A Review and Analysis of Placebo Treatments, Placebo Effects, and Placebo Controls in Trials of Medical Procedures When Sham Is Not Inert

STEPHEN BIRCH, Ph.D., Lic.Ac. (US), M.B.AcC. (UK)

ABSTRACT

Researchers examining the efficacy of medical procedures make assumptions about the nature of placebo. From these assumptions they select the sham interventions to be used in their trials. However, placebo is not well defined. A number of definitions are contradictory and sometimes misleading. This leads to problems in sham-controlled studies of medical procedures and difficulties interpreting their results. The author explores some of the contradictory definitions of placebo and assumptions and consequences of these. Principal among these is the assumption that the placebo is inert when it is not, which introduces bias against the tested medical procedures and devices. To illustrate the problem, the author examines the use of sham procedures in clinical trials of the medical procedures surgery and acupuncture in which the sham was assumed to be inert but was not. Trials of surgery and acupuncture should be re-examined in light of this.

INTRODUCTION

The literature on placebo is confusing evidenced by recent books and conferences. Authors claim placebo effects are small and large. Some claim placebo does not exist. Placebo effects vary with different treatments; effects for drugs are different than for devices; different-colored placebo pills change the placebo effect; how treatment and placebo are explained in the study changes the placebo effect and can trigger its opposite, the nocebo effect. Placebo effects interact with other treatment effects, making it difficult if not impossible to control for them. Some argue that placebo controlled trials may not be possible in complex interventions because they cannot separate placebo from other treatment effects. How should this bewildering array of opinions, claims, counterclaims, and contradictory findings be understood? When something supposedly so ubiquitous as the placebo effect is not really understood, what does it mean for clinical research?

Placebo is a very difficult concept to define. Although commonly understood as, “A pharmacologically inactive agent given to a patient as a substitute for an active agent and where the patient is not informed whether he is receiving the active or inactive agent,” recent descriptions highlight some of the inherent ambiguities.

One assumption is that placebo is by definition inert. “The one thing of which we can be absolutely sure is that placebos do not cause placebo effects. Placebos are inert and don’t cause anything.” The inert placebo triggers complex internal mechanisms in a patient that lead to observable placebo effects, a process especially related to how the patient attributes meaning. However, contrary definitions of placebo treatment include noninert therapies, so long as they are not thought to be specific for the condition under treatment.

The Shapiros described placebo as any therapy that has not been proved effective in double-blind controlled studies and that “The placebo may be an inert sugar pill, an active drug, or any treatment no matter how potentially specific or by whom administered.” The assumption that an unproven treatment can be considered a valid placebo leads to the paradox in placebo trials that a placebo is being compared to a placebo. Further, what if it is discovered later that
what was called a “placebo” according to this definition was later shown to be an active specific treatment?26–28 Because the definition will have contradicted itself, is not the definition now logically false?

Some have discussed the distinction between “perceived placebo effects” and “true placebo effects.”17,23,29,30 Perceived placebo effects lump the “true placebo effect” together with a number of “nonspecific effects,” including regression to the mean, natural course of the disease.30 However, this model makes no clear statement about the nature of the placebo treatment itself, and leaves it up to the imagination of the researcher to define. The same confusion can be introduced as was seen in the last definition.

Finally, a recent approach is based on an “operational” definition: “For the purpose of this article, placebo is an intervention used in a clinical trial that is administered with the intention of mimicking some other intervention so that an unbiased comparison can be made.”31 This definition explicitly assumes that any attempt at mimicking a therapy is a placebo intervention and is thus similar to the definition by Shapiro and susceptible to the same errors and contradictions as that definition. It includes in the definition all mistakes made by researchers in defining, constructing, and implementing their mimic therapy.

Although there is a lot of agreement that a number of different factors such as treatment context, patient expectations, and enthusiasm of the practitioner can contribute to placebo effects,1–3,5,8,23 there is considerable confusion about what constitutes a placebo therapy. Is the placebo to be an inert therapy or is it not inert, and include other “non-specific effects” caused by trying to mimic a therapy? This paper explores some of the consequences of this confusion.

Why the placebo control?

