RESEARCH ETHICS

Sham surgery controls: intracerebral grafting of fetal tissue for Parkinson's disease and proposed criteria for use of sham surgery controls

R L Albin

J Med Ethics 2002;**28**:322–325

Correspondence to:
Dr R L Albin, Department of Neurology, University of Michigan, 4412D Kresge III, 200 Zina Pitcher Place, Ann Arbor, MI, 48109–0585, USA;
ralbin@umich.edu

Revised version received
28 February 2002
Accepted for publication
4 March 2002

Sham surgery is a controversial and rarely used component of randomised clinical trials evaluating surgical interventions. The recent use of sham surgery in trials evaluating efficacy of intracerebral fetal tissue grafts in Parkinson's disease has highlighted the ethical concerns associated with sham surgery controls. Macklin, and Dekkers and Boer argue vigorously against use of sham surgery controls. Macklin presents a broad argument against sham surgery controls while Dekkers and Boer present a narrower argument that sham surgery is unnecessary in the specific setting of fetal tissue engraftment for Parkinson's disease. I defend sham surgery controls against both these criticisms. Appropriate clinical trial design, sometimes including sham surgery, is needed to ensure that false positive trial results do not occur and endanger public safety. Results of a completed trial of fetal tissue grafting for Parkinson's disease are used to illustrate the potential benefits of, and problems associated with, sham surgery controls. Sham surgery controls, however, should be employed only when absolutely necessary. I suggest criteria for appropriate use of sham surgery controls.

Sham surgery (surgical placebo) in clinical trials is a rare event.[1–4] It is easy to understand the infrequent use of sham surgery controls and the discomfort aroused by their employment. A basic principle of clinical trials is that the risk-benefit ratio must be favourable.[5] In typical placebo controlled medical trials, the treatment group is exposed to the risks of the novel treatment but this is compensated for by the chance of receiving the hoped for benefits of the novel treatment. The placebo group forego the chance of benefits from the novel treatment but incurs no additional risk. With sham surgery, the control has no chance of gaining additional benefit but is exposed to additional risks. Dekkers and Boer provide a nice analysis of the ways in which sham surgery controls challenge ethical norms.[6]

Sham surgery, none the less, has been performed as part of small trials evaluating internal mammary artery ligation for relief of angina and arthoscopy for treatment of osteoarthritis.[1 2] More recently, sham surgery has been incorporated into two substantial trials evaluating the efficacy of intracerebral fetal tissue tissue grafting in Parkinson's disease (PD).[3 4] These studies were preceded by a considerable body of animal experiments and unblinded procedures in humans suggesting substantial benefit from intracerebral grafts of fetal tissue in PD. The use of sham surgery in these trials generated considerable debate, which was summarised well in a pair of articles in the *New England Journal of Medicine*.[3 7] In one of these articles, the bioethicist Ruth Macklin delivered a stringent critique of the concept of sham surgery controls.[7] Dekkers and Boer subsequently published a thoughtful analysis of Macklin's arguments and provided additional arguments against sham surgery controls. To contribute to this debate, I will defend sham surgery controls in the context of studies evaluating fetal tissue engraftment for PD. My analysis has two primary goals. First, to demonstrate that sham surgery controls are sometimes methodologically necessary and ethically justifiable, and second, to suggest general criteria for the employment of sham surgery controls.

ETHICAL ARGUMENTS AGAINST AND FOR SHAM SURGERY CONTROLS

Macklin identified three primary problems with sham surgery controls. She argued that the concept of sham surgery controls produced "tension between the highest standard of research design and the highest standard of ethics", that there were intractable problems with assessing risks and benefits in this situation, and that the informed consent doctrine could not be used as a blanket assurance justifying sham surgery controls. The first of Macklin's charges is probably the most serious but Macklin's conclusion is based on a narrow view of research ethics. As pointed out by Dekkers and Boer, Macklin seemingly concentrates solely on the relationship between researchers and subjects/patients.[6] Dekkers and Boer point out that research ethics "takes into account not only the interests of research subjects, but also the interests of biomedical science, of the category of patients to which the research subjects belong, and of society at large".[6] Macklin ignores the most important justification of sham surgery controls: the need for rigorous studies that will exclude false positive results. As pointed out by Freeman *et al* in their defence of sham surgery controls, it is common for surgical techniques to be introduced into clinical practice without rigorous evaluation.[3] The result can be exposure of substantial numbers of patients to procedures that incur significant risks and have no benefit. In addition to becoming a public health hazard, inadequately evaluated surgical methods can consume valuable societal resources. As stated by Emanuel *et al*: "valuable research must be conducted in a methodologically rigorous manner".[5] The problem is not tension between the highest standard of research design and the highest standard of ethics, the problem is tension between obligations to individual research subjects/patients and obligations to the larger group of patients and the general public.

