# Review

# Evidence-Based Evaluation of the Ethics of Sham Surgery for Parkinson's Disease

Stephen Polgar[a,*] and Sheeza Mohamed[b]
[a]*School of Public Health and Psychology, La Trobe University, Bundoora, Melbourne, Australia*
[b]*School of Life Sciences, La Trobe University, Bundoora, Melbourne, Australia*

**Abstract**. The stated purpose of sham or placebo surgery is to enable the implementation of surgical placebo-controlled trials (SPTs) for evaluating the safety and efficacy of surgical interventions. Exposing the participants to the burdens and harms of sham surgery has been justified on the grounds of the absolute necessity for controlling large placebo effects and observer bias, assumed to be associated with surgical procedures. In the present review, we argue that evidence obtained from SPTs of cellular therapies for the treatment of Parkinson's disease (PD) has failed to demonstrate either large and consistent placebo effects or decisive methodological advantages for relying on sham surgical controls. We outline several alternative assessment strategies and designs available to establish the efficacy of cellular therapies. It is concluded that the evidence evaluated in the present analysis indicated that use of sham surgery in the context of developing novel surgical procedures for PD is not necessary, and therefore, unethical under a utilitarian model.

Keywords: Ethics of sham surgery, evidence-based ethics, Parkinson's disease, patient-centered methods, stem cells, surgical placebo-controlled trials

## INTRODUCTION

Sham or placebo surgery imitates an actual surgical procedure but without providing research participants with the assumed therapeutic benefits of an intervention. The purpose of placebo surgery is to enable the implementation of double-blind surgical placebo-controlled trials (SPTs) for evaluating the safety and efficacy of novel procedures [1, 2]. In comparison to the moderate risks and burdens associated with assignment to the control arm of a non-surgical double-blind trial, participants in the control groups of SPTs suffer additional physical injuries and burdens. The implementation of SPTs in the late 1990 s initiated a vigorous and ongoing debate, which has

been framed as "a clash between the highest standards of research and the highest standards of research ethics . . . " [3, p.2].

Macklin [4] argued that SPTs were, in principle, unethical as their use failed to ensure the moral treatment of seriously ill people who volunteer to participate in surgical research. The principal reason for this claim has been 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" [4, p. 993]. The challenges for patients who volunteer as participants is not limited only to the physical injuries caused by the sham procedures but also the disruption of their normal therapeutic schedules and everyday lives due to participating in trials which may continue for several years. We need to keep in mind that even though patients give informed consent to participate, most

*Correspondence to: Stephen Polgar, Public Health and Psychology, La Trobe University, Melbourne 3086, Australia. E-mail: s.polgar@latrobe.edu.au.

people in the community generally reject the idea that it is morally acceptable to conduct surgical procedures for reasons other than the provision of treatment [5].

In contrast, utilitarian approaches to medical ethics allow for a reasonable degree of risks, harm and burdens to a small number of volunteers for the sake of the 'greater good', referring in the present context to the accurate determination of the efficacy and safety of surgical procedures. The harms and burdens associated with implementing SPTs are presented as the unavoidable consequences of developing and evaluating new treatments, including innovative surgical procedures, in a methodologically rigorous fashion [6]. This position was succinctly expressed by Savulescu, et al. [2] as, "SPTs are necessary and ethical as long a certain conditions are fulfilled" (p. 776). Bioethicists and researchers explored and debated the details of what constitutes these conditions [3, 7–9].

Many well-considered recommendations have been advanced to minimise the probability of physical harms and to ensure the safety, autonomy and respectful treatment of the participants of SPTs. However, the relevance of these recommendations is contingent on the acceptance of the absolute necessity of selecting SPTs as the trial design. It is understood that if there are research designs available which are less onerous for the participants but nevertheless provide valid evidence to advance a research program, then there is no ethical justification for implementing SPTs [10, 11].

There has been a strong impetus to shift medical ethics from a theoretical, normative base to a consequentialist, evidence-based approach. As noted by Kim [12], "What sounds good in theory can in fact produce counterproductive results if not implemented with a thorough understanding of the facts. Facts matter in ethics" [12, p. 373]. The facts that inform utilitarian ethical decisions for the methodological benefits of SPTs are embedded in the evidence provided by completed and published trials [6, 8, 13]. As Miller [14] originally proposed, the risk to benefit ratios for implementing sham surgical procedures are very different for distinct health disorders, as for example brain surgery for neurodegenerative disorders compared to peripheral pain conditions. Therefore, it is not useful to attempt to evaluate the universal utility of SPTs, but rather we need to focus on the evidence for the development of treatments for specific classes of disorders.

In this paper, we investigate the evidence for the necessity of SPTs as specifically applicable to the translation of recently discovered stem cell lines into safe and effective cell-based therapies for people with PD.

## ARGUMENT FOR THE NECESSITY OF SHAM SURGERY

The theoretical justification for the necessity of sham surgery is based on the methodological advantages of implementing double-blind trials for evaluating the safety and efficacy of interventions. In addition to controlling for confounding extraneous variables, the concealment of allocation enables researchers to control for placebo effects and observer bias. Therefore, double-blind randomised trials ensure internal validity and are held by many neuroscientists and clinicians as the 'gold standard' of research designs for evaluating safety and efficacy of interventions [15].

The rationale for interpreting the results of an SPT for PD involves subtracting the average changes found in the placebo-controlled group from the results of the group receiving the active surgical procedure as indicated in Fig. 1. This adjustment enables researchers to reduce the probability of 'false positive' inferences, that is the overestimation of the true efficacy of cell transplantation. The importance of controlling for false-positive decisions is the crucial justification for using SPTs, aiming to protect patients and the community from the introduction of ineffectual and possibly harmful surgical interventions [1]. The position adopted by the majority of North American neuroscientists and bioethicists has been that
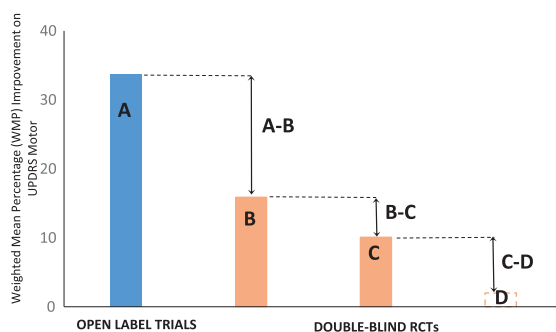


Fig. 1. Overall results for five open label and six double-blind RCTs (Gross et al., 2011 [24]). A) Average score of active groups in open-trials [27–30, 32]. B) Active groups in SPTs [6, 23, 24, 32–34]. C) Sham controls in SPTs [6, 23, 24, 32–34]. D) Unoperated controls (no data). A, B) Difference between the active groups under the two trial conditions. B, C) Estimated effect size in SPTs. C, D) Difference between placebo groups and unoperated controls. WMP, weighted mean percentage.

randomised controlled trials (RCTs) which include sham operated control groups are necessary when conducting research in the field of cellular therapies [16–18].

Although RCTs are generally recognised as excellent designs for ensuring the internal validity of a trial, we also need to consider the issue of external validity [15, 19]. Specifically it has been suggested that the double-blind characteristics of the SPTs may compromise their ecological validity, that is the accuracy of generalising the results of a set of trials to real life settings [20, 21]. Randomised trial designs have their strengths and limitations; therefore, their selection and methodological advantages have to be evaluated in the contexts in which the trials are conducted [15, 19]. Given that there are published SPTs developed to evaluate the efficacy of cellular therapies for PD, it is essential to examine the evidence produced by these trials in order to determine whether the results of published trials demonstrate the evidence-based utility of sham surgery.

We are proposing that the assumptions for the theoretical necessity of SPTs for PD may be stated as a set of empirically testable hypotheses as listed in Table 1. The key question investigated in the present article is whether the evidence provided by published trials confirms the hypotheses below.

Table 1
Hypotheses for the necessity of sham surgery in evaluating cell transplantation

1. There are large and consistent placebo effects associated with cellular therapies for PD.
2. The placebo effects associated with surgical therapies for PD are significantly greater than those associated with non-surgical interventions.
3. Subtracting the results of the placebo arm from those of the active group provides the most accurate estimate of the true efficacy of cellular therapies.
4. Without implementing SPTs, it is not possible to avoid the burdens and expenses associated with the clinical introduction of inadequately evaluated cellular therapies.

## EVIDENCE FOR THE UTILITY OF SHAM SURGERY

Before evaluating the evidence, it is relevant to clarify terminology associated with the placebo concept as defined in recent reviews [8, 22]. The improvement from baseline to the designated endpoint in sham operated control groups is referred to as the 'overall placebo effect' or the 'placebo response'. The overall placebo effect is influenced by a variety of confounding factors, as for example, the natural progression of a disease or the impact of life events experienced by patients during post-surgery recovery [8].

The term 'true placebo effect', refers to the actual therapeutic benefits associated with sham surgery; that is the recovery targeted by the treatment and objectively demonstrated by the participants. It has been suggested that the additional non-specific components of the placebo effect may be estimated by including a non-surgical pharmacological treatment arm in an SPT [8]. It is important to note that all the participants in SPTs reviewed here continued to receive 'best medical therapy', therefore, the control group would simply represent the group which continued to receive optimal pharmacological intervention. However, as Fig. 1 indicates, there were no additional controls other than the sham operated control groups included in any of the SPTs.

Recently several meta-analyses have been published, aiming to identify the differences between placebo and active groups evident in completed SPTs. Each of these meta-analyses [8, 22] relied on different criteria for selecting the trials which were the most suitable for answering the research questions being investigated by the reviewers. Even though a number of SPTs of cellular therapies for PD were included, the meta-analyses included a very heterogeneous set of trials. For example, Wartolowska, et al. [22] analysed 47 SPTs, targeting a broad range of disorders including pain, snoring, heart burn as well as PD. Analysing the ethics of sham surgery needs to take into account the fact that the severity of surgical interventions range from relatively minor incisions such as those for arthroscopic knee surgery to the very demanding neurosurgical procedures for the imitation of the transplantation of cells for PD [14, 23]. As stated earlier, it is essential, therefore, to focus on specific categories of interventions in order to evaluate the utility of sham surgery for PD [14].

In the present analysis, we relied on data collected by Gross and colleagues [24] as the source for the statistics used to estimate the numerical values for the variables shown in Fig. 1 and Table 2. This collection of results for cell and gene therapy trials were published in the context of an SPT aimed at evaluating the safety and efficacy of microcarrier-bound human retinal pigment cells [24]. Although this collection of trials in no way constitutes a systematic review conforming to current 'PRISMA' [25] standards, there were several advantages to using this database in the absence of a relevant meta-analysis.

