The use of placebos in controlled trials of surgical interventions: a brief history

KA Wartolowska, DJ Beard and AJ Carr
NIHR Musculoskeletal Biomedical Research Unit, Botnar Research Centre, Headington, Oxford OX3 7LD, UK
Corresponding author: KA Wartolowska. Email: karolina.wartolowska@phc.ox.ac.uk

Introduction

Inferences about the effects of treatments, including surgical treatments, rely on making comparisons. These comparisons may be with patient’s symptoms before a treatment has been applied. For example, the return of hearing after puncturing the ear drum (tympanotomy) in certain kinds of longstanding deafness can be so dramatic that the change can be confidently ascribed to the treatment.1,2 More usually, treatments have more modest effects, and alternative treatments may differ from each other only slightly, if at all. In these circumstances, disagreements are common about the mechanisms, the magnitude of any effects and the value of a particular treatment. Examples include disputes about different ways of treating wounds,3,4 the timing and methods of limb amputations,5–10 and about lithotripsy as an alternative to lithotomy.11–13

Reducing biases in assessing the effects of surgery

Obtaining reliable estimates of treatment effects, especially when differences between the active and control treatment are moderate, requires the effects of treatments to be distinguished from placebo effects, the effects of biases and the play of chance. Of the many biases that can lead to unreliable estimates of the effects of treatments, there are two in particular that researchers have attempted to reduce over the past century.

The first of these, allocation bias, happens when the patients in treatment comparison groups differ in prognosis, so that it is impossible to assess whether any differences in outcomes following treatment are due to differential effects of the treatments or differences in their prognoses before treatments. Towards the end of the 19th century, prospective, alternate allocation to treatment comparison groups was introduced to generate groups of patients with similar prognoses.14 Surgeons were among those who contributed to this methodological development,15–18 which, half a century later, led on to the adoption of concealed random allocation to treatment comparison groups.14,19,20

The second of the two key biases, observer bias, happens when knowledge of treatment allocation may influence the way that treatments are administered and their possible effects measured. This bias was recognised to be a problem and steps taken to reduce it towards the end of the 18th century. A Royal Commission in Paris used blindfolds to assess whether, as claimed by Anton Mesmer, a force he had dubbed ‘animal magnetism’ was real.21–23 A few years later, Haygarth24 compared the effects of metallic ‘tractors’ with sham, wooden ‘tractors’ in patients with arthritis. These studies helped to demonstrate that patients’ expectations and perceptions could result in ineffective treatments being deemed effective. The need for blinding and placebo controls may be debatable if outcome measures are objective, such as, for example mortality. However, when outcome measures are subjective and reflect satisfaction with the treatment in domains important to patients, such as pain, function and quality of life,25 comparisons of active treatment with placebo are important to reduce biased estimates of treatment effects.26–29

This bias may be caused by patients’ expectations of treatment effects and what they were told to expect from the treatment. They may report symptoms based on their ‘hunches’ about treatment being effective or give answers they believe are ‘correct’ or expected from them, for example, because it would be regarded as impolite not to report improvement.30,31

Blinding means concealing the treatment allocation from patients and any other people involved in the trial, such as assessors or researchers collecting data, caregivers, data analysts, who may bias the results of the trial by knowing the group to which patients have been allocated. Blinding also improves patient retention in trials, for example, across the trials we reviewed the withdrawal rate was low.
Blinding of patients may also reduce the likelihood of patients in the control group not following the assigned treatment, the so-called adherence bias, and help to prevent the control group from seeking and receiving additional treatment outside the trial, which may lead to contamination of the control group. Although designing and carrying out unbiased-controlled trials of surgery is undoubtedly challenging, failure to take steps to reduce observer biases in surgical trials results in untrustworthy estimates of treatment effects.

A few examples of sham/placebo treatments in the 1950s

Although adoption of blinding and use of placebos in drug trials began in the 19th century, these measures were not used in surgical trials until the middle of the 20th century; even today, surgical comparison groups receiving sham treatments or placebos remain rare.

The earliest example of the use of placebo treatments in surgical research of which we are aware is a little known psychosurgical study reported in 1953. Some patients with psychoses underwent anterior cingulate cortex isolation, i.e. excision of a part of cingulum, but four patients had only a skin incision and removal of a fragment of skull bone without any alteration to brain anatomy. None of the patients who had placebo treatment improved and all of them subsequently underwent excision of a fragment of cingulum.

Better known early placebo-controlled trials of surgical treatment compared the effects of internal mammary artery ligation with an inactive procedure consisting of bilateral skin incisions without actual artery ligