from differential provider engagement, as the therapist's intention and passion to heal (here viewed as nonspecific components, although the discussion is not settled whether these might be key components) would not be included in the nonspecific effect of the control intervention. Ernst and Canter (2003) summarize this issue in their review cited earlier by stating, "There is no credible placebo."

Only two of the studies we reviewed used placebo in the control groups: In these studies (by the same author), a placebo relaxation instruction was developed for patients with asthma (Loew et al., 1996a, 2001): A 10-minute standardized but substantially shortened version of the investigated Breath Therapy was compared to an equally timed placebo method that lacked the focus on body awareness. The same therapist performed both the placebo and the study intervention, thus controlling for therapist personality but not for differential provider engagement or intention to heal.

In summary, our review of studies including touch and verbal guidance did not discover any credible placebo intervention to control for nonspecific treatment factors.

In two studies, a waitlist–control study design was used in which the waitlisted controls received no treatment (Kirkby, 1994; Lundblad et al., 1999): Feldenkrais for neck–shoulder problems or for premenstrual syndrome. This design clearly does not blind participants, does not control for expectations or biased self-report outcome ratings, does not control for nonspecific treatment effects, and may even introduce a disillusionment effect (Hart, 2003) or resentful demoralization (Torgenson and Sibbald, 1998) in the patients who have "drawn the short straw," thus weakening internal and external validity. However, a delayed intervention might be less demoralizing than no intervention, although this is not known.

Nonspecific treatment effects or context effects are those that are not specific to the treatment given, but influenced by patient expectations, patient–provider interaction, treatment appearance, and the healing environment (Di Blasi and Kleijnen, 2003; Gotzsche, 1994). From the perspective of patient outcomes, both specific and nonspecific effects could be viewed as important, and nonspecific effects could be viewed less as a trial nuisance than as potentially meaningful mediators and moderators of therapeutic outcomes in clinical trials (Di Blasi and Reilly, 2004). When trying to answer questions around the efficacy of a treatment, however, the distinction between specific and nonspecific effects becomes important. Out review found several ways to maintain systematic control over some nonspecific factors:

1. Control for time, setting, and practitioner: Time spent with therapist and different settings in complementary and alternative medicine (CAM) therapy are individual components of a nonspecific treatment effect that can be equalized between study arms. Seven reviewed studies made great efforts to equalize duration and provider characteristics (Johnson et al., 1999; Kirkby, 1994; Loew et al., 2000; Lowe et al., 2002; Manocha et al., 2002; Smith et al., 2001; Stallibrass et al., 2002) and three studies of Feldenkrais and Alexander Techniques reported on

equalized room settings (Lowe et al., 2002; Stallibrass et al., 2002; Stephens, 2001).