Table 2
Summary of the weighted mean percentage improvement(s) (WMP). Based on Gross et al. [24]

| Factors | Study groups | Combined Sample Size ($N_i$) | Weighted Mean Percentage Improvement (WMP) |
|---|---|---|---|
| A | Active groups in open-trials | 41 | 33.7 |
| B | Active groups SPTs | 148 | 15.9 |
| A-B | Difference between the active groups under the two trial conditions | – | 17.8 |
| C | Sham controls in SPTs | 125 | 10.1 |
| B-C | Estimated effect size in SPTs | – | 5.8 |
| D | Unoperated controls | 0 | No results |
| C-D | Difference between placebo groups and unopened controls | 0 | No results |

Combined Sample Size ($N_i$) represents the number of participant in each of treatment categories, as defined in Fig. 1.

Table 3
Summary of key methodological problems associated with SPTs for the development of cellular therapies for PD

- The placebo effects in the SPTs of moderate in size and demonstrate a great variability across trials.
- There is no evidence available that placebo effects associated with SPTs are actually greater than placebo effects in non-surgical RCTs.
- There is no evidence for true placebo effects in SPTs as represented by objective, structural and functional improvement in people with PD.
- The ecological validity of SPTs are problematic, the results may not accurately generalise to patients living in everyday environments.
- It is uncertain if the SPTs provide accurate evidence for the true benefits of cellular therapies.
- Prematurely implemented SPTs may lead to the premature discontinuation of research into potentially beneficial cell lines.

Considering the strong commitment to the necessity of STPs by a large proportion of researchers in the field [17, 18, 26], we think that there is a strong level of confirmatory bias involved in the selection and interpretation of data for evaluating the utility of sham surgery. Therefore, given that we are challenging the use of SPTs for evaluating neurosurgical interventions for PD [13, 20, 21], it seemed appropriate to use the database collected by Gross and colleagues. It is very clear, as indicated by their concluding comment in their SPT, "most importantly, the findings emphasise the importance of using a placebo controlled, double-blind design when moving from small, open-label, single-centre studies to large, multi-group, multicentre studies of new surgical treatments", [24, p. 517], that this database was produced by researchers who strongly support the use of SPTs.

The collected results used in the present analysis included outcomes from 5 open-label trials [27–31] and 6 associated SPTs [6, 23, 24, 32–34] for evaluating the efficacy and safety of cellular therapies for PD. These published studies provided the database for comparing effect sizes in the active groups in open-label and SPT studies.

The outcome measure for determining the effect sizes was the United Parkinson's Disease Rating Scale, motor component, under no medication conditions (UPDRS, motor off). This outcome measure seemed relevant given that it has emerged as the most favoured primary outcome measure for determining the efficacy of cellular therapies for PD [35].

The results of the present meta-analysis are shown in Table 2 and graphed in Fig. 1. A, B and C, represents effect sizes and placebo effects. The associated numerical values (Table 2) represent the averaged changes on UPDRS motor, expressed as the Weighted Mean Percentage (WMP) change in the outcomes (Fig. 1 and Table 2). UPDRS (motor) is a standardised measure of the degree of movement problems including tremor, rigidity and bradykinesia experienced by patients. Therefore, high scores are associated with more severe motor signs and a decrease in scores indicates improvements in motor functioning (Fig. 1 and Table 2). In order to facilitate visual presentation of the averaged results, improvements on UPDRS (motor, off) were represented as positive outcomes on Fig. 1 and Table 2.

Regardless of the advantages of using the database prepared by Gross and colleagues, when we take into account the significant differences in the timing for the endpoints in the SPTs (ranging from 6 months to 24 months) and the diverse nature of the transplants, strong scepticism is required regarding the precision of the values shown in Table 2. The averaged scores are reported here should be read as approximate indicators for major trends evident in the database. The averaged effect sizes were quantified as percentages, but the interpretation of the evidence refers to previously identified clinically important improvements in motor functioning in people with PD [36].

## PLACEBO EFFECT IN SURGICAL TRIALS

A key premise of the argument for the necessity of SPTs is that the strength of the placebo effect associated with surgical interventions is much greater than the placebo effects associated with non-surgical trials [1, 37]. Consequently, it is a strongly held belief that control groups provided with the best available pharmacological treatment are unsuitable, as their use would lead to the overestimation of the benefits of the intracerebral transplantation of cells [1, 3].

In the present analysis, the value of C, representing the overall placebo effect (Table 2) was found to be an average of 10.1% change from baseline, corresponding to an approximately 5 point improvement on UPDRS (motor off) in the sham operated group. This is equivalent to a modest to moderate effect size as determined by Shulman and colleagues [36] on UPDRS (motor). Therefore, the results clearly support the presence of an overall placebo effect in SPTs for evaluating the efficacy of cellular therapies for PD. However, there was a remarkable degree of variability in the size of the placebo effect across the 6 sham operated groups, ranging from an 18% deterioration [34] to a 20.7% improvement [24].

The moderate and highly variable placebo effect size associated with sham surgery as identified in the present review is consistent with the results of a meta-analysis conducted by Wartolowska and colleagues [22]. In this meta-analysis, the SPTs selected were categorised based on the way in which the primary outcomes were measured. The category of 'Assessed' outcomes included 3 SPTs which were also included in the present review. The reported pooled standardised mean difference (SMD) for this category was 0.22 with CI 95%: –1.16 to 1.5. This indicates a very modest and highly variable placebo effect with the wide CI 95% indicating the absence of statistical significance.

One of the many sources of variability relevant to the present analysis is the year of publication. The average change from baseline in the placebo response published in the first three trials [6, 23, 34] between 2001 and 2006 was calculated as a reduction of 2.6%. That is, the sham operated groups demonstrated a modest degree of deterioration on UPDRS (motor) at the endpoint of the trials, indicating no overall placebo effect. In contrast, the second set of three trials [24, 32, 33] published between 2010 and 2011 indicated an average improvement of 17.9% in the sham operated groups. In effect, there has been a

remarkable, clinically important 20.1% [38] increase in the placebo effect between the years of 2001 [6, 23, 24] and 2011 [32–34].

