THE JOURNAL OF ALTERNATIVE AND COMPLEMENTARY MEDICINE
Volume 8, Number 1, 2002, pp. 5–6

Methodological Changes in the Evaluation of Complementary and Alternative Medicine: Issues Raised by Sherman et al. and Hawk et al.

DANIEL REDWOOD, D.C.

The quest for methodologies appropriate for evaluating complementary and alternative medicine (CAM) continues to challenge the research community. Two papers in this issue, Sherman et al.'s "Description and Validation of a Noninvasive Placebo Acupuncture Procedure," (pp. 11–19) and Hawk et al.'s "Issues in Planning a Placebo-Controlled Trial of Manual Methods: Results of a Pilot Study," (pp. 21–32) demonstrate efforts by thoughtful investigators to develop legitimate placebo interventions and to address standardization of delivery.

Sherman et al. show great ingenuity in their use of a toothpick and a plastic insertion tube to mimic acupuncture needling. In their first experiment, the toothpick insertions were perceived as slightly more like real needling than the real needling itself; in the second, a still impressive 53% of those receiving the simulated needling believed it "definitely" or "probably" was real acupuncture, compared to 65% of those receiving actual acupuncture. Further supporting the legitimacy of this method as a placebo or control is the difference in reported therapeutic response: 33% of subjects receiving real acupuncture reported significant improvement in their back pain compared to only 4% receiving imitation acupuncture.

Yet questions remain. The ideal placebo should be therapeutically inert; a reasonable placebo should be reasonably inert. But the authors note that noninsertive needling is actually practiced as part of some styles of acupuncture and that the effects of gentle stimulation are not well understood at this time. Therefore, they suggest that the toothpick method be considered a "minimal sham treatment," which should not be assumed to be inert. Nonetheless, based on the efficacy of patient blinding and the differences in therapeutic response to toothpick versus acupuncture needling, they recommend their method as a "reasonable control treatment for acupuncture-naïve individuals in randomized controlled trials assessing the efficacy of acupuncture for low-back pain."

Although Sherman and colleagues are clearly conscious of the subtleties in these distinctions among placebos, shams, and minimal shams, other investigators using this method in the future may lack this awareness or fail to convey it to consumers of their research. Unless a clear disclaimer to the contrary is noted in the abstract as well as the text of an article, most readers assume that methods used as placebos, controls, or shams are actually inert. A broader context for interpretation may be achieved by including a "no treatment" and/or a conventionally treated group along with the minimal sham and real acupuncture groups.

In the Hawk et al. trial, a multidisciplinary team of researchers began with the assumption that fundamental differences exist between the administration of medications and the application of manual procedures. They designed a multisite pilot study on chiropractic care for women with chronic pelvic pain, seeking to address the methodological challenges inherent in these differences. They also hoped to lay the groundwork for a larger randomized controlled trial (RCT).

Virginia Beach, VA.

The results of their pilot study appear to have raised more questions than were answered, questions that are of real importance to the future of manual therapy and CAM research. The most striking point is the wide variation of clinical response at the three sites. At site 1, the active group's improvement was substantially greater than that of the placebo group. (The placebo consisted of sham manipulation utilizing a spring-loaded chiropractic adjusting instrument set to zero, plus light effleurage-type massage.) At site 2, both groups fared about equally well. At site 3, the placebo group substantially outperformed the group receiving active treatment.

What is one to make of this? It is true that small samples can produce misleading results, but something more appears to be happening here. Despite the laudable efforts of the researchers to achieve uniformity of delivery, different chiropractors have different degrees of effectiveness, even when utilizing the same technique system. Of interest in this study is the fact that the chiropractor who elicited the most positive therapeutic responses was the one who went "outside the box" most often, bending the rules of the study when he judged that this would be most helpful to the individual patient.

The treating chiropractors were supposed to apply flexion-distraction technique (a variation of an osteopathic technique using angled traction with a specially designed treatment table) and manual trigger point (TP) therapy on all patients in the active treatment group. However, the doctor at site 1 performed the TP work on only half of his active treatment patients and added an additional procedure in 20% of cases. He appears to have prioritized the interest of the patients above that of the data collectors, and his results were by far the best in the study. There is a very important message here regarding individuality of treatment, one that we must never push too far into the background in our quest for data validity.

Hawk et al. also note "the lack of a definitive, single 'active' agent in chiropractic care," opining that "the reason for patients' improvement may not be the specific manual procedures but a psychophysical *gestalt* encompassing the entire clinical encounter and belief systems of the practitioner and patient."

Therein lies a great conundrum of CAM research. Even if it were possible to standardize adequately the delivery of manual therapy, acupuncture, or massage, standardizing the doctor–patient *gestalt* is an absurdity. This is cause for deep humility about what we know and what we are capable of knowing. As we seek to refine our research methodologies, we must be ever mindful that the healing encounter likely includes significant factors that may never be quantifiable. These perhaps ineffable qualities, which may be crucial to many a patient's recovery, must not be swept out of the research environment through attempts at enforced standardization that shackle the treating doctor and fail to fully honor the fullness of the therapeutic *gestalt*.

Address reprint requests to:
Daniel Redwood, D.C.
1645 Laskin Road, Suite 103
Virginia Beach, VA 23451

E-mail: danredwood@aol.com