This is not a theoretical concern. There are abundant examples of widely adopted surgeries that were abandoned subsequently for lack of efficacy. A good example in the area of neurology is carotid endarectomy (CEA). This procedure has been used for decades as primary and secondary prophylaxis for stroke. Scientific evaluation of CEA was performed many

years after its incorporation into clinical practice. While these trials confirmed benefits of CEA, the magnitude of the benefits and the number of eligible patients proved to be less than predicted by CEA advocates. Since CEA had been performed on millions of patients prior to proper evaluation, it is likely that hundreds of thousands of patients were exposed to unnecessary risks and that substantial resources were wasted. Even small trials incorporating sham surgery controls can have a major impact. As Beecher pointed out in his comments on sham surgery controlled trials of internal mammary ligation, a pair of small but well constructed studies involving 35 subjects probably spared thousands the risks of unnecessary surgery.[8]

Macklin's other criticisms have greater force. There are inherent difficulties in assessing risk/benefit ratios for sham surgical controls. There are both empirical and theoretical reasons to avoid reliance on the doctrine of informed consent as a convenient escape from the ethical dilemmas raised by sham surgical controls. There are documented limitations of the informed consent process. For example, a high percentage of adult patients enrolled in cancer clinical trials at three prestigious American institutions exhibited poor understanding of the nature of trials.[9] Informed consent may be distorted by the so-called "therapeutic misconception" in which sometimes desperate subjects have unrealistic expectations about benefits of trial participation. Primary reliance on informed consent to justify use of placebos may place undue emphasis on the underlying principle of autonomy.[5]

Where Macklin issued a blanket critique of sham surgery controls, Dekkers and Boer published a thoughtful critique of sham surgery controls in the specific context of intracerebral grafting of fetal tissue for Parkinson's disease.[6] Acknowledging the power of much of Macklin's general analysis, Dekkers and Boer condemned sham surgery controls in this context not only as undesirable but also as unnecessary. Their suggested alternative is a study in which subjects are evaluated in a rigorous manner both before and after surgery in parallel with a matched group of controls. They suggest that, with an appropriate set of standard measures and sufficient follow up, decisive information could be accumulated. While not mentioned by Dekkers and Boer, this type of design could incorporate blinded evaluations. This is a rational proposal, but would it be sufficiently rigorous to exclude false positive effects? In particular, does the parallel design proposed by Dekkers and Boer deal adequately with possible placebo effects of surgery?

PLACEBO EFFECTS IN PARKINSON'S DISEASE

There is considerable evidence of placebo effects in clinical trials of PD. Shetty *et al* reviewed a large number of medical clinical trials for PD and found evidence of placebo effects in a large proportion of the studied trials.[10] In a careful analysis of placebo effects during a well designed drug trial for PD, Goetz *et al* documented significant and persistent placebo effects over a six month period with no evidence of a transient early effect.[11] Placebo effects were evident with the most objective component of the evaluation, the motor subscale of the Unified Parkinson's Disease Rating Scale (UPDRS). Unlike most placebo effects, there may be a physiological explanation for placebo effects in PD. De la Fuente-Fernandez *et al* studied the effect of placebo administration on dopamine release in PD subjects.[12] In a remarkable result, placebo administration evoked substantial dopamine release. The effect on dopamine receptor occupancy was similar to parenteral administration of the dopamine agonist apomorphine or oral levodopa. One of the functions of the nigrostriatal dopamine projection may be to signal expectation of reward—for example, expectation of therapeutic benefit. This result is consistent with our present understanding of nigrostriatal neuron function. De la Fuente-Fernandez *et al*'s results are not, however, a general explanation for placebo effects but are probably specific to nigrostriatal dopaminergic function.