It can be said with hindsight that regardless of the actual state of the evidence, the advocates of sham surgery have consistently expressed a resolute conviction in the presence of a massive placebo effect associated with surgical interventions for PD [18, 24, 38]. It is interesting to note that a parallel increase in the size of the placebo effect has been found in clinical trials evaluating the efficacy of psychoactive drugs. These apparent increases in the placebo effect appear most strongly in trials conducted in the USA, but not in Europe or Asia [39]. The question of why the placebo effect varies so dramatically requires close scrutiny.

Another unanswered question is the level of improvement required to provide acceptable evidence for identifying persons as 'placebo responders' in sham arm of an SPT. For example, Ko and colleagues [40] aimed to study the patterns of activity in neural networks associated with mediating placebo effects in patients with PD. The participants included in this study were previously assigned to the sham arm of an SPT investigating the safety and efficacy of 'AAV-GAD' infusion [32].

Ko, et al. [40] selected as 'sham responders' the participants who demonstrated improvements of 2 points or more on UPRDS motor under blinded conditions. In contrast, in the original AAV-GAD trial, LeWitt and colleagues [32] defined a minimum 9 point improvement as the indication for a clinically meaningful response. Previously, Shulman and colleagues [36] indicated that a minimal 2.3 point improvement on UPDRS motor was required to demonstrate even a modest clinically important effect size. We suggest that a degree of consistency in referring to a trial participant as placebo responder would be achieved if researchers used the same standards as for identifying clinically meaningful outcomes in both the placebo and the active group.

Based on evidence from surgical and non-surgical RCTs, it is apparent that subjective outcome measures are more sensitive to placebo effects than objective measures [41]. For example, as we mentioned earlier, Wartolowska and colleagues [22] reviewed a heterogeneous set of SPTs and reported the outcomes under different classes of measures for the placebo responses. The pooled standardised effect for 'Subjective' outcomes was moderate, (0.64, CI 95%: 0.51–0.77) indicating a significant placebo effect. In contrast, the pooled standardised effect size for

'Objective' outcomes indicated that there was no significant placebo effect (0.11, CI 95%: 0.04–0.26). The implication here is that where researchers are concerned with a strong bias being associated with the assessment of treatment outcomes then the use of objective outcome measures is advisable. A combination of both objective and subjective outcomes has been proposed to ensure the optimal accuracy and reliability of the assessments of recovery in patients with PD [35].

As discussed earlier, only an unspecified component of the overall effect size in a sham operated group can be attributed to the true placebo effect [8, 22]. This component refers to the clinically meaningful improvements in the targeted treatment response due to the placebo. We suggest that in context of dopamine (DA) replacement therapies the essential source of evidence for a true placebo effect is the significant recovery of the basal ganglia systems affected by the degenerative processes which characterise PD. However, as shown in two SPTs which involved the transplantation of DA expressing cells [23, 34], there were no increases in intrastriatal DA activity in the sham surgery groups. That is, there is no evidence based on objective PET scan outcomes that the expectation of benefits by research participants resulted in the long-term structural reconstruction of the dopamine expressing nigrostriatal system of human brain.

As stated earlier, none of the SPTs evaluating cellular therapies for PD included an additional arm of non-surgical controls such as best available medical therapies. It is not possible, therefore, on the grounds of the currently available evidence to determine if the placebo effect 'D' (Table 2) due to non-surgical interventions is actually less than those associated with surgical interventions. Hence, it seems that there is no direct evidence to support or reject Hypothesis 2 (Table 1) that placebo effects associated with sham surgery are greater than those associated with other control groups.

While there was clear evidence for an overall placebo effect, there was no convincing evidence for the large and consistent placebo effects associated with sham surgical procedures as predicted by Hypothesis 1. Also, there was no evidence for a true placebo effect, as indicated by the absence of long-term recovery of the nigrostriatal system following sham surgery. Furthermore, there is no direct evidence that the placebo effects due to sham surgical interventions for PD result in greater placebo effects than alternative controls (Hypothesis 2).

## METHODOLOGICAL NECESSITY FOR SPTS

The fact that placebo responses were found to be highly variable across the SPTs reviewed in this paper does not rule out the necessity for selecting sham operated control groups for assessing neurosurgical interventions [42]. The point here is that it is not simply the size of the placebo effect, but rather the difference between the means of the active and placebo arms of a trial that determine decisive evidence for treatment efficacy.

In the present analysis, the overall effect size (WMP) for the active groups in the SPTs, as represented by B was equal to 15.9% (Table 2). Therefore, the true effect size, as represented by B-C (15.9–10.1) is 5.8% (Table 2). Although the 5.8% average improvement must be seen as an approximation, it clearly represents a weak effect size [36]. Consistently with this interpretation, none of the experimental interventions reviewed in the present paper have been introduced as viable, evidence-based treatments for people with PD [24].

In contrast, as indicated by A = 33.7% (Table 2), the open-label trials which preceded the SPTs, indicated significant improvements in the transplanted groups. Similarly, an earlier meta-analysis indicated an average effect size of approximately 39.8% in open-label trials for groups transplanted DA expressing cells [43]. Such improvements could have been judged as evidence for meaningful clinical benefits [36] and initiated the implementation of cellular therapies for PD. It has been argued that the differences between the outcomes in the open-label and double-blind trials constitutes evidence for the scientific and public health advantages of SPTs [17]. Therefore, it appears that the available evidence supports both Hypotheses 3 and 4.