2. Control for attention: Empathic attention can be controlled for by using an active and equally credible control intervention that is similar but lacks the specific element that characterizes the study intervention. In this study design, the control group receives similar personal attention, empathy, intention to heal, listening, and some explanations, but not the therapeutic ingredient(s) seen as specific for the studied modality (i.e., guided awareness of physical sensations and breath movement, evoking a state of alert and relaxed presence). This approach would test the effect of the specific components and control for some nonspecific effects. It was used in 9 of the 20 reviewed trials (Kirkby, 1994; Loew et al., 1996a, 2000, 2001; Lowe et al., 2002; Lundblad et al., 1999; Manocha et al., 2002; Smith et al., 2001) of which six were randomized: Breath Therapy was compared to relaxation technique (Manocha et al., 2002), cognitive-behavioral therapy to Feldenkrais (Kirkby, 1994), a shortened version of Breath Therapy to placebo-relaxation (Loew et al., 2000), and partially audiotaped Feldenkrais to audiotaped narrative (Smith et al., 2001); in two studies participants served as their own crossover controls, comparing a shortened version of Breath Therapy to placebo-relaxation (Loew et al., 1996a, 2000). The more nonspecific elements of a therapy are controlled for, the more explanatory a study becomes. There appear to be different degrees of rigor for explanatory trials. The most rigorous has been called fastidious trial, which does not permit individualized therapy (Schwartz and Lellouch, 1967). Although it has been repeatedly put into question whether this level of rigor is desirable and appropriate for CAM and other therapies (Hyland, 2003; Lewith et al., 2002; Walach et al., 2002), explanatory trials appear to be feasible when the study protocol allows for some individualized variations around a core of standardized treatment guidelines. This design would be a compromise between explanatory trials assessing treatment efficacy above and beyond nonspecific treatment effects in the control group and pragmatic trials assessing effectiveness of a real-life treatment. Pragmatic trials use flexible, individualized protocols, in which the nonspecific components of extra attention and empathy are allowed to add to and freely interact with the specific effects of the test intervention thus maximizing outcome. Rather than being mutually exclusive and contradicting each other, both approaches complement each other (Lewith et al., 2002), and may be the poles of a continuum with both explanatory and pragmatic value to varying degrees.
3. A third no-treatment arm: If a credible placebo intervention as control can still be conjured up and patients can be blinded to group allocation and the investigator's intentions, that placebo's nonspecific effect can be estimated by comparison with a third no-treatment arm. A three-arm design was used in several studies (Kirkby, 1994; Lowe et al., 2002; Lundblad et al., 1999; Stallibrass et al., 2002). Interestingly, one of these studies used Feldenkrais as the active, supposedly ineffective placebo for a condition (premenstrual syndrome), for which it was not expected to help, and included a third no-treatment arm thus providing us with valid data of the nonspecific effect of Feldenkrais in that condition (Kirkby, 1994). The others used Feldenkrais versus PT versus no-treatment for shoulder-neck problems (Lundblad et al., 1999), Alexander Technique versus massage versus no-treatment for Parkinson's disease (Stallibrass et al., 2002), or Feldenkrais versus muscle relaxation versus no-treatment for acute myocardial infarction (Lowe et al., 2002). These studies, however, either did not test for statistical differences across groups (Stallibrass et al., 2002) or tested across all three groups together (Lowe et al., 2002; Lundblad et al., 1999), thus missing the opportunity to test discriminatively for a specific efficacy beyond nonspecific effects or the nonspecific effect compared to no-treatment. Nevertheless, these studies provide data on the amount of specific and nonspecific effects, even if they were not analyzed or presented that way. Using a third no-treatment arm can help to differentiate between specific and measurable nonspecific effects of a therapy (Ernst, 2001). Furthermore, it can help control for other elements that frequently contribute to nonspecific or placebo effects: regression to the mean, spontaneous recovery, use of concurrent therapies, and the Hawthorne effect (Ernst, 2001).
4. A couched study intervention: Using a large multimodal control intervention package identical to the treatment package except for the study intervention may make it possible to couch the study intervention in a bundle of cointerventions. This was done in a nonrandomized controlled study of inpatients with eating disorder treated over 5 weeks primarily with intensive individual and group psychotherapy and optional additional movement therapy (including Feldenkrais and dance) (Laumer et al., 1997). In another study (Elkayam et al., 1996), Alexander Technique was one of seven modalities in an uncontrolled comprehensive outpatient treatment plan for chronic low-back pain (including back schooling, intensive psychologic interventions, muscle relaxation training, acupuncture, chiropractic, Alexander Technique, diet counseling, and a pain specialist). Theoretically, this treatment could have been compared to an almost identical control group without Alexander technique, thus possibly blinding the patients to some (measurable) degree and reduce bias from differential patient expectations. However, such design has not yet been employed in a randomized trial and would not provide definitive data as, theoretically, treatment interactions might modify the effect of the treatment of interest.

Volunteer bias in recruitment and attrition

Expectations and preferences for CAM therapies often are highly emotional and, as a result, "patients may not want to take a chance with randomization . . . in an environment, where patients' enthusiasm is often in favor of CAM and against receiving a control treatment" (Ernst, 2003; Kalauokalani et al., 2001) or no treatment. This volunteer bias generates difficulties with obtaining informed consent and recruitment (Ernst, 2003) and with preventing differential postrandomization attrition, thus weakening subsequent interpretation of the results.

There may be a wide variance in patient expectations following different recruitment modes. Self-referral (i.e., in response to radio or newspaper advertisements or from CAM practitioners) likely introduces a higher degree of volunteer bias compared to sequential referral from physicians. Consequently, recruitment sources should be reported. Only one of the reviewed studies did not report any recruitment details (Engel and Andersen, 2000).

Only one study with dropout rates that differed between groups discussed causes and conclusions (Manocha et al., 2002), four reported clearly different rates without further analysis or comment (Kirkby, 1994; Loew et al., 2000; Lundblad et al., 1999; Malmgren-Olsson and Branholm, 2002). No study reported attempts in assessing or reducing differences in volunteer bias. We recommend reducing this bias and the subsequent difficulties with consent and attrition by the following measures:

1. Offering an attractive, high-quality control intervention.
2. Recruitment from university-based clinics or other settings, where patients are often glad to contribute to research and not already committed to using a particular CAM therapy.
3. Recruitment through physician referral: This recruitment mode was used in 10 of the reviewed studies, whereas the remaining studies used multiple sources (e.g., flyers, newspaper advertisement). Data from 10 studies reporting dropout rates and recruitment mode suggest that recruitment via physician referral may decrease differential drop-out rates. Recruitment through direct physician referral, however, can be cumbersome particularly from university primary care clinics, as one of the authors learned during a recently completed study of Breath Therapy for chronic low-back pain (Mehling, 2004). He obtained independent review board (IRB) approval to use a different mode of recruitment that accelerated enrollment: The university's IRB waived the prior patient authorization for accessing the electronic medical center database to obtain lists of potentially eligible patients sorted by primary care providers and allowed mass mailings of provider-signed information letters to these patients. It took 3 months to recruit a total of 8 participants by individual direct primary care provider referral and it took 2 weeks after mailing to recruit another 25 from the same clinics. Similarly, the electronic database of an integrated health care system could be used.

Even these three measures combined might not be able to overcome volunteer bias and differential dropout rates: In the study of Breath Therapy by one of the authors (Mehling, 2004), an attractive, free, high-quality, individualized PT from highly motivated and qualified providers was offered as control intervention. Nevertheless, several patients were disappointed with their allocation to the control intervention, confirming the disillusionment effect cited above. This was reflected in higher drop-out rates in the control arm (7 of 18 in control versus 3 of 18 in Breath Therapy group). Therefore, another recently tried measure might be considered:

4. Partial randomization or patient-preference trials, whereby patients are given a choice according to their treatment preferences, and only patients without strong preferences get randomized (i.e., standard of care versus CAM method versus randomization to one of the two) (Carter, 2003; Zelen, 1979). In a similar design, patients get randomized to usual care or a choice between usual care and a CAM method (Eisenberg, 2002). Both types of studies in patients with low-back pain are underway (Eisenberg, 2002; North American Spine Society Board of Directors, 2003). To our knowledge, this method has not been used yet in this field.

CONCLUSION

Rigorous clinical trials of hands-on complementary and alternative therapy interventions are scarce and clearly needed. They face a series of particular challenges that can be strategically and systematically addressed in order to minimize bias. Among methodological difficulties, issues with blinding, choice of control intervention, and volunteer bias are foremost.

When therapist blinding is not possible, control for therapist or investigator bias can be partially maintained by blinding an independent outcome assessor. When patient blinding is not possible, control for patient expectations can be partially maintained by assessing the success of attempted blinding, assessing patient expectations, using a large multimodal intervention package in the main arm with an identical treatment package minus the study intervention for the control, and preconsent randomization. When a placebo control is not feasible, control for nonspecific treatment effects can be partially maintained by carefully and systematically controlling for their individual elements, such as time spent with therapist, therapist engagement and attention, settings, and room environment, use of a similar active control in-

tervention lacking the specific ingredient that characterizes the study intervention, and by collecting data on nonspecific effects by a third no-intervention study arm. While an absolute bias control is not realistic, these methods may help to minimize various key biases encountered in trials of bodywork.

Variations in expectations of treatment outcomes emerge as the central theme connecting the challenges in the three discussed areas of blinding, choice of controls, and volunteer bias. Although the assessment of expectations appears to be an important response to the reviewed methodological challenges, it has been subject of only limited research (Di Blasi et al., 2001). That such a measure can be included was demonstrated in a study by Kalauokalani et al. (2001). More research on the complex theme of expectations and their assessment is needed.

ACKOWLEDGMENTS

We would like to thank Susan Folkman, Ph.D., Dan Cherkin, Ph.D., and Karen Sherman, Ph.D., M.P.H., for their constructive criticism during the preparation of this manuscript.

REFERENCES

Anonymous. Back pain and related Conditions. Levels of scientific evidence for specific therapies. Natural Standard Online document at: www.naturalstandard.com 2003. Accessed February 15, 2005.

Bergmann JF, Chassany O, Gandiol J, Deblois P, Kanis JA, Segrestaa JM, Caulin C, Dahan R. A randomised clinical trial of the effect of informed consent on the analgesic activity of placebo and naproxen in cancer pain. Clin Trials Metaanal 1994;29: 41–47.

Berman JD, Straus SE. Implementing a research agenda for complementary and alternative medicine. Annu Rev Med 2004;55: 239–254.