Hrobjartsson and Gotzsche argued recently that the magnitude of the placebo effect is exaggerated.[13] In their analysis of many trials they found little evidence of a placebo effect in studies with objective or subjective binary outcomes. They did, however, find evidence of significant placebo effects in trials involving continuous subjective outcomes and in studies of pain interventions. Since PD trials focus on clinical effect, subjective measures are inescapable and will be susceptible to placebo effects. Double blind placebo designs are used also for two purposes: as a control for the placebo effects proper, and to eliminate other forms of bias. These two purposes can be related. For example, it is possible to blind investigators without placebo controls but it is still possible for subjects to communicate their status as controls or experimental subjects, even unconsciously, to investigators. Hrobjartsson and Gotzsche stress also the importance of placebo controlled clinical trials in avoiding several forms of bias and not just the placebo effect.

Placebos are important components of clinical trials in PD. There is a real placebo effect in PD, PD trials often involve subjective outcome measures likely to be affected by placebo effects, and placebos are important in avoiding other sources of bias.

SHAM SURGERY CONTROLLED TRIALS OF INTRACEREBRAL GRAFTING OF FETAL TISSUE IN PD: INITIAL RESULTS

Two trials of sham surgery controlled intracerebral grafting of fetal tissue in PD have been undertaken in the USA. One of these trials has been completed.[4] This is not the trial discussed specifically by Macklin and Freeman *et al* and has some different design features. In this study, 40 subjects were randomised to graft or no graft arms. The sham surgery subjects had an outer skull table burr hole and sham use of injection cannulae. General anaesthesia and immunosuppressive agents were not used in either arm. The physicians performing the engraftment were entirely separate from the physicians caring for and evaluating the subjects. After one year of double blind follow up, all subjects had the option of receiving grafts. Subjects were followed with a subjective self reported global rating scale as the primary outcome measure, the Schwab and England activities of daily living scale, and the UPDRS. Successful engraftment was confirmed by [^{18}F]Fluorodopa positron emission tomography. At one year, only younger subjects (<60 at time of engraftment) had significant benefits from grafting. With longer follow up, the magnitude of the benefit increased but was not substantial in older subjects and some of the subjects with a good response to engraftment developed disabling dyskinesias. There was also a significant placebo effect of sham surgery in the first months after engraftment.[14] At the end of the one year double blind period, it was impossible to demonstrate an improvement in the global rating scale because of the magnitude of the placebo response. There was also a significant placebo response in the UPDRS scores. Serious adverse events related to trial procedures did not occur in the sham surgery arm. Results of the second trial are not yet available.

The documentation of a significant and persistent placebo effect due to sham surgery raises real concerns. This trial was constructed thoughtfully and executed carefully. It is easy to imagine the consequences of an uncontrolled and less carefully run study. The early and apparently positive results could result in premature termination of the study and premature endorsement of intracerebral grafting of fetal tissue for PD. Overenthusiastic interpretation of trial data is a familiar phenomenon. For example, the large Deprenyl and Tocopherol Antioxidative Therapy in Parkinsonism (DATATOP) trial, designed by a highly competent group of investigators, was terminated early because of an apparently

favourable outcome and then had to be restarted because of conflicting data from another trial.[15] In DATATOP, the final interpretation of study data was different from the initial enthusiastic response.[15] In the published grafting trial, the actual benefits of grafting were shown to be limited and the adverse consequences of grafting significant, so the potential for harm to PD patients from a poorly controlled and prematurely terminated trial was substantial.

SHAM SURGERY CONTROLS FOR INTRACEREBRAL GRAFTING OF FETAL TISSUE IN PD: JUSTIFIED OR NOT?