Why are randomized placebo-controlled clinical trials, or “efficacy”32,33 trials conducted? The simplest answer is that “. . . tests of efficacy always presume an underlying mechanism, and so a test of efficacy is necessarily a test of an assumed mechanism.”32 Because a therapy has specific effects (mechanisms or active ingredients) associated with it and a variety of other effects not specific to that therapy, then placebo-controlled studies are conducted to examine the size of the specific effects by controlling for (subtracting) the nonspecific effects.13,33–35 These nonspecific effects include effects caused by placebo responses, and effects such as regression to the mean, natural course of the disease.29,30 Randomization is thought to distribute the latter effects equally between treatment groups.5,30,33,36 Placebo effects are controlled for if the test or active treatment is compared to a “sham” treatment. In drug trials this is not difficult to do provided double blinding (blinding of therapist and patient) is used and the active and sham treatments appear the same.33 The sham treatments involve using a chemically inert similar-looking pill with no active chemicals in it and blinding those administering the treatment so that neither they nor the patient know what is being given. In such trials, the sham intervention is inert. Randomization is also usually thought to distribute the amount of placebo responses equally between the treatment groups. Under these basic conditions efficacy or randomized placebo-controlled clinical trials are conducted. This type of study is also known as an “explanatory” trial33 because it is designed not only to examine if the treatment is more effective than its comparison treatment, but also to explain the effectiveness in relation to the active components of the test treatment (e.g., chemical compound).

However, what about therapeutic techniques such as surgery and acupuncture? How can one guarantee that the sham treatment is inert like a placebo pill? How can one imitate surgery while not initiating other nonplacebo-related treatment effects resulting from trauma, etc.? Is it possible to make inert sham treatments in ACT (alternative and complementary therapies) like acupuncture?

If it is not possible to use inert sham treatments, how does one deal with the additional beyond-placebo “nonspecific” effects of the technique or device? When sham treatments are used that are not inert, what impact does this have on the understanding and interpretation of sham-controlled clinical trials? Many researchers have assumed that the sham treatments in medical technique trials are inert when they are not. This opens the door to difficult questions about how to interpret sham trials of, for example, acupuncture and surgery.

The inertness of “sham”: implications when wrong

What are the implications of a faulty assumption that the sham procedure is inert? It introduces bias against finding the tested therapy to be effective. As de Craen and colleagues demonstrated, “if the placebo has a large deviation from inertness it is obvious that the results of any trial will be biased, but even low placebo activity could bias the comparison when the absolute treatment effect is small.”26 Besides the threat to internal validity that a noninert placebo can create,32,39 three other complicating problems occur.

Researchers risk overestimating the size of the placebo effects because they have inadvertently included other effects that are not placebo related,5,36 making it more difficult to show that the real therapy is significantly more effective than sham.23 This problem is compounded by the next.

Sample sizes often are based on an estimate of the expected size of the test and control treatment effects. If the sham intervention is assumed to be inert, no greater than placebo, calculations often are based on general ideas about the size of the placebo effect. Usually this is taken to be around 30% based on Beecher’s original calculations (however, these have since been found to be erroneous).17,19 To demonstrate a significant difference between the 30% ef-
fectediveness of the sham (placebo) and the expected effec-
tiveness of the intervention itself (say 60% to 70%), rela-
tively small sample sizes are necessary.37 However, if the
sham intervention turns out to be greater than the estimated
30% effectiveness because the sham is not inert (e.g., if the
sham is 50% effective), then the smaller sample size based
on a comparison of 30% versus 70% will be inadequate to
demonstrate the effectiveness of the therapy.23,37 The ma-
majority of acupuncture trials have suffered from inadequate
sample size,37–39 and a number of the surgery trials have
had small sample sizes.40,41 Trials that assume their sham
intervention is inert when it is not are at greater risk of mak-
ing false-negative judgments. De Craen and colleagues
showed that if the effect of the noninert placebo is small, it
can bias a study if the treatment effects are small. This bias
increases if the effects of the noninert placebo are larger and
especially if they are specific to the condition under study.26

Finally, whenever a trial concludes that “therapy X is not
more effective than placebo,” it is either implied by the re-
searchers and/or understood by readers that the placebo was
inert.11 The training of most people reading published trials is
not adequate to understand the complexity of the placebo
concept and people revert to the original idea from the 1950s
that the sham procedure was inert, which usually damnsthat
therapy as ineffective.19 These various factors compound
each other to produce bias against the test therapy.

Biologic and therapeutic effects of surgical
procedures

In sham surgical trials, both groups of patients receive the
surgical incision or other invasive procedure to expose the rel-
vant organs or anatomical structures. The real surgical group
receives the surgical procedure, whereas the sham group has
the incision closed. Both groups receive the same postopera-
tive care. In the 1950s two trials of mammary artery ligation
surgery were performed for the treatment of angina.40,41 The
mock or sham surgery included making an incision and ex-
posing the relevant blood vessels, and then surgical closure of
the wound. Although these trials described the procedure as a
placebo procedure, one implied that the additional postsurgi-
cal care may have had an effect41 the other that “spontaneous
improvement in collateral circulation cannot be ruled out.”40
Thus these trials seem to imply that their “placebo” may not
have been inert. However, they have been discussed in the lit-
erature as though they were inert placebos.24,42 In the 1990s
a trial of the surgical implantation of fetal tissue in the treat-
ment of Parkinson disease used a sham procedure in which a
hole was drilled in the skull and then the wound was closed
without transplantation of tissue.43 Although this trial did not
refer to the sham surgery as a placebo procedure, an earlier
discussion of the trial44 and later researchers45–47 described it
as placebo.

It is generally held that surgery is associated with strong
placebo effects. Various reasons are offered: the absence of
regulatory demands of proof of efficacy,25 the degree of
physical discomfort from the procedure,17 and the meaning
attributed to the procedure and shedding of blood.24 The last
two reasons touch on a key and usually ignored aspect of
any surgical intervention, including the sham. The human
body has protective and reparative mechanisms that are trig-
ergged by injury. These are not placebo effects, they are di-
rect and expected by-products of the actual trauma of the
surgical incision. All surgical wounds provoke a cascade of
physiologic responses,48 which not only repair the wound,
but also trigger a range of biologic effects that could affect
other (patho)physiologic processes in the body. Local and
systemic reparative mechanisms are triggered by surgical in-
clusions, including: vasoconstriction followed by clot forma-
tion, vasodilation, migration of white blood cells into the
damaged tissues, inflammatory reactions, fibroplasia, ep-
ithelialization, and wound contraction.48,49 Additionally, the
body responds to traumatic painful injury with analgesic
mechanisms.50 It cannot be ruled out that these effects may
affect problems near the site of the injury as well as other
parts of the body. If the biologic activities resulting from
surgical incisions have not been adequately investigated for
their potential therapeutic effects, they cannot be assumed
to be therapeutically inert, a necessary feature if sham stud-
ies are to be considered placebo controlled.

Additionally, one can speculate that other effects might
result from the loss or letting of blood that accompanies the
procedure. Bloodletting has been used as a medical proce-
dure over more than two thousand years in multiple cul-
tures.51–53 It is used in modern medicine for treatment of
conditions such as hemochromatosis.54 It also has been in-
dicated for a number of other problems,55–58 and often is in-
dicated because it helps improve blood circulation.59–61 It is
not known whether the loss of blood from the surgical in-
cision on the chest of angina patients40,41 can be therapeu-
tic for angina, but bloodletting on the torso has been indi-
cated for the treatment of some cardiac diseases.61 Blood
loss triggers sympathetic responses that cause decreased
blood flow in most parts of the body, but it can cause slight
vasodilation of the coronary arteries,62 which could in prin-
ciple help trigger improved coronary artery blood flow. One
report indicates that bloodletting was specifically helpful for
angina pectoris.63 Likewise, it is not known if the letting of
blood from incisions on the skull of patients with Parkinson
disease undergoing surgical implantation of fetal tissue is
therapeutic for Parkinson’s syndrome,43 but bloodletting of
the head, neck, or upper torso has been indicated for a wide
range of problems, including some neurologic disorders,60
with one report indicating successful use for improvement of
Parkinsonian symptoms.64

The same problem is found when other sham surgical tri-
als are scrutinized. The sham and real surgical procedure for
osteoarthritis of the knee had similar effects,65 whereas the
use of leeches for bleeding on the knee was more effective
than the usual treatment.66 This suggests that the sham
surgery may have been an active therapeutic method rather than a placebo treatment because it involved making incisions that cause bleeding, and bleeding of the knee was demonstrated effective for osteoarthritis of the knee. Researchers conducting sham surgical trials have not considered the possibility of these effects; thus, they have not been subjected to relevant investigations and have not been controlled for.

However one interprets placebo responses, they will most likely be inseparable from the biologic effects of the surgical incision (both in the sham and real procedures). When placebo effects interact with or are inseparable from other effects, placebo-controlled trials become difficult if not impossible,\(^5,15\) as they cannot test the purported mechanism (specific effects) of the therapy.\(^13,32\) In other words, placebo-controlled trials of surgical interventions may not be possible because there is no way to factor out or control for the actual placebo effects.