Breen A. Manual therapies. In: Lewith G, Jonas W, Walach H, eds. Clinical Research in Complementary Therapies. London: Harcourt, 2002:247–261.

Carter B. Methodological issues and complementary therapies: Researching intangibles? Complement Ther Nurs Midwifery 2003; 9:133–139.

Cassileth BR, Vickers AJ. Massage therapy for symptom control: Outcome study at a major cancer center. J Pain Symptom Manage 2004;28:244–249.

Cawley N. A critique of the methodology of research studies evaluating massage. Eur J Cancer Care (Engl) 1997;6;23–31.

Chalmers I. Comparing like with like: Some historical milestones in the evolution of methods to create unbiased comparison groups in therapeutic experiments. Int J Epidemiol 2001;30: 1156–1164.

Cherkin DC, Deyo RA, Street JH, Hunt M, Barlow W. Pitfalls of patient education. Limited success of a program for back pain in primary care. Spine 196;21:345–355.

Crow R, Gage H, Hampson S, Hart J, Kimber A, Thomas H. The role of expectancies in the placebo effect and their use in the delivery of health care: A systematic review. Health Technol Assess 1999;3:1–96.

Deyo, R. Measuring the functional status of patients with low back pain. Arch Phys Med Rehabil 1988;69:1044–1053.

Di Blasi Z, Kleijnen J. Context effects. Powerful therapies or methodological bias? Eval Health Prof 2003;26:166–179.

Di Blasi Z, Harkness E, Ernst E, Georgiou A, Kleijnen J. Influence of context effects on health outcomes: A systematic review. Lancet 2001;357:757–762.

Di Blasi Z, Reilly D. Placebos in practice and research: Disentangling medical paradoxes. BMJ 2004;329:927–928.

Diego MA, Field T, Sanders C, Hernandez-Reif M. Massage therapy of moderate and light pressure and vibrator effects on EEG and heart rate. Int J Neurosci 2004;114:31–45.

Eisenberg DM. Patient and provider expectations as predictors of outcome in a trial of complementary therapies for acute low back pain. Presentation at the International Scientific Conference on Complementary, Alternative, and Integrative Medicine Research, Boston: April 2, 2002.

Elkayam O, Ben Itzhak S, Avrahami E, Meidan Y, Doron N, Eldar I, Keidar I, Liram N, Yaron M. Multidisciplinary approach to chronic back pain: Prognostic elements of the outcome. Clin Exp Rheumatol 1996;14:281–288.

Engel L, Andersen LB. Effects of body-mind training and relaxation stretching on persons with chronic toxic encephalopathy. Patient Educ Couns 2000;39:155–161.

Ernst E. Obstacles to research in complementary and alternative medicine. Med J Aust 2003;179:279–280.

Ernst E. Towards a scientific understanding of placebo effects. In: Peters D, ed. Understanding the Placebo Effect in Complementary Medicine. London: Harcourt, 2001:17–29.

Ernst E, Canter PH. The Alexander Technique: A systematic review of controlled clinical trials. Forsch Komplementarmed Klass Naturheilkd 2003;10:325–329.

Field T. Massage therapy research methods. In: Lewith G, Jonas W, Walach H, eds. Clinical Research in Complementary Therapies. London: Harcourt, 2002:263–287.

Frenkel MA, Borkan JM. An approach for integrating complementary-alternative medicine into primary care. Fam Pract 2003;20:324–332.

Freund J, Krupp G, Goodenough D, Preston LW. The doctor–patient relationship and drug effect. Clin Pharmacol Ther 1972; 13:172–180.

Gotzsche PC. Is there logic in the placebo? Lancet 1994;344:925–926.

Gracely RH, Dubner R, Deeter WR, Wolskee PJ. Clinicians' expectations influence placebo analgesia. Lancet 1985;1:43.

Grahn B, Ekdahl C, Borgquist L. Effects of a multidisciplinary rehabilitation programme on health-related quality of life in patients with prolonged musculoskeletal disorders: A 6-month follow-up of a prospective controlled study. Disabil Rehabil 1998; 20:285–297.

Grahn B, Ekdahl C, Borgquist L. Motivation as a predictor of changes in quality of life and working ability in multidisciplinary rehabilitation. A two-year follow-up of a prospective controlled study in patients with prolonged musculoskeletal disorders. Disabil Rehabil 2000;22:639–654.