In the case of intracerebral grafting of fetal tissue for PD, a good case can be made for the scientific necessity of a sham surgery control arm. Both medical trials and the experience of the first completed sham surgery controlled trial substantiate the existence of important placebo effects. Parkinson's disease studies often involve subjective ratings, which are apparently particularly prone to placebo effects. Finally, placebo controls are an important element of avoiding other forms of bias. Dekkers and Boer proposed an alternative parallel trial design to sidestep the problems inherent in sham surgery controls. This proposal, however, has significant flaws. It does not directly address the problem of the placebo effect, continues to rely on subjective measures, and by eliminating the placebo arm, runs the risk of introducing other forms of bias. Use of this design could be justified only if it were proven to be equivalent to a sham surgery controlled trial, which then requires a sham surgery controlled trial for comparison. Since the goal of the parallel design study is to avoid sham surgery controlled trials, this proposal is self defeating.

The need to avoid dissemination of unsafe treatments mandates critical evaluation of proposed new medical or surgical therapies. Novel therapies must be scrutinised with the highest possible level of scientific rigour. What is possible, however, is limited by other ethical considerations. Use of sham surgery is unattractive because the increased risk to control subjects is not accompanied by any possibility of benefit. In some cases, however, sham surgery controls are strongly preferred on scientific grounds and may be necessary to answer the key questions. Sham surgery controls cannot be prohibited absolutely but their use must be balanced carefully against the safety of research subjects.

CRITERIA FOR USE OF SHAM SURGERY CONTROLS

Because of the necessity of minimising risk for research subjects, sham surgery controls should not be the default method of constructing human clinical trials involving surgical interventions. Sham surgery controls should be used only with careful justification and I believe that these circumstances will be rare. When is sham surgery justified? I propose formal criteria as a decision aid.

First, all the general standards for ethical conduct of clinical research must be satisfied. Emanuel *et al* describe seven ethical requirements that must be met by all clinical research.[5] These comprise value, scientific validity, fair subject selection, favourable risk-benefit ratio, independent review, informed consent, and respect for potential and enrolled subjects.

Second, there cannot be reasonable alternative research designs. There must be a legitimate and substantiated concern about placebo effects or other forms of potential bias that cannot be defused by use of an alternative design. This concern must be well documented. For example, information from medical clinical trials or other studies demonstrating significant potential for placebo effects must be available. Other forms of potential bias that would require a sham surgery control would have to be identified in the same way. Trial designs that obviate potential for bias, even if more expensive or time consuming, are mandated. For example, in situations where a placebo effect was a significant concern, use of subjective binary designs or objective outcome measures without sham surgery controls would be satisfactory. To require sham surgery controls, there would have to be an absence of useful objective measures addressing the study outcomes. The following hypothetical example illustrates the way in which an objective measure could eliminate the need for sham surgery controls. Suppose there was a proposed trial of a gene therapy in chronic progressive multiple sclerosis. Suppose also that this therapy required intraventricular administration of the vector. A rational approach to avoid exposing controls to risks of vector administration would be a parallel design with a matched group of conventionally treated patients and magnetic resonance imaging (MRI) measurements of white matter lesion changes as the primary outcome variable. Any objective measure, however, will have to be well validated in terms of clinical significance. The ultimate measure of interventions is beneficial clinical outcomes, and objective measures that correlate poorly with clinical outcomes or have little biological significance would be of little use. Again, data from medical trials or observational studies validating objective measures would be crucial. In some cases, such as studies of pain interventions, objective measures may be impossible.

Third, there must be a procedure for minimising the risk-benefit ratio. It is difficult to delineate definitive standards for accomplishing this end. A useful procedure might be to examine the proposed intervention, identify the major risks, and to attempt to construct a sham procedure that eliminates these risks. In the recently completed trial of intracerebral fetal tissue grafting for PD, the investigators made a serious effort in this direction. The major risks associated with grafting are the danger of infection resulting from introduction of foreign material, the risk of intracerebral haemorrhage from introduction of the injection cannulae, the risk of general anesthesia, and the risk from the use of potentially toxic antirejection drugs. The sham procedure was designed to avoid these risks. No tissue was injected or cannulae inserted. The dura was not breached. General anaesthesia and antirejection drugs were not used in either arm. Similarly, in a small trial of arthroscopy for osteoarthritis pain, the sham surgery group did not receive endotracheal general anaesthesia, superfical cuts were made in the skin but no arthroscope was inserted.[2] The operative team then carried out a dummy procedure. If no sham procedure avoiding the major risks of the intervention can be established, then sham surgery controls should be abandoned.