**General biologic effects of acupuncture**

Acupuncture is an ancient therapy with a variety of different explanatory models in its clinical practice.\(^57-70\) However, there are a few important principles that are generally agreed upon and common to virtually all these models. Treatment involves applying selected techniques to selected points. Traditional and modern theories and diagnoses guide and dictate the selection of which techniques at which points.\(^68,71\) Clinical trials testing the specific claims of acupuncture generally have tried to focus on testing the efficacy of applying specific techniques and/or specific points.\(^72-74\) However, insertion of needles into the body can stimulate other effects not dependent on the locations of stimulation and are thus nonspecific. Regardless of where the needles are placed, a range of nonspecific mechanisms can be activated.\(^77,74-80\) These include heterosegmental analgesic mechanisms,\(^80,81,82\) homosegmental analgesic effects,\(^83\) microcirculatory effects,\(^84\) and relaxation response effects.\(^78,85\) One can speculate a cascade of normal physiologic effects involved in prevention of infection and healing of damaged tissues after the insertion of any needles anywhere on the body.\(^49\) These effects occur in addition to placebo effects. An international group of researchers discussing these effects recently concluded,

we cannot be satisfied that a truly inert intervention is possible as the control treatment in acupuncture studies. Therefore the sham intervention cannot be considered equivalent to placebo in the same way that a placebo pill is considered to be an inert intervention in a placebo controlled pharmaceutical trial.\(^74\)

This is a point raised by others.\(^11,79,80\) However, it has been common for researchers to not separate these different effects and lump them together with any placebo effects.\(^74,75\) It is also possible that the different treatment effects of acupuncture may interact,\(^74\) evidenced by the finding of the endorphin involvement in specific needling of acupoints,\(^86\) heterosegmental analgesic nonspecific effects,\(^81\) and placebo effects.\(^87\) Such interactions probably make placebo-controlled trials of acupuncture difficult if not impossible.\(^5,15\) Further, it has been postulated that the principal purpose of the traditionally based diagnoses and treatments in acupuncture is to target a specific improvement in the innate healing abilities of each patient, which may employ the same mechanisms of action as those harnessed by placebo.\(^74\)

There is growing evidence for the claim that placebo harnesses self-healing mechanisms.\(^22\) Because the traditionally based acupuncture treatment targets improvement of these mechanisms, controlling for this in placebo-controlled trials of acupuncture may thus require attempting to control for the specific mechanisms and effects of the therapy being investigated, which contradicts the reasons for conducting placebo-controlled trials.\(^32,74\) Some of the nonspecific effects that have been documented with needling are involved in the natural healing mechanisms of the body, which also may employ some of the same pathways that placebo activates. These potential problems with sham acupuncture treatments also need to be investigated further.

Different sham methods have been employed in sham trials of acupuncture, including invasive and noninvasive sham methods.\(^74,75,88\) Although some researchers consider that invasive sham techniques produce only placebo effects,\(^45,72\) there is general agreement that any invasive sham acupuncture cannot be inert.\(^74,79,88\) It is not clear how active the noninvasive sham methods are. The question of whether they are truly inert has been raised.\(^74,75,79\) Putting aside questions about the practical value of the noninvasive sham for research,\(^75\) any noninvasive sham acupuncture method at least must be investigated to establish if it produces unintended physiologic effects from touch, pressure, and so on, so that they can be controlled for, or whether it is truly inert.\(^79\) The recently developed noninvasive sham needle of Streitberger and Kleinhenz is considered to be a “placebo needle” by its inventors,\(^89\) but reports of its use clearly indicate otherwise. One researcher in Germany caused bleeding in one patient as a result of the required mechanical stimulation of the noninvasive sham (Birgit Seybold, personal communication, 2003).

Recently various authors have suggested that acupuncture is a complex intervention,\(^15,16\) because for example, the interview and discussion with patients are a necessary part of treatment, and a number of different methods may be used in treatment. Thus, it may not be possible to conduct placebo-controlled studies of acupuncture because they run the risk of mixing other specific and nonspecific effects with any placebo effects in the sham group and thus generate false-negative results.\(^15\) Other research models need to be investigated and developed for complex interventions.\(^14-16\) This argument parallels the preceding discussions and compounds the problems identified.
Validity of controls in sham intervention trials

