Carnegie Mellon University
Research Showcase @ CMU

Department of Philosophy | Dietrich College of Humanities and Social Sciences

2002

Placebos That Harm: Sham Surgery Controls in Clinical Trials

Alex John London
Carnegie Mellon University

Joseph B. Kadane
Carnegie Mellon University

Follow this and additional works at: http://repository.cmu.edu/philosophy

Part of the Philosophy Commons

Recommended Citation
.

This Article is brought to you for free and open access by the Dietrich College of Humanities and Social Sciences at Research Showcase @ CMU. It has been accepted for inclusion in Department of Philosophy by an authorized administrator of Research Showcase @ CMU. For more information, please contact research-showcase@andrew.cmu.edu.

Statistical Methods in Medical Research 2002; **11**: 413–427

Placebos that harm: sham surgery controls in clinical trials

Alex John London Department of Philosophy, Carnegie Mellon University, Pittsburgh PA, USA and **Joseph B. Kadane** Department of Statistics, Carnegie Mellon University, Pittsburgh PA, USA

Recent debates over the use of sham surgery as a control for studies of fetal tissue transplantation for Parkinson's disease have focused primarily on rival interpretations of the US federal regulations governing human-subjects research. Using the core ethical and methodological considerations that underwrite the equipoise requirement, we find strong prima facie reasons against using sham surgery as a control in studies of cellular-based therapies for Parkinson's disease and more broadly in clinical research. Additionally, we believe that these reasons can be generalized to apply to the use of other placebo controls that carry significant risks of positive harms in and of themselves. As a result, our arguments are centrally relevant to the emerging drive to subject therapies with a surgical component to the same rigorous standards of evaluation as those governing the approval of new pharmaceuticals.

1 Introduction

Ethically responsible and scientifically sound human-subjects research seeks to advance the state of medical knowledge and clinical practice without knowingly compromising the welfare and integrity of individual trial participants in the process. The degree of difficulty that one sees for achieving this goal in actual practice depends on a variety of factors, including one's conception of the purpose of clinical research as a form of scientific inquiry, and one's views on what is required in order to safeguard the welfare and integrity of individual trial participants. Despite a burgeoning literature on research ethics and methodology, divergent views concerning some of these issues continue to generate controversy about the choice of appropriate controls for certain kinds of clinical research. This was illustrated most recently by the public debate that erupted over the use of sham surgery as a control for trials of cellular-based therapies for Parkinson's disease.[1,2]

Parkinson's disease is a degenerative neurological disorder characterized by a loss of dopaminergic neurons in the basal ganglia of the brain, producing tremors, muscle rigidity, and abnormal movements. The standard treatment, oral doses of the dopamine precursor levodopa, reverses these symptoms in most patients, but over time its effects tend to wear off as its side effect profile increases. In the clinical trial described by Freeman *et al.*,[1] 36 subjects were randomized to one of three arms, two

Address for correspondence: Alex John London, Assistant Professor, Department of Philosophy and Centre for the Advancement of Applied Ethics, Carnegie Mellon University, 135 Bakes Hall, Pittsburgh, PA 15213, USA. E-mail: ajlondon@andrew.cmu.edu

© Arnold 2002 10.1191/0962280202sm300ra

receiving bilateral fetal nigral transplantation and one receiving bilateral placebo surgery. Throughout the study all subjects continued to receive standard medical therapy. Just as members of the first two arms underwent two surgical procedures, the control group received two placebo surgical procedures that were designed 'to provide an equivalent experience for the subjects and their family members.' Each placebo procedure involved the placement of a stereotactic frame—a frame attached to the cranium with surgical screws, which allows for accurate location of targeted areas in the brain—a magnetic resonance imaging scan, a positron-emission tomography (PET) scan, the administration of general anesthesia, and the drilling of two dime-sized burr holes into the skull through scalp incisions. In the control group, the burr holes did not penetrate the dura and no material was injected into the brain. All subjects, however, received intravenous antibiotics and cyclosporin for six months after surgery.

In its choice of controls, this study is not unique. Controls that are similar in their substantive details have been used in at least one other trial of fetal nigral cells,[3] and another trial of porcine nigral cells.[4] Additionally, the use of sham surgeries in Parkinson's research appears to be part of a larger movement to subject therapeutic interventions that involve a surgical component to the same rigorous standards of evaluation as those governing the approval of new pharmaceuticals.[5] As a result, there is a pressing need for a more general analysis of the conditions under which it is acceptable, if ever, to use controls of this kind in human-subjects research.

In what follows, we argue that the use of sham surgery in clinical research raises fundamental questions about the goals of human-subjects research and the best ethical and methodological framework for evaluating research protocols. In particular, we argue that one of the main sources of disagreement about the use of sham surgery controls stems from a conflict between two basic conceptions of clinical research, what we refer to as the clinical or pragmatic model and the basic science model. We intend to illustrate how a careful examination of this conflict provides an especially rich context in which to explore the relationship between two conceptual frameworks for evaluating clinical research. In particular, the debate over the use of sham surgery in trials for Parkinson's disease has focused largely on whether the trial design violates the requirements set out in the US Code of Federal Regulations. It is interesting to note, however, that at a time when the concept of equipoise is widely used to evaluate the choice of controls in pharmaceutical research, it has received little explicit attention in the debate over sham surgery. In its most basic formulation, equipoise represents a state of genuine and credible doubt about the relative therapeutic merits of some set of interventions that target a specific medical condition, and to many it represents a necessary condition for ethically acceptable human-subjects research.[6–8] To some degree, the relative inattention to equipoise in this debate no doubt reflects the fact that it does not have the same regulatory status as the rules articulated in the federal regulations. It may also reflect ongoing uncertainties about the best way to explicate the concept of equipoise itself. Nevertheless, we argue that critical reflection on the core ethical and methodological considerations that underwrite the equipoise requirement support strong *prima facie* reasons for rejecting the use of sham surgical controls. We elucidate some of the conditions that would have to be obtained to override these *prima facie* reasons.

2 Sham surgery and the current US regulatory framework

In the USA, all federally funded research—as well as nonfederally funded research conducted at institutions that have agreed to abide by federal regulations—must comply with federal regulations for the protection of human subjects.[9] The recent debate over the use of sham surgical controls in clinical research has focused special attention on two of the requirements laid out there. The requirements are:

1) Risks to subjects are minimized: (i) by using procedures that are consistent with sound research design and that do not unnecessarily expose subjects to risk, and (ii) whenever appropriate, by using procedures already being performed on the subjects for diagnostic or treatment purposes.
2) Risks to subjects are reasonable in relation to anticipated benefits, if any, to subjects, and the importance of the knowledge that may reasonably be expected to result. In evaluating risks and benefits, the IRB (Institutional Review Board) should consider only those risks and benefits that may result from the research (as distinguished from risks and benefits of therapies subjects would receive even if not participating in the research). The IRB should not consider possible long-range effects of applying knowledge gained in the research (for example, possible effects of the research on public policy) as among those research risks that fall within the purview of its responsibility.

In defense of their use of a placebo surgical control, Freeman *et al.* argue that their research meets these conditions. First, they argue that the inclusion of a placebo surgery arm is essential to answering the chosen research question: is fetal-tissue transplantation a safe and effective treatment for Parkinson's disease and, if so, are the observed benefits the result of the fetal-tissue transplant or of some associated placebo effect? They then assert that within their clinical trial 'the risks to participants are reasonable and have been minimized as far as possible' (p. 991). In particular, members of the control group continue to receive standard medical therapy for Parkinson's disease, a partial burr hole is used instead of penetrating the dura and inserting inert material in the brains of subjects, and their renal function is monitored to detect adverse reactions to cyclosporine. As interpreted by Freeman *et al.*, if their study design is sound, and if the risks to subjects have been minimized as far as is consistent with sound research design, then their research meets the first of the above regulatory requirements. As a result, they seem committed to the view that one does not *unnecessarily* expose subjects to risk if one ensures that the risks have been minimized as far as is consistent with a sound trial design.

Although the condition (2) in the above cited regulations requires that risks must be reasonable in relation to the benefits subjects may receive from participating in the research, it is also clear that subjects need not themselves receive any such benefit for research to be acceptable. What the second condition does require is an evaluation of the reasonableness of the risks posed to participants in relation to the importance of the knowledge that may result from the research.[10] To be clear, the second condition does not indicate what kind of questions clinical research ought to target, only that the knowledge to be gained from the research be important enough to justify the risks incurred by subjects. For Freeman *et al.*, then, the importance of the research question,

combined with their attempts to minimize risks within the boundaries of a rigorous clinical trial design, is also sufficient to meet the second of the above regulatory requirements.

It is worth noting that within this general understanding of the requirements for acceptable medical research, the use of sham surgical controls does not raise special or unique ethical concerns. Although one could accept the outlines of this approach and still object to the use of sham surgical controls on independent grounds, there is a fairly straightforward sense in which this framework militates against such a move. More specifically, this is because the questions of whether risks have been minimized and whether they are necessary are seen as largely posterior to, and circumscribed by, what are portrayed as the independent requirements of a sound clinical trial design. In other words, as long as the chosen research question cannot be settled by a trial that poses less risk to subjects, then the use of a sham surgery control may be a necessary component of a sound clinical trial design. Additionally, as long as the risks associated with the sham surgery have been minimized as far as possible without vitiating the integrity of the trial design, and the remaining risks are reasonable in relation to the value of the knowledge to be gained from the study, then sham surgical controls are on a par with other nontherapeutic elements of a clinical trial that may pose some degree of risk to participants.

Freeman *et al.* bolster their claims by noting two additional facts. First, their protocol was approved by several institutional review boards and monitored by an independent performance and safety monitoring board appointed by the National Institutes of Health (NIH). Second, each of the 36 subjects gave their informed consent to participate and Freeman *et al.* argue that this not only constitutes the fulfillment of a necessary condition for ethical research, but represents the judgment of the subjects that the risks of participation were reasonable in relation to the potential benefits.