Gryll SL, Katahn M. Situational factors contributing to the placebos effect. Psychopharmacology (Berl) 1978;57:253–261.

Harlan WR Jr. New opportunities and proven approaches in complementary and alternative medicine research at the National Institutes of Health. J Altern Complement Med 2001;7(Suppl 1):S53–59.

Hart A. What is the research question? A case study in the early stages of design of a randomised controlled trial for a complementary therapy. Complement Ther Med 2003;11:42–45.

Haugli L, Steen E, Laerum E, Nygard R, Finset A. Learning to have less pain—Is it possible? A one-year follow-up study of the effects of a personal construct group learning programme on patients with chronic musculoskeletal pain. Patient Educ Couns 2001;45:111–118.

Hyland ME. Methodology for the scientific evaluation of complementary and alternative medicine. Complement Ther Med 2003; 11:146–153.

Ives JC, Sosnoff J. Beyond the mind–body exercise hype. Phys Sportsmed 2000;28:67–81.

Johnson SK, Frederick J, Kaufman M, Mountjoy B. A controlled investigation of bodywork in multiple sclerosis. J Altern Complement Med 1999;5:237–243.

Kalauokalani D, Cherkin DC, Sherman KJ, Koepsell TD, Deyo RA. Lessons from a trial of acupuncture and massage for low back pain: Patient expectations and treatment effects. Spine 2001;26:1418–1424.

Kaptchuk TJ. Intentional ignorance: A history of blind assessment and placebo controls in medicine. Bull Hist Med 1998;72:389–433.

Kirkby RJ. Changes in premenstrual symptoms and irrational thinking following cognitive–behavioral coping skills training. J Consult Clin Psychol 1994;62:1026–1032.

Latey P. Placebo responses in bodywork. In: Peters D, ed. Understanding the Placebo Effect in Complementary Medicine. London: Harcourt, 2001:147–163.

Laumer U, Bauer M, Fichter M, Milz H. Therapeutic effects of the Feldenkrais method "awareness through movement" in patients with eating disorders [in German]. Psychother Psychosom Med Psychol 1997;47:170–180.

Levin JS, Glass TA, Kushi LH, Schuck JR, Steele L, Jonas WB. Quantitative methods in research on complementary and alternative medicine. A methodological manifesto. NIH Office of Alternative Medicine. Med Care 1997;35:1079–1094.

Lewith G, Jonas W, Walach H. Clinical Research on Complementary Therapies. London: Harcourt, 2002.

Lewith G, Walach H, Jonas W. Balanced research strategies for complementary and alternative medicine. In: Lewith G, Jonas W, Walach H, eds. Clinical Research in Complementary Therapies. London: Harcourt, 2002:3–27.

Loew TH, Tritt K, Siegfried W, Bohmann H, Martus P, Hahn EG. Efficacy of 'functional relaxation' in comparison to terbutaline and a 'placebo relaxation' method in patients with acute asthma. A randomized, prospective, placebo-controlled, crossover experimental investigation. Psychother Psychosom 2001;70:151–157.

Loew TH, Martus P, Rosner F, Zimmermann T. The efficiency of "functional relaxation" in comparison with salbutamol and a placebo-relaxation-technique in asthmatics: A prospective randomized study in children and adolescents [in German]. Monatsschr Kinderheilkd 1996a;144:1357–1363.

Loew TH, Siegfried W, Martus P, Tritt K, Hahn EG. 'Functional relaxation' reduces acute airway obstruction in asthmatics as effectively as inhaled terbutaline. Psychother Psychosom 1996b; 65:124–128.

Loew TH, Sohn R, Martus P, Tritt K, Rechlin T. Functional relaxation as a somatopsychotherapeutic intervention: A prospective controlled study. Altern Ther Health Med 2000;6:70–75.

Long AF. Outcome measurement in complementary and alternative medicine: Unpicking the effects. J Altern Complement Med 2002;8:777–786.

Lowe B, Breining K, Wilke S, Wellmann R, Zipfel S, Eich W. Quantitative and qualitative effects of Feldenkrais, progressive muscle relaxation, and standard medical treatment in patients after acute myocardial infarction. Psychother Res 2002;12:179–191.

Lundblad I, Elert J, Gerdle B. Randomized controlled trial of physiotherapy and Feldenkrais Interventions in female workers with neck-shoulder complaints. J Occup Rehabil 1999;9:179–194.

Malmgren-Olsson EB, Branholm IB. A comparison between three physiotherapy approaches with regard to health-related factors in patients with non-specific musculoskeletal disorders. Disabil Rehabil 2002;24:308–317.