It has been argued that in sham studies in which the therapist cannot be blinded, it is essential to assess the credibility of the treatment and success of blinding,\textsuperscript{73,75,79,90} something that is curiously missing in sham surgical trials. Methods have been developed to assess the blinding and credibility of the treatment,\textsuperscript{79,90} but these are not sufficient to allow control of the physiologic nonspecific effects. Further methods have been developed to attempt these additional controls,\textsuperscript{78,91} but so far only a few trials have attempted to make distinctions between placebo and control for the physiologic “nonspecific” effects and placebo “nonspecific” effects.\textsuperscript{85,91,92} Although the methodology\textsuperscript{74,75,78,93} has not yet been generally accepted, it is the only published method developed to try to address these nonplacebo nonspecific physiologic effects. Trials that cannot or have made no attempt to separate placebo from these other effects become difficult to interpret. Reviewers may need to reclassify many trials before they attempt to interpret their results. Some reviews have attempted to analyze efficacy of acupuncture compared to placebo,\textsuperscript{94–96} which for the reasons discussed here cannot be accurate. Such reviews need to restate the comparison as acupuncture versus sham. Some reviews have made the comparison to sham acupuncture,\textsuperscript{99,111} which is a more accurate comparison. However, in both cases it needs to be explicitly stated that the sham is an active treatment of unknown effectiveness. If in attempting to investigate the efficacy of purported mechanisms of acupuncture it is not possible or feasible to control for all the nonspecific effects of needling, trials that investigate the effectiveness of acupuncture might do better to use pragmatic models\textsuperscript{53} together with laboratory methods to investigate purported mechanisms.\textsuperscript{92}

Attributing meaning: the need for change and reevaluation

In sham trials of acupuncture researchers project their definition of placebo effects onto patients and assume that observed effects in the control group match the projected placebo effects. Sham surgical trials can be seen to make similar conceptual mistakes. Surgeons do not really believe that the physiologic effects of making incisions and causing blood loss can be therapeutic; rather, they believe that the specific surgical procedure produces the therapeutic effects. These assumptions lead them to conclude that their sham is inert. This introduces bias against surgery and acupuncture in these studies.

Clinical trials of medical procedures need to be re-examined and redesigned to remove this bias. “Studies involving placebo effects must be designed to separate actual placebo effects from various artefacts. These artefacts include investigator, observer, and patient bias, specific biological effects attributable to physical or chemical properties of the placebo” (emphasis added).\textsuperscript{99} The biologic effects of any intervention need to be understood in addition to any projected therapeutic effects of that intervention. When these unintentionally produce therapeutic effects in addition to the projected therapeutic effects, trials must take this into account. The practice of calling noninert sham treatments placebo in publications should stop.\textsuperscript{11}

Early trials of cholesterol-lowering agents in heart disease used capsules with olive oil or corn oil as the placebo treatment. However, later research showed that these two oils are both capable of reducing low-density lipoprotein and thus were inappropriate as placebo treatments.\textsuperscript{56} It is now clearly known that surgical and acupuncture sham procedures are not inert; this must be dealt with in future trials and interpretation of previous trials.

CONCLUSIONS

The assumption that a sham medical procedure is inert when it is not introduces bias against the medical procedure for which the sham is used as a control. It makes it difficult to find significant differences between the sham and real procedure. The common conclusion from trials of medical procedures that the procedure was no better than placebo is misleading to the reader and is inaccurate. In these studies the conclusion is either wrong, or it is not possible to show that it is correct, thus making many of these trials difficult if not impossible to interpret and reviews of these trials open to question. Trials of medical procedures should stop stating that they are “placebo-controlled” or that their sham procedure is a placebo procedure unless they can present evidence that their sham is in fact inert and have validated it as a placebo treatment. The practice of making \textit{a priori} assumptions about what are acceptable placebo effects in a trial needs to be re-examined, because this has been responsible for introducing faulty assumptions about the nature of placebo and bias against the therapy being investigated. Often these \textit{a priori} assumptions are based on ignoring published literature about treatment effects and biologic mechanisms that are known to arise in response to the proposed treatment methods. Placebo-controlled surgical trials appear to be impossible. Placebo-controlled acupuncture trials cannot use invasive sham procedures. If a noninvasive sham procedure is used in an acupuncture trial, the sham must be investigated for physiologic effects to establish that it is in fact inert, and more complex and complete control procedures and methods must be used to separately control for other nonspecific physiologic effects. Additionally, the conceptual and theoretical basis of acupuncture practice needs to be better considered, to ensure that the natural healing processes that are the target of therapy are not confused with placebo, and are “controlled” for in sham trials. This is an extremely complex and difficult area in which few studies have attempted
comprehensive controls and it remains to be seen how practical, acceptable, and feasible proposed methods are for doing this.

ACKNOWLEDGMENTS

Thanks to Jackie C. Wootton, M.Ed., for her helpful comments on the manuscript.

REFERENCES


Send reprint address requests to:
Stephen Birch, Ph.D., Lic.Ac. (US) M.B.Ac.C. (UK)
Stichting (Foundation) for the Study of Traditional East Asian Medicine (STEAM)
W.G. Plein 330
1054 SG Amsterdam
The Netherlands
E-mail: sjbirch@gmail.com