In her critical evaluation of the use of sham surgical controls in studies for Parkinson's disease, Macklin[2] challenges the justification offered by Freeman *et al.*, on each of the above points. To begin with, Macklin rightly notes that the informed consent of participants is a necessary, though not sufficient condition for ethical research. As such, researchers have an independent obligation to ensure that the terms to which subjects are asked to consent meet additional ethical standards for at least two reasons. The first, noted by Macklin, is primarily subject-centered and takes the persistence of what is referred to as the 'therapeutic misconception' as evidence that the process of informed consent is not always an adequate procedural safeguard of subject welfare. The second, which she might have noted, is researcher-centered and focuses on the researcher's moral responsibility for foreseeable and preventable outcomes that result from trials that they design and implement.[11]

Next, Macklin argues that, 'performing a surgical procedure that has no expected benefit other than the placebo effect violates the ethical and regulatory principle that the risk of harm to subjects must be minimized in the conduct of research' (p. 993). On this reading of the regulatory requirements, the injunction to minimize risks to subjects is not subordinated to or circumscribed by the terms of the chosen clinical trial design. Rather, it stands as an independent consideration that must be factored into the choice of trial design and which, as a result, may provide a reason to reject what might otherwise appear to be an optimal trial design.

Additionally, Macklin sees a fundamental difference between the use of placebo controls in pharmaceutical research and the use of so-called placebo surgery. In the former case, because the placebo is not known to produce any adverse effects, subjects randomized to the placebo arm do not risk positive harms or what we will call harms of commission. Rather, if there are possible harms associated with receiving a placebo in pharmaceutical research they are what we will call harms of omission or exclusion. These are harms that subjects might incur as the result of opportunity costs associated with receiving a placebo instead of some more effective treatment option. In contrast, placebo surgeries carry with them known risks of positive harm and Macklin uses this fact to support her judgment that 'it is undeniable that performing surgery in research subjects that has no potential therapeutic benefit fails to minimize the risk of harm.'

Finally, for our present purposes, Macklin argues that there is widespread disagreement over whether the risks associated with the use of sham surgery in the present case are outweighed by the associated potential benefits. While the research of Freeman *et al.* was approved by a set of review boards, other institutional review boards have denied approval to similar research and some researchers have refused to include sham surgical controls for studies of Parkinson's disease on the ground that they were judged to be unethical.[5] As a result, she argues, at best the risk–benefit ratio is uncertain and, at worst, unfavorable. Given this lack of consensus on an issue of such significance, she argues that a more cautionary stance should be taken and that sham surgery controls ought not to be used.

Macklin might have argued, additionally, that the above federal regulations endorse a presumption that the current treatment is a suitable comparison for a new treatment. Freeman *et al.* reduced the risks to the control group by maintaining their standard medical therapy in accordance with 1.ii, but the question remains whether the inclusion of the sham procedure was an 'appropriate' addition to the procedures already being performed on the subjects for diagnostic or treatment purposes. Again, ambiguity is introduced by the words 'whenever appropriate' and it remains unclear how an IRB is supposed to distinguish appropriate from inappropriate cases.

It is likely that a number of disputes in research ethics turn on the differences of interpretation illustrated in this case.[12] Both sides of this debate can draw support from the letter of the federal regulations and from broader intuitions about the goals of research. In the following sections, however, we argue that the intuitive core of the principle of equipoise provides a less ambiguous standard for evaluating the use of sham surgical controls. First, however, we turn to a point about the purpose of clinical research.

3 The purpose of clinical trials

There are two kinds of purposes for clinical trials we think it is useful to distinguish. The first relates directly to clinical practice and is aimed at comparing the net therapeutic effects of two or more therapies for an illness. The second relates more narrowly to specific scientific issues concerning, for example, the mechanisms by which therapies operate. Following Schwartz, Flamant, and Lellouch[13] we refer to the first purpose as 'pragmatic' and the second as 'explanatory.' Of course, understanding mechanisms can

lead to new candidate therapies in the future, so this distinction is not as sharp as it might be. Roughly the difference we have in mind is analogous to the difference between a practicing civil engineer, who wants to build a bridge that will stand, and a physicist, who wants to understand how Newton's laws apply to bridges. In medicine, pragmatic trials compare alternative treatments with the goal of determining which one is preferable under conditions that most closely resemble those of the current clinical practice. They answer the clinician's question: Is Pill A better than aspirin for relieving certain aches and pains? In contrast, explanatory trials fix the conditions of the study so as to maximize the likelihood of discovering the biological properties or mechanisms of the study intervention. They answer the scientist's question: What are the chemical pathways through which aspirin achieves the effects that it has?