However, at the time when the SPTs for evaluating the efficacy of cell-based therapies for PD were commenced, there was controversy regarding the ethics and the methodological advantages for initiating these trials [44]. In particular, there were concerns that insufficient preliminary evidence was available to justify the initiation of confirmatory STPs [21, 45]. There were numerous unknown parameters, which were likely to have influenced the benefits of cell transplantation, such as criteria for selecting suitable patients or the level of immunosuppression required to ensure the adequate survival of the DA expressing cells [23, 34, 46].

The long-term use of immunosuppressants illustrates the ethical complexities of relying on sham

surgery for creating control groups. Depending on the source and preparation of the transplanted tissue, the optimal survival and functional integration of the DA expressing cells is influenced by the host immune system [47]. The brain is known to be an immunologically privileged site and accordingly Freed and colleagues [23] conducted the first SPTs for evaluating the efficacy of intracerebral cell transplantation without the use of immunosuppressants for either the transplanted or the sham operated participants. In contrast, Olanow and colleagues [34] treated all participants with the immunosuppressant cyclosporine for 6 months following the transplantation procedure. Clearly, from an ethical perspective, the withholding of a powerful but useless drug is preferable to administering cyclosporine for 6 months post sham surgery.

A complication emerging [34] was that when the administration of cyclosporine was stopped at 6 months the initial improvements in the transplanted participants declined and failed to improve from baseline at the 2 year endpoint (Fig. 1, p. 406). Also, even though there was evidence for good graft survival and re-innervation, post-mortem examination of 4 deceased participants indicated inflammatory responses around the grafted cells [34]. Subsequently, a recent protocol called for 12 months of triple immunotherapy in the grafted group, a currently ongoing randomised trial [46, 47]. The design of this RCT did not call for a sham operated control group. Perhaps, the administration of therapeutically needless immunotherapy for 12 months to a group of older people with PD may have been seen as ethically problematic, even by regulatory bodies which are comfortable with the use of sham surgery.

It has been suggested the SPTs were prematurely implemented and may have damaged the progress of the research program of cellular transplantation for PD [21, 26]. That is, the essential lesson learned from the trials was not that SPTs are absolutely necessary, but rather that there is a need for detailed exploratory investigations for enabling the understanding of how new cell lines work and can be effectively translated into best practice [48].

The issue here is that the effect sizes in the active groups were found to be very different under open-label and double-blind conditions. The difference between the endpoints in the outcomes of the active groups is A-B = 17.8% (Fig. 1 and Table 2). This represents a greater than 50% difference, which was also described in other analyses comparing effect sizes in active groups under open-label and SPTs conditions

[21, 49]. This is puzzling, as it is generally assumed that in rigorously executed RCTs the true treatment effects are revealed by subtracting the results of the control group, rather than by compromising the performance of the active group. In other words, it is not entirely clear, that the approximately 17.8% difference is in fact, an accurate representation of the true effect size. An alternative point of view is that the early SPTs may have underestimated the true effect size and generated false negative inferences regarding the efficacy of cell transplantation for PD [22, 26, 50, 51].

Several attempts have been made to explain the reasons for the more than 50% reduction of the recovery of transplanted group when moving from open-label to blinded trials. One hypothesis has been that concealing from the participants the true nature of the treatment they received prevented them from achieving the best possible outcomes following the transplantation of the cells [20, 45, 52]. In open-label trials, in contrast to SPTs, the patients and their caregivers are confident that the cells have been implanted and are therefore, in a position to implement activities which may optimise recovery [20, 21]. Of course, other explanations are possible, as for instance different levels of investigator and patient bias operating under the two different trial paradigms [21].

It is evident that there are numerous methodological problems emerging regarding the use of SPTs for evaluating the efficacy of sham surgery for PD. That is, the difference between the active and the sham-operated arms (i.e., B-C) may not represent accurately the true therapeutic benefits of cellular therapies for PD. Therefore, we suggest that there has been no convincing evidence provided by the present analysis in support of Hypothesis 3.

## ALTERNATIVE APPROACHES

While we are well aware of the methodological advantages of double-blind randomised trials [53], it is essential to consider alternative methodological approaches which enable rigorous research while imposing lesser harms on research participants [19]. There is a range of alternative research designs which are applicable to the advancement of stem cell research into PD, four of which are discussed below.

Several research groups have adopted data collection by blinded assessors as a means of reducing observer bias when implementing trials for evaluating the safety and efficacy of surgical interventions [46,

54]. Although the exact magnitude of observer bias is unknown, the introduction of single-blind RCTs should diminish further the placebo effects from the currently estimated moderate level of approximately 10.1% on UPDRS (motor) to less significant levels. If very strong effect sizes are actualised in the active groups, then placebo effects will have little impact on the accuracy of inferences for clinical efficacy.

As suggested earlier, the available evidence does not indicate structural improvements in the DA expressing nigrostriatal systems of the participants in sham operated control groups [23, 34]. Considering that PD is a progressive neurological disease, we suggest that on the basis of current evidence it is unlikely, that the expectation of treatment benefits by patients would result in the reconstruction of impaired neural systems [40]. The use of a set of interrelated co-primary outcome measures, which would include neural imaging techniques to demonstrate increased striatal structural and functional improvements would provide a comprehensive outcome measure for the efficacy of novel regenerative therapies [35]. The demonstration of strong correlation between objective measures and clinically significant outcomes should obviate the need for sham operated controls [35].