Manocha R, Marks GB, Kenchington P, Peters D, Salome CM. Sahaja yoga in the management of moderate to severe asthma: A randomised controlled trial. Thorax 2002;57:110–115.

Mason S, Tovey P, Long AF. Evaluating complementary medicine: Methodological challenges of randomized controlled trials. BMJ 2002;325:832–834.

Mehling W. Breath therapy for chronic low back pain. A randomized controlled trial. Poster presentation at 2nd Bay Area Research Symposium, San Francisco, 2004. October 2004.

Mehling WE. The experience of breath as a therapeutic intervention. Forsch complementärmed Klass Naturheilkd 2001;8:359–367.

Miller FG, Emanuel EJ, Rosenstein DL, Straus SE. Ethical issues concerning research in complementary and alternative medicine. JAMA 2004;291:599–604.

North American Spine Society Board of Directors. Spine Patient Outcome Research Trial (SPORT): Multi-center randomized clinical trial of surgical and non-surgical approaches to the treatment of low back pain. Spine J 2003;3:417–419.

Pocock S. Clinical Trials: A Practical Approach. New York: Wiley, 1983.

Psychology Department of the University of Utah. Feldenkrais® Research Archive, 2004. Online document at: www.psych.utah.edu/feldenkrais/ Accessed December 22, 2004.

Redwood D. Methodological challengs in the evaluation of complementary and alternative medicine: Issues raised by Sherman et al. and Hawk et al. J Altern Complement Med 2002;8:5–6.

Richardson J. The use of randomized controlled trials in complementary therapies: Exploring the issues. J Adv Nurs 2000;32:398–406.

Riley D, Berman B. Complementary and alternative medicine in outcomes research. Altern Ther Health Med 2002;8:36–37.

Schulz KF, Chalmers I, Altman DG. The landscape and lexicon of blinding in randomized trials. Ann Intern Med 2002;136:254–259.

Schwartz D, Lellouch J. Explanatory and pragmatic attitudes in clinical trials. J Chron Dis 1967;20:637–648.

Smith AL, Kolt GS, McConville JC. The effect of the Feldenkrais method on pain and anxiety in people experiencing chronic low back pain. NZ J Physiother 2001;29:6–14.

Smith EB. Effect of investigator bias on clinical trials. Arch Dermatol 1989;125:216–218.

Smith WB., Research methodology: Implications for CAM pain research. Clin J Pain 2004;20:3–7.

Stallibrass C. An evaluation of the Alexander Technique for the management of disability in Parkinson's disease—A preliminary study. Clin Rehabil 1997;11:8–12.

Stallibrass C, Sissons P, Chalmers C. Randomized controlled trial of the Alexander technique for idiopathic Parkinson's disease. Clin Rehabil 2002;16:695–708.

Stephens J. Use of Awareness Through Movement improves balance and balance confidence in people with multiple sclerosis: A randomized controlled study. Neurol Rep 2001;25:39–49.

Torgerson DJ, Sibbald B. Understanding controlled trials. What is a patient preference trial? BMJ 1998;316:360.

van Dixhoorn JJ, Duivenvoorden HJ. Effect of relaxation therapy on cardiac events after myocardial infarction: A 5-year follow-up study. J Cardiopulm Rehabil 1999;19:178–185.

Vickers A, Cassileth B, Ernst E, Fisher P, Goldman P, Jonas W, Kang SK, Lewith G, Schulz K, Silagy C. How should we research unconventional therapies? A panel report from the Conference on Complementary and Alternative Medicine Research Methodology, National Institutes of Health. Int J Technol Assess Health Care 1997;13:111–121.

Walach H, Jonas W, Lewith G. The role of outcome research in evaluationg complementary and alternative medicine. Altern Ther Health Med 2002;8:88–95.

Williams NH, Wilkinson C, Russell I, Edwards RT, Hibbs R, Linck P, Muntz R. Randomized osteopathic manipulation study (ROMANS): Pragmatic trial for spinal pain in primary care. Fam Pract 2003;20:662–669.

Zelen M. A new design for randomized clinical trials. N Engl J Med 1979;300:1242–1245.

Address reprint requests to:
Wolf E. Mehling, M.D.
Osher Center for Integrative Medicine
University of California, San Francisco
1704 Divisadero Street, Suite #150
San Francisco, CA 94115

E-mail: mehling@itsa.ucsf.edu