Clearly there is social benefit to the pursuit of each goal and it would be unwise not to recognize the importance of greater understanding in the realm of medical science for the advancement of clinical practice. At the same time, if the history and philosophy of science reveal anything about scientific progress, it is that the boundaries of knowledge are often pushed back in a less orderly and systematic way than science text books would lead us to believe and that they have yet to be pushed back by a single critical experiment. While responsible research must strive for rigor, our understanding of what counts as rigorous inquiry must also reflect the fact that not all questions can be settled at once by a single, compendious trial. Special caution is warranted, however, when the most comprehensive trial design pushes the boundaries of what is ethically acceptable.

The study of Freeman *et al.* attempts to span both of the above purposes. The first, to establish whether fetal-tissue transplantation is a safe and effective treatment, is the most clinically relevant goal. The second, to establish whether observed outcomes are due to specific effects of the implanted tissue or some set of nonspecific effects, is more scientific or explanatory in nature. We believe that the first of these purposes could be accomplished by a trial that would be both scientifically sound and ethically acceptable and, furthermore, that the results of this trial would have important implications for whether and how the second question ought to be pursued.

A trial in which fetal-tissue transplantation (along with standard medical therapy) is compared against a control group receiving only standard medical therapy would have addressed the question that is most relevant from the clinical perspective: what is the relative net therapeutic advantage of standard medical therapy versus fetal-tissue transplantation as it would be performed on patients with Parkinson's disease? Because the net therapeutic advantage of an intervention represents a robust measure of its performance, this measure includes effects on treatment goals as well as side effects associated with the intervention's biological action and mode of delivery.[7] Currently, it is not possible to implant fetal tissue into the brains of subjects without performing a host of diagnostic and surgical procedures. Furthermore, unless they are performed in the service of fetal-tissue transplantation, or some other therapeutic endeavor, there would be no reason to perform these diagnostic and surgical procedures on patients. In other words, fetal-tissue transplants cannot occur without these procedures and these procedures do not in themselves represent viable treatment alternatives for Parkinson's disease.

Trials that take a more analytical approach to fetal-tissue transplantation have the explanatory merit of providing data about the specific impact of a particular array of fetal negral cells on measured outcomes. However, they provide this data by exposing

subjects in the control arm to a range of diagnostic and surgical procedures that carry significant risks and burdens in their own right. Furthermore, although they are able to isolate the specific contributions of the cell transplants, the extrapolation of this data to the standard clinical context is not straightforward. In order to isolate the component variable to be studied, the clinical circumstances of the control group have to be altered beyond the standard baseline as when, in the current case, the control group is subjected to an array of procedures that carry significant risks and burdens in their own right.

In the pragmatic trial that we are recommending, the treatments are defined broadly. Standard medical therapy is compared against standard therapy with the addition of the package of diagnostic and surgical procedures that are required to implant a particular array of fetal negral cells. The data provided by such a trial address the question that is most relevant from the clinical perspective, namely, which of these two interventions has the superior net therapeutic advantage? Furthermore, if fetal-tissue transplantation is found to be inferior to standard medical treatment, then it is not clear that there would be further scientific questions to pursue. At best, further basic research would have to be performed before the procedure in question could be a candidate for therapeutic use.

As a result, we believe that in this case, explanatory trials should not be carried out until it is clear that there is a scientific issue worth pursuing in a trial involving human subjects. If fetal-tissue transplantation were found to be superior to standard medical therapy in a pragmatic trial then, having demonstrated a therapeutic benefit, any number of scientific questions might arise concerning the causal pathways through which that benefit was achieved. We believe, however, that the nature of the specific questions that might thus arise will depend on the actual results of such a trial. If such a trial were to demonstrate that transplanted fetal nigral cells successfully implanted and increased dopamine production and if subjects receiving the transplants scored significantly better than the controls on other measured outcomes, the probability that the observed effects were due to the surgical component alone might be judged to be low. If fetal-tissue transplantation compared favorably to standard medical therapy, and there were credible doubts among informed members of the medical community as to whether the surgical component of the procedure alone was responsible for improved outcome measures, it might then be possible to justify the inclusion of a sham surgical component in the control arm.

In the following section we argue that our proposal has the significant merit of ensuring that each of these possible clinical trials will begin in, and be designed to disturb, the state of equipoise. We also argue that without some reason to believe that transplantation is a viable treatment option, and that sham surgery might be as beneficial as transplantation, but perhaps have fewer risks of dangerous side effects, such a trial would appear to violate the equipoise requirement.

4 Sham surgery and the principle of equipoise

The concept of equipoise holds strong intuitive appeal because of the way it proposes to reconcile the interests of scientific progress and the interests of individual

research subjects. The basic core of the concept can be stated as follows (adapted from London, 2001):[14]

> **Principle of Equipoise**: (1) Equipoise exists between interventions I_1 and I_2 relative to problem P for subjects with well-defined medical characteristics S, just in case credible doubts exist about the relative net therapeutic advantage of I_1 and I_2 for treating P in subjects S and (2) there is no intervention I_3 that is preferable to either or both I_1 and I_2 for treating P in subjects S.