The ongoing remarkable advances in electronics and information technologies are enabling the development of movement and activity sensing tools for the objective and accurate assessment of changes in the symptoms and activity levels of patients with neurological disorders [55]. A recent systematic review [56] provides an overview of the ways in which wearable sensors may be applied to provide objective, unobstructive, long term and accurate information for the symptoms and functional activities in patients with PD. It is understood that this technology needs further development, but it is likely that by the time the safety and efficacy of novel, cell-based treatments for PD are ready to be evaluated, wearable sensors will be available to provide objective and detailed evidence to evaluate the functional efficacy of these interventions.

A variety of research designs other than SPTs is available to determine the safety and preliminary efficacy of new stem cell lines. Evidence from regression analyses and open-label comparative designs are well suited to generate reliable data for answering key research questions such as the optimum quantity and the location for the transplantation of stem cells. We suggest that by the time that a methodologically adequate treatment protocol for the transplantation of cell lines is developed, sufficient evidence will have been collected to identify the key factors which exercise the most important causal effects on the outcomes. Therefore, the statistics for the difference between the transplanted and sham operated controls are not essential to determining the preliminary safety and efficacy of stem cell transplantation [26, 46].

Most importantly, the results of SPTs do not confirm that a given stem cell line is the best available treatment for PD. The benefits of a number of established and/or emerging treatments for PD need to be taken into consideration when selecting acceptable control groups conducting confirmatory RCTs [57]. Also, factors such as the severity and duration of PD or the age of the patients also determine the selection of the most appropriate active control group for valid comparison [51]. It follows that comparison with competing viable interventions such as DBS [10, 11, 20, 51] rather than sham operated controls will provide the decisive evidence required for initiating the clinical introduction of stem cell transplantation for people with specific variants of PD.

All considered, given the availability of alternative methodological approaches, we suggest that the evidence did not support Hypothesis 4 (Table 1). It is important to note that the transnational TRANSEURO trials for cell transplantation are progressing without relying on SPTs [58].

## CONCLUSION

SPTs were introduced to ensure the rigorous evaluation of the safety and efficacy of surgical procedures. We recognise that the earlier SPTs commenced in the 1990 s were ethically justifiable on the well-founded basis of the theoretical advantages of blinded clinical trials. However, the results of published trials have not provided convincing evidence to support the claim that sham operated control groups guarantee accurate decisions regarding the benefits of cellular therapies for PD. The available evidence does not support any of the four hypotheses for justifying the use of sham surgical controls. Moreover, it is even possible that prematurely implemented SPTs may have impeded the progress of the research program.

Ultimately, in the context of the above discussion, the question of how much and what type of information should be provided to patients with PD who volunteer to participate in research involving the transplantation of cells, is critical. While we recognise that the current analysis did not con-

form to the highest current standards for conducting meta-analyses, and while it could be argued that a more rigorously conducted meta-analysis might have uncovered additional evidence for inferring the utility of sham surgery, nonetheless, we maintain that greater disclosure of the ongoing debate regarding alternate research protocols, in addition to SPTs, should form part of the process of obtaining fully informed consent..

Alternative control groups and improved outcome measures may be preferable to placebo controls and provide more valid and clinically applicable evidence. Based on the available facts emerging from the present review, we argue that the use of sham surgery for the evaluation of cellular therapies for PD is unnecessary, and therefore, should be considered unethical from an evidence-based, utilitarian perspective.

## ACKNOWLEDGMENTS

## CONFLICT OF INTEREST

The authors have no conflict of interest to report.

## REFERENCES

[1] Freeman T, Vawter D, Leaverton P, Godbold J, Hauser R, Goetz C, Olanow C (1999) Use of placebo surgery in controlled trials of a cellular based therapy for Parkinson's disease. *New Engl J Med* **341**, 988-992.

[2] Savulescu J, Wartolowska K, Carr A (2016) Randomised placebo-controlled trials of surgery: Ethical analysis and guidelines. *J Med Ethics* **42**, 776-783.

[3] Rogers W, Hutchison K, Skea ZC, Campbell MK (2014) Strengthening the ethical assessment of placebo-controlled surgical trials: Three proposals. *BMC Med Ethics* **15**, 78.

[4] Macklin R (1999) The ethical problems with sham surgery in clinical research. *N Engl J Med* **341**, 992-996.

[5] Leeds HS (2003) Social aspects of sham surgeries. *Am J Bioethics* **3**, 70-71.

[6] Lang AE, Gill S, Patel NK, Lozano A, Nutt JG, Penn R, Brooks DJ, Hotton G, Moro E, Heywood P (2006) Randomized controlled trial of intraputamenal glial cell line–derived neurotrophic factor infusion in Parkinson disease. *Ann Neurol* **59**, 459-466.

[7] Horng S, Miller FG (2003) Ethical framework for the use of sham procedures in clinical trials. *Crit Care Med* **31**, S126-S130.

[8] Probst P, Grummich K, Harnoss JC, Hüttner FJ, Jensen K, Braun S, Kieser M, Ulrich A, Büchler MW, Diener MK (2016) Placebo-controlled trials in surgery: A systematic review and meta-analysis. *Medicine* **95**, e3516.

[9] Swift TL (2012) Sham surgery trial controls: Perspectives of patients and their relatives. *J Empir Res Hum Res Ethics* **7**, 15-28.

[10] Dekkers W, Boer G (2001) Sham neurosurgery in patients with Parkinson's disease: Is it morally acceptable? *J Med Ethics* **27**, 151-156.

[11] London AJ, Kadane JB (2002) Placebos that harm: Sham surgery controls in clinical trials. *Stat Methods Med Res* **11**, 413-427.