Fourth, the minimum necessary number of subjects should be enrolled. This places considerable emphasis on careful biostatistical formulation of the trial with reliable estimates of power and sample size needs. This may require significant preliminary data accumulation from open studies, medical trials, and studies evaluating the measuring instruments.

Fifth, there should be an exceptionally vigilant, independent safety monitoring board. Safety monitoring boards with predefined stopping rules are an important component of clinical trials, though small trials frequently do not use them. All trials involving sham surgery controls should have an independent safety monitoring board with well defined stopping rules and unusually frequent oversight.

CONCLUSIONS

Ethical clinical research requires the most rigorous scientific methods possible. Sham surgery controls are likely to be highly desirable in some situations and may be necessary safeguards against false positive trial outcomes. The recent experience with intracerebral fetal tissue grafting for Parkinson's disease suggests that sham surgery controls can be done in a safe and ethical manner. Criteria are proposed for the use of sham surgery controls.

FUNDING

Supported by a Merit Review Grant, and AG08671 from the National Institutes of Health.

REFERENCES

1 **Cobb LA**, Thomas GI, Dillard DH, *et al*. An evaluation of internal-mammary-artery ligation by a double blind technic. *New England Journal of Medicine* 1959;**260**:1115–18.
2 **Moseley JB**, Wray NP, Kuykendall D, *et al*. Arthroscopic treatment of osteoarthritis of the knee: a prospective randomized placebo-controlled trial. Results of a pilot study. *American Journal of Sports Medicine* 1996;**24**:28–34.
3 **Freeman TB**, Vawter DE, Leaverton PE, *et al*. Use of placebo surgery in controlled trials of a cellular-based therapy for Parkinson's disease. *New England Journal of Medicine* 1999;**341**:988–91.
4 **Freed CR**, Greene PE, Breeze RE, *et al*. Transplantation of embryonic dopamine neurons for severe Parkinson's disease. *New England Journal of Medicine* 2001;**344**:710–19.
5 **Emanuel EJ**, Wendler D, Grady C. What makes clinical research ethical? *Journal of the American Medical Association* 2000;**283**: 2701–11.
6 **Dekkers W**, Boer G. Sham neurosurgery in patients with Parkinson's disease: it is morally acceptable? *Journal of Medical Ethics* 2001;**27**:151–6.
7 **Macklin R**. The ethical problems with sham surgery in clinical research. *New England Journal of Medicine* 1999;**341**:992–6.
8 **Beecher HK**. Surgery as placebo. A quantitative study of bias. *Journal of the American Medical Association* 1961;**176**:1102–7.
9 **Joffe S**, Cook EF, Cleary PD. Quality of informed consent in cancer clinical trials: a cross-sectional study. *British Medical Journal* 2001;**358**:1772–7.
10 **Shetty N**, Friedman JH, Kieburtz K, *et al*. The placebo response in Parkinson's disease. *Clinical Neuropharmacology* 1999;**22**:207–12.
11 **Goetz CG**, Leurgans S, Raman R, *et al*. Objective changes in motor function during placebo treatment in PD. *Neurology* 2000;**54**: 710–14.
12 **De la Fuente-Fernandez R**, Ruth TJ, Sossi V, *et al*. Expectation and dopamine release: mechanism of the placebo effect in Parkinson's disease. *Science* 2001;**293**:5532–3.
13 **Hrobjartsson A**, Gotzsche PC. Is the placebo powerless? An analysis of clinical trials comparing placebo with no treatment. *New England Journal of Medicine* 2001;**344**:1594–602.
14 **Freed CR**, Breeze RE, Greene PE, *et al*. Double-blind placebo-controlled human fetal dopamine cell transplants in advanced Parkinson's disease. *Society for Neuroscience Abstracts* 1999:**25**:212.
15 **Shoulson I**, for Parkinson Study Group. DATATOP: a decade of neuroprotective inquiry. *Annals of Neurology* 1998;**44**(suppl 1):160–6S.