For this principle to provide meaningful guidance in clinical practice a number of additional clarifications need to be made, and there remains considerable debate over how to do this in a way that achieves the requisite clarity without sacrificing the plausibility of the initial intuitive idea. For example, *which factors* ought to be included in the measure of an intervention's net therapeutic advantage? This is an issue about which researchers and subjects may disagree since interventions may differ in their attractiveness to subjects depending on their individual treatment goals and health priorities. The possibility of such a divergence raises the further question of *who* ought to decide whether equipoise exists, expert physicians, some broader group that would include nonphysicians, or perhaps individual patients? Additionally, what *epistemic standard* must be met before equipoise is disturbed? In other words, when are doubts credible and where should the burden of proof be placed?

An additional issue that has not received sufficient critical attention in the literature on equipoise concerns the level at which equipoise should be applied to a trial population. To see why this issue deserves more attention in this context than it has received, we must first look briefly at the moral underpinnings of the equipoise requirement. One reason the existence of equipoise is widely accepted as a necessary condition for the initiation and continuation of a clinical trial is that it is thought to be ethically permissible to allow a subject's care to be determined by a random process when equipoise exists because there is genuine disagreement among informed medical experts about whether one treatment option dominates the others. As an ethical requirement, therefore, the principle of equipoise seeks to operationalize the epistemic component of what is sometimes called the researcher's 'therapeutic obligation'[8] or 'duty of personal care.'[6] This is because the content of that obligation—what it is the duty of a physician or a researcher to provide in a given instance—depends in part on what is known about the likely effects of the available treatment options.

It is a well-known statistical fact, however, that an intervention may prove to be beneficial for a population in the aggregate, but fail to be so for identifiable subgroups of that population. If equipoise is going to operationalize the epistemic component of the duty of personal care then we believe that it needs to be applied to patients insofar as they are characterized by variables that are determined by a clinically adequate description of the characteristics that are deemed to be medically relevant by the appropriate experts. If, at the inception of a trial, there is general agreement among the relevant decision makers that equipoise is disturbed for subjects with specified characteristics, then it would violate the equipoise requirement to assign such patients the disfavored treatment, even if equipoise exists for the larger trial population. Additionally, as information is gathered from an ongoing trial there may come a point at which equipoise is disturbed for subjects with some identifiable characteristics

but not for others. As a general point, the question of how finely it is possible to characterize the medically relevant characteristics of subjects is conceptually distinct from the question of when it is appropriate to review the data pertaining to them. However, we also believe that these issues have broader implications for the ethics clinical trial design. For instance, a sequentially dynamic trial can maintain a state of equipoise at the level we are recommending by automatically assigning to patients the treatment deemed by the relevant decision makers not to be dominated by the alternatives.[15]

For the purposes of the present discussion, we simply want to note that the above issues are open questions that would have to be addressed by any precise articulation of the principle of equipoise. Even without a precise clarification of these technical issues, however, we believe that the core of the equipoise requirement elaborated above is sufficient to state in broad terms an objection that may account for the lingering if ambiguous resistance that many experience concerning sham surgical controls.

As mentioned above, one goal of the equipoise requirement is to ensure that the welfare and integrity of research participants are not knowingly sacrificed for the interests of future patients or for the advancement of scientific progress. Another goal of the requirement is to ensure that research addresses an important health question in a way that will yield reliable, generalizable information. Trials that begin in and that are designed to disturb equipoise will provide information that the medical community can use to improve its current practice and advance the quality of care. Applying the intuitive core of the principle of equipoise to the use of sham surgical controls yields a strong, though rebuttable presumption that any protocol involving a sham surgery control violates condition (2) of the equipoise requirement. In particular, the sham surgery component of the placebo arm will be dominated by a more benign alternative that substitutes a less burdensome placebo for the surgical component of the control.

For example, if there already exists an intervention of proven effectiveness for some condition, and if there is genuine disagreement among informed medical experts as to whether the study intervention is equally or more effective, then the study intervention must be compared against the established alternative.[16] Adding a sham surgery control to the established intervention adds significant risk of foreseeable and easily preventable harm to the control group without the promise of a corresponding benefit. Similarly, if no such established intervention exists, and there is genuine disagreement among informed medical experts as to whether the study intervention is equally or more effective than a *benign* placebo, then the study intervention may be compared against a benign placebo. Using a sham surgery component in the control again adds risks of foreseeable and preventable harm without a corresponding benefit to subjects in the control arm. As a result, it is difficult to see how the use of sham surgery controls might be reconciled with the duty of personal care.

While the concept of equipoise has not been invoked explicitly in the debate about the use of sham surgery in trials of Parkinson's disease, it has not been altogether absent either. For example, Freeman *et al.*, note that a condition for ethically acceptable clinical research is that there 'must be preliminary but not conclusive evidence that the intervention is effective' (p. 989) and they cite Banjamin Freedman's article 'Equipoise and the Ethics of Clinical Research' in support of this requirement.[16] Perhaps Freeman *et al.* can charitably be read as embracing the first condition of the principle

of equipoise (1) stated above and then claiming that there existed a state of genuine uncertainty regarding the comparative merits of fetal negral cell transplantation (plus standard medical therapy) versus sham surgery (plus standard medical therapy).