[12] Kim SY (2004) Evidence-based ethics for neurology and psychiatry research. *NeuroRx* **1**, 372-377.

[13] Polgar S, Ng J (2005) Ethics, methodology and the use of placebo controls in surgical trials. *Brain Res Bull* **67**, 290-297.

[14] Miller FG (2003) Sham surgery: An ethical analysis. *Am J Bioethics* **3**, 41-48.

[15] Frieden TR (2017) Evidence for health decision making—beyond randomized, controlled trials. *N Engl J Med* **377**, 465-475.

[16] Kim SY, Frank S, Holloway R, Zimmerman C, Wilson R, Kieburtz K (2005) Science and ethics of sham surgery: A survey of Parkinson disease clinical researchers. *Arch Neurol* **62**, 1357-1360.

[17] Vawter D, Gervais K, Prehn A, DeVries R, Freeman T, Garrett J, Mcindoo T (2004) Placebo-controlled surgical trials: Perspectives of Parkinson's disease (PD) researchers. *Exp Neurol* **187**, 221-222.

[18] Galpern WR, Corrigan-Curay J, Lang AE, Kahn J, Tagle D, Barker RA, Freeman TB, Goetz CG, Kieburtz K, Kim SY (2012) Sham neurosurgical procedures in clinical trials for neurodegenerative diseases: Scientific and ethical considerations. *Lancet Neurol* **11**, 643-650.

[19] Cartwright N (2007) Are RCTs the gold standard? *BioSocieties* **2**, 11-20.

[20] Polgar S, Karimi L, Morris M (2013) Stem cell therapy for Parkinson's disease: Are double-blind randomized control trials the best design for quantifying therapy outcomes. *J Neurol Neurophysiol* **4**, 1-6.

[21] Polgar S, Karimi L, Buultjens M, Morris ME (2016) A critical evaluation of the methodological obstacles to translating cell-based research into an effective treatment for people with Parkinson's disease. *Neurorehabil Neural Repair* **30**, 845-853.

[22] Wartolowska KA, Feakins BG, Collins GS, Cook J, Judge A, Rombach I, Dean BJ, Smith JA, Carr AJ (2016) The magnitude and temporal changes of response in the placebo arm of surgical randomized controlled trials: A systematic review and meta-analysis. *Trials* **17**, 589.

[23] Freed CR, Greene PE, Breeze RE, Tsai W-Y, DuMouchel W, Kao R, Dillon S, Winfield H, Culver S, Trojanowski JQ (2001) Transplantation of embryonic dopamine neurons for severe Parkinson's disease. *N Engl J Med* **344**, 710-719.

[24] Gross RE, Watts RL, Hauser RA, Bakay RA, Reichmann H, von Kummer R, Ondo WG, Reissig E, Eisner W, Steiner-Schulze H (2011) Intrastriatal transplantation of microcarrier-bound human retinal pigment epithelial cells versus sham surgery in patients with advanced Parkinson's disease: A double-blind, randomised, controlled trial. *Lancet Neurol* **10**, 509-519.

[25] Liberati A, Altman DG, Tetzlaff J, Mulrow C, Gøtzsche PC, Ioannidis JP, Clarke M, Devereaux PJ, Kleijnen J, Moher

D (2009) The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: Explanation and elaboration. *PLoS Med* **6**, e1000100.

[26] Katsnelson A (2011) Why fake it? How 'sham' brain surgery could be killing off valuable therapies for Parkinson's disease. *Nature* **476**, 142-145.

[27] Gill SS, Patel NK, Hotton GR, O'Sullivan K, McCarter R, Bunnage M, Brooks DJ, Svendsen CN, Heywood P (2003) Direct brain infusion of glial cell line–derived neurotrophic factor in Parkinson disease. *Nat Med* **9**, 589.

[28] Hauser RA, Freeman TB, Snow BJ, Nauert M, Gauger L, Kordower JH, Olanow CW (1999) Long-term evaluation of bilateral fetal nigral transplantation in Parkinson disease. *Arch Neurol* **56**, 179-187.

[29] Kaplitt MG, Feigin A, Tang C, Fitzsimons HL, Mattis P, Lawlor PA, Bland RJ, Young D, Strybing K, Eidelberg D (2007) Safety and tolerability of gene therapy with an adeno-associated virus (AAV) borne GAD gene for Parkinson's disease: An open label, phase I trial. *Lancet* **369**, 2097-2105.

[30] Marks Jr WJ, Ostrem JL, Verhagen L, Starr PA, Larson PS, Bakay RA, Taylor R, Cahn-Weiner DA, Stoessl AJ, Olanow CW (2008) Safety and tolerability of intraputaminal delivery of CERE-120 (adeno-associated virus serotype 2–neurturin) to patients with idiopathic Parkinson's disease: An open-label, phase I trial. *Lancet Neurol* **7**, 400-408.

[31] Stover NP, Bakay RA, Subramanian T, Raiser CD, Cornfeldt ML, Schweikert AW, Allen RC, Watts RL (2005) Intrastriatal implantation of human retinal pigment epithelial cells attached to microcarriers in advanced Parkinson disease. *Arch Neurol* **62**, 1833-1837.

[32] LeWitt PA, Rezai AR, Leehey MA, Ojemann SG, Flaherty AW, Eskandar EN, Kostyk SK, Thomas K, Sarkar A, Siddiqui MS (2011) AAV2-GAD gene therapy for advanced Parkinson's disease: A double-blind, sham-surgery controlled, randomised trial. *Lancet Neurol* **10**, 309-319.