Unfortunately, however, uncertainty of this sort is not sufficient to justify the use of a placebo surgery component in the control group. Accepting equipoise condition (1) without (2) severs the connection between the principle of equipoise and the duty of personal care, opening the door to trials in which some or all subjects are knowingly randomized to care that is dominated by an available alternative. Accepting both (1) and (2), however, rules out the use of placebo surgery in this case, because the intervention used in the control arm is dominated by the more benign option of a control that provides standard medical therapy without the sham surgery component.

In light of the above analysis, several things seem clear. First, there are strong *prima facie* objections to sham surgery in general and in recent Parkinson's research in particular. Second, while there may be circumstances in which these *prima facie* objections may not obtain, there are good reasons to think that these conditions presuppose the results of clinical trials that do not use sham controls. Third, the above analysis should clarify the sense in which clinical trials run without sham surgery controls can generate important and scientifically valuable results. In part, such trials address questions that bear directly on clinical practice while raising fewer ethical objections. Additionally, their results will help to determine whether subsequent studies may permissibly use sham surgical controls.

For these reasons, like Macklin, we challenge the ethical basis of sham surgical controls in trials of fetal-tissue transplantation for the treatment of Parkinson's disease. Although Macklin argues primarily on the basis of her interpretation of the US Code of Federal Regulations, however, our arguments draw on the basic ethical and methodological underpinnings of equipoise. The fact that we draw such similar conclusions from largely independent starting points lends further credence to our reservations about the ethical status of sham surgery controls in this context.

5 Thinking more deeply about placebo effects

It might be argued that our claims in the previous section hinge on the assumption that the risks associated with sham surgical controls are not compensated for or outweighed by corresponding therapeutic effects. It will be helpful, therefore, to respond to an idea that seems to complicate such an analysis. The basic idea is this: Although there are risks associated with the sham surgical procedure, these risks might be compensated for by placebo effects that result from the placebo surgery. We are dubious about this for several reasons.

First, a general cautionary note. In evaluating the use of any placebo control, one must be careful not to confuse two uses of the term 'placebo effect.' In controlling for placebo effects in a clinical trial, the term is being used in what we will call its broad, 'operational' sense. The function of a placebo control in a properly designed clinical trial is to assist in blinding and to hold constant as many variables as possible, other than the fact of whether a subject receives the control substance or the new intervention. Ideally, this allows one to isolate the specific effects of an investigational agent from the

array of background influences that may affect the condition of subjects in both the experimental and control arms. In the operational use of the term, 'placebo effects' are changes in measured variables that may result *from any of a number of sources*, ranging from natural variations in disease cycle, regression to the mean, or what is sometimes called the Hawthorne effect or the caring effect, among others. The primary importance of the operational placebo effect is epistemic. That is, one might say that the effect of using a placebo as a control in a well-designed clinical trial is to hold this array of background influences constant so that their impact can be separated from the specific effects of the investigational intervention.

In contrast, what we will call the 'therapeutic' use of the term refers to an improvement in a person's condition that is causally dependent on the patient's perception of having received something beneficial. While various physiological pathways may be instrumental in realizing whatever benefit a patient receives from the placebo, the perception of having received something beneficial is itself causally sufficient to activate these underlying mechanisms. As a result, to postulate a therapeutic placebo effect is to postulate a particular causal story, one in which certain physiological mechanisms—about which we may know very little—would not have been activated without the subject's perception that he or she had received something that would benefit her. The primary importance of the therapeutic placebo effect is not epistemic but clinical. That is, one might say that on this use of the term, the effect of administering a placebo is to provide the patient with a benefit that he or she would not have otherwise received.

In light of the above distinction we want to stress the following point. If the condition of members of the control group appears to improve over the course of a double blind, randomized, placebo-controlled trial, we can say with certainty that what has been measured is an operational placebo effect. From such evidence alone, however, it cannot be inferred that this operational placebo effect has been caused by a therapeutic placebo effect. This is because such trials cannot themselves isolate what we are calling a therapeutic placebo effect from the background array of possible influences on the measured variables.[17,18]

In itself, this point provides strong reason to be circumspect about the way we think about 'the' placebo effect. It takes on special significance, however, in light of a recent meta-analysis that casts doubt on the therapeutic placebo effect as a general phenomenon.[18] The authors of this study looked at clinical trials in which patients were randomized to both a placebo and a no-treatment arm. They found 'little evidence that placebos in general have powerful clinical effects' (p. 1599) and they report that the study was unable to detect any significant difference between placebo and untreated groups as a result of bias associated with a lack of double blinding (p. 1599). They are clear that the study could not estimate the overall influence of bias due to lack of double blinding, but they note that this has been estimated to be exaggerated by 17% (p. 1599). With respect to this last point it is worth mentioning that some recent studies have found little evidence that estimates of treatment effects in observational studies reported after 1984 are either consistently larger than or qualitatively different from those obtained in randomized, controlled trials.[19,20]

For our present purposes, these studies simply provide further reason to shift the burden of proof to those who would postulate a therapeutic placebo effect as a potential

benefit that might compensate for the risks incurred in receiving sham surgery. We are not claiming that they constitute proof that there is no therapeutic placebo effect or that subjective bias can be effectively controlled in observational studies. In fact, we have very recently become aware of two studies that attempt to measure placebo-associated improvements in patients with Parkinson's disease.[21,22] In terms of the distinction articulated above, these studies are designed to determine whether an operational placebo effect has been caused by a therapeutic placebo effect. Since so much clinical trial practice is built on a belief in the power of a therapeutic placebo effect, it will be interesting to see the extent to which the findings of the above studies can be replicated and whether they come to be widely accepted. At a very general level, however, they should remind us that just as researchers have to question dogmas of clinical practice that may need to be changed, we have to question the dogmas of clinical trial design in light of ongoing empirical investigation and advances in statistical methods.

This brings us to a second general point about placebo effects: We should not simply presume that they will be beneficial. When placebos are associated with adverse events the resulting effect is sometimes referred to as a 'nocebo effect.'[23] If, in fact, the patient's perception of his or her treatment or care can influence his or her health status, it is reasonable to think that the influence can be either beneficial, giving rise to a therapeutic placebo effect, or harmful, giving rise to a nocebo effect. Although much less attention has been paid to negative influences of placebos, there have been reports that placebos can worsen pre-existing conditions and produce pain in normal subjects.[17] Presumably, those who might countenance a therapeutic placebo effect as a possible benefit of receiving sham surgery should also countenance the possibility of nocebo effects as well. Just as a nocebo effect could result from a negative experience within a clinical trial, such an effect could be delayed, affecting patients when they learn that they received the sham rather than the real procedure. If the burdens of a placebo control are going to be justified on the basis of potential therapeutic placebo effects then surely this dual potential should be considered when evaluating the potential risks and benefits associated with sham surgical procedures.

The upshot of the previous two points is to place the burden of proof squarely on those who claim that the risks of sham surgery are outweighed by the possible benefits of a therapeutic placebo effect. If this is not where the burden of proof is located then hope of potential therapeutic placebo effects could be used to support even the most dangerous or ridiculous of controls. Responsible research, however, must be based on more than vague possibilities.

We therefore want to emphasize that before we contemplate subjecting patients to the burdens of sham surgery controls, we ought to first conduct trials that establish the merits of fetal negral cell transplants as they would be executed in clinical practice relative to the more benign alternative of standard medical care only. Even if it is true that those who receive fetal-tissue transplants will also benefit from a therapeutic placebo effect, it does not follow that fetal-tissue transplants will therefore dominate a standard therapy control in a pragmatic clinical trial. After all, such a trial seeks to compare the net therapeutic advantages of the alternative treatments and it may well be that such a therapeutic placebo effect is outweighed by the burdens associated with the biological effects of the implants, the burdens of the surgical procedures, or some other set of factors.

As a result, substantiating the existence of a genuine therapeutic placebo effect in studies of Parkinson's disease would not, in itself, constitute an objection to the position we are defending. That is because our position is that there are important ethical reasons that support giving *prima facie* priority to studies that answer pragmatic questions before we use human subjects to answer more analytical, explanatory questions. To impugn the value of a pragmatic trial for not answering explanatory questions is simply to assert that the explanatory questions are the ones we ought to be asking. In the present case, we have argued that this is not the case and that the pragmatic question should take priority.

6 Active placebos in drug trials

A discussion of the use of various controls in surgery trials would be incomplete if it did not acknowledge the background of the widespread use of placebos in pharmaceutical research. In studies of treatment for depression, some researchers have found that 75% of the apparent effect of the drugs is also present in the placebo, and that 80% of the patients in some trials guessed correctly to which group they had been assigned.[24] It should be noted that the severity of depression is standardly measured with Hamilton Scores, a self-report of a patient's functioning. Patients can probably tell which group they belong to by the presence or absence of side effects of the drugs, such as dry mouth, nausea, dizziness, or sexual dysfunction. This has led to suggestions to use an 'active placebo', which would mimic the side effects of the drug but not its main effect on seratonin and/or norepinephrine in the brain. Presumably such an active placebo would better protect the blinding in such a trial.

Sham surgery as a control for fetal-tissue transplantation strikes us as being analogous to active placebos in drug trials, both in its goals and the ethical considerations that apply to it. In both cases, patients in the control arm are being given a treatment that disadvantages them in order to exclude certain causal hypotheses in explaining results of a trial. To the extant that evaluation of treatments are based on subjective quality of life measures, it is important in both cases that the information about the treatment arms be presented in a fair and balanced way, emphasizing the uncertainty faced by both patient and researcher. Not only is this required by duty to secure adequately informed consent, but it also helps to ensure more realistic patient expectations.