[33] Marks Jr WJ, Bartus RT, Siffert J, Davis CS, Lozano A, Boulis N, Vitek J, Stacy M, Turner D, Verhagen L (2010) Gene delivery of AAV2-neurturin for Parkinson's disease: A double-blind, randomised, controlled trial. *Lancet Neurol* **9**, 1164-1172.

[34] Olanow CW, Goetz CG, Kordower JH, Stoessl AJ, Sossi V, Brin MF, Shannon KM, Nauert GM, Perl DP, Godbold J (2003) A double-blind controlled trial of bilateral fetal nigral transplantation in Parkinson's disease. *Ann Neurol* **54**, 403-414.

[35] Polgar S, Karimi L, Buultjens M, Morris ME, Busse M (2018) Assessing the efficacy of cell transplantation for Parkinson's disease: A patient-centered approach. *J Parkinsons Dis* **8**, 375-383.

[36] Shulman LM, Gruber-Baldini AL, Anderson KE, Fishman PS, Reich SG, Weiner WJ (2010) The clinically important difference on the unified Parkinson's disease rating scale. *Arch Neurol* **67**, 64-70.

[37] Albin R (2002) Sham surgery controls: Intracerebral grafting of fetal tissue for Parkinson's disease and proposed criteria for use of sham surgery controls. *J Med Ethics* **28**, 322-325.

[38] Albin R (2005) Sham surgery controls are mitigated trolleys. *J Med Ethics* **31**, 149-152.

[39] Alphs L, Benedetti F, Fleischhacker WW, Kane JM (2012) Placebo-related effects in clinical trials in schizophrenia: What is driving this phenomenon and what can be done to minimize it? *Int J Neuropsychopharmacol* **15**, 1003-1014.

[40] Ko JH, Feigin A, Mattis PJ, Tang CC, Ma Y, Dhawan V, During MJ, Kaplitt MG, Eidelberg D (2014) Network modulation following sham surgery in Parkinson's disease. *J Clin Invest* **124**, 3656-3666.

[41] Hróbjartsson A, Gøtzsche PC (2001) Is the placebo powerless? An analysis of clinical trials comparing placebo with no treatment. *N Engl J Med* **344**, 1594-1602.

[42] Holtedahl R, Brox JI, Tjomsland O (2015) Placebo effects in trials evaluating 12 selected minimally invasive interventions: A systematic review and meta-analysis. *BMJ Open* **5**, e007331.

[43] Polgar S, Morris M, Reilly S, Bilney B, Sanberg P (2003) Reconstructive neurosurgery for Parkinson's disease: A systematic review and preliminary meta-analysis. *Brain Res Bull* **60**, 1-24.

[44] Thompson L, Cohen J (1994) New fight over fetal tissue grafts. *Science* **263**, 600-602.

[45] Dunnett SB, Rosser AE (2014) Challenges for taking primary and stem cells into clinical neurotransplantation trials for neurodegenerative disease. *Neurobiol Dis* **61**, 79-89.

[46] Barker RA, Barrett J, Mason SL, Björklund A (2013) Fetal dopaminergic transplantation trials and the future of neural grafting in Parkinson's disease. *Lancet Neurol* **12**, 84-91.

[47] Freeman TB, Brundin P (2006) Important aspects of surgical methodology for transplantation in Parkinson's disease. In *Restorative Therapies in Parkinson's disease* Springer, pp. 131-165.

[48] Barker RA, Parmar M, Kirkeby A, Björklund A, Thompson L, Brundin P (2016) Are stem cell-based therapies for Parkinson's disease ready for the clinic in 2016? *J Parkinsons Dis* **6**, 57-63.

[49] Alterman RL, Tagliati M, Olanow CW (2011) Open-label surgical trials for Parkinson disease: Time for reconsideration. *Ann Neurol* **70**, 5-8.

[50] Barker RA, Drouin-Ouellet J, Parmar MJNRN (2015) Cell-based therapies for Parkinson disease—past insights and future potential. *Nat Rev Neurol* **11**, 492.

[51] Buttery PC, Barker RA (2014) Treating Parkinson's disease in the 21st century: Can stem cell transplantation compete? *J Comp Neurol* **522**, 2802-2816.

[52] Polgar S, Borlongan CV, Koutouzis TK, Todd SL, Cahill DW, Sanberg PR (1997) Implications of neurological rehabilitation for advancing intracerebral transplantation. *Brain Res Bull* **44**, 229-232.

[53] Polgar S, Thomas SA (2013) *Introduction to Research in the Health Sciences*, Elsevier Health Sciences.

[54] Weaver FM, Follett K, Stern M, Hur K, Harris C, Marks WJ, Rothlind J, Sagher O, Reda D, Moy CS (2009) Bilateral deep brain stimulation vs best medical therapy for patients with advanced Parkinson disease: A randomized controlled trial. *JAMA* **301**, 63-73.

[55] Steins D, Dawes H, Esser P, Collett J (2014) Wearable accelerometry-based technology capable of assessing functional activities in neurological populations in community settings: A systematic review. *J Neuroeng Rehabil* **11**, 36.

[56] Rovini E, Maremmani C, Cavallo F (2017) How wearable sensors can support Parkinson's disease diagnosis and treatment: A systematic review. *Front Neurosci* **11**, 555.

[57] Stoker TB, Barker RA (2018) Regenerative therapies for Parkinson's disease: An update. *BioDrugs* **32**, 357-366.

[58] Barker RA, Drouin-Ouellet J, Parmar M (2015) Cell-based therapies for Parkinson disease—past insights and future potential. *Nat Rev Neurol* **11**, 492.