Yet the stakes and the situations do not strike us as completely analogous, for the following reasons. First, the disadvantage to patients assigned to sham surgery strikes us as more damaging to patients than the side effects of active pharmaceutical placebos. This consideration tends to put a greater ethical burden on sham surgery for transplantation than on active placebos for depression. However, as the significance of the burdens imposed on subjects by active pharmacological placebos increases, so do the *prima facie* reasons not to use them. Second, in cases where antidepressants have demonstrated an attractive net therapeutic index in pragmatic trials, it may be possible to justify the use of active placebos in further explanatory trials if there remains uncertainty in the expert community about the mechanisms responsible for the effects measured in the pragmatic trials.

7 Conclusion

Recent debates over the use of sham surgery as a control for studies of fetal-tissue transplantation for Parkinson's disease have focused primarily on rival interpretations of the US federal regulations governing human-subjects research. Using the core ethical and methodological considerations that underwrite the equipoise requirement, we find strong *prima facie* reasons against using sham surgery as a control in studies of cellular-based therapies for Parkinson's disease and in clinical research more broadly. Additionally, we believe that these reasons can be generalized to apply to the use of other placebo controls that carry significant risks in and of themselves of positive harms.

Although sham surgery and so-called active placebo controls in pharmaceutical research raise similar ethical issues in this regard, practical considerations lead us to the conclusion that sham surgery controls in trials for Parkinson's disease are more dubious than active placebos for depression. Nevertheless, those who wish to use either must meet the burden of proof. We suggest that they justify their choice of control within the framework of equipoise.

References

1 Freeman TB, Vawter DE, Leaverton PE *et al.* Use of placebo surgery in controlled trials of a cellular-based therapy for Parkinson's disease. *The New England Journal of Medicine* 1999; **341**: 988–92.

2 Macklin R. The ethical problems with sham surgery in clinical research. *The New England Journal of Medicine* 1999; **341**: 992–96.

3 Freed CR, Green PE, Breeze RE *et al.* Transplantation of embryonic dopamine neurons for severe Parkinson's disease. *The New England Journal of Medicine* 2001; **344**: 710–19.

4 Pollack A. Companies announce setback in treatment for Parkinson's. *New York Times* 2001; May **19**: C2.

5 Johannes L. Sham surgery is used to test effectiveness of novel operations. *Wall Street Journal* 1998; **Dec 11**: A1, A8; Moran M. Placebo surgeries test new procedures. *American Medical News* 1999; **42**: 1, 34–35.

6 Fried C. *Medical experimentation: Personal integrity and social policy.* Amsterdam: North-Holland Publishing, 1974.

7 Freedman B. Placebo-controlled trials and the logic of clinical purpose. *IRB* 1990; **12**: 1–6.

8 Marquis D. Leaving therapy to chance. *Hastings Center Report* 1983; **13**: 40–47.

9 56 Federal Register 28012, 45 Code of Federal Regulations 46.

10 Weijer C. The ethical analysis of risk. *Journal of Law, Medicine & Ethics* 2000; **28**: 344–61.

11 London AJ. The ambiguity and the exigency: Clarifying 'standard of care' arguments in international research. *Journal of Medicine and Philosophy* 2000; **25**: 379–97.

12 Rothman KJ, Michels KB. The continuing unethical use of placebo controls. *The New England Journal of Medicine* 1994; **331**: 394–98.

13 Schwartz D, Flamant R, Lellouch J. *Clinical trials* [translated by Healy MJR]. London: Academic Press, 1980.

14 London AJ. Equipoise and international human-subjects research. *Bioethics* 2001; **15**: 312–32.

15 Kadane, JB (Editor) *Bayesian methods and ethics in a clinical trial design.* New York: John Wiley & Sons, Inc., 1996.

16 Freedman B. Equipoise and the ethics of clinical research. *The New England Journal of Medicine* 1987; **317**: 141–45.

17 Turner JA, Deyo RA, Loeser JD, Von Korff M, Fordyce WE. The importance of placebo effects in pain treatment and research. *JAMA* 1994; **217**: 1609–14.

18 Hrobjartsson A, Gotzsche PC. Is the placebo powerless. *The New England Journal of Medicine* 2001; **344**: 1594–602.

19 Benson K, Hartz AJ. A comparison of observational studies and randomized

controlled trials. *The New England Journal of Medicine* 2000; **342**: 1878–86.

20 Ioannidis JPA, Haidich AB, Pappa M *et al.* Comparison of evidence of treatment effects in randomized and nonrandomized studies. *JAMA* 2001; **286**: 821–30.

21 Goetz CG, Leurgans S, Raman R, Stebbins GT. Expectation and dopamine release: mechanism of the placebo effect in Parkinson's disease. *Neurology* 2000; **54**: 710–14.

22 de la Fuente-Fernández R, Ruth TJ, Sossi V, Schulzer M, Calne DB, Stoessl AJ. Expectation and dopamine release: mechanism of the placebo effect in Parkinson's disease. *Science* 2001; **293**: 1164–66.

23 Voelker R. Nocebos contribute to host of ills. *JAMA* 1996; **275**: 345–47.

24 Enserink M. Can the placebo be the cure? *Science* 1999; **284**: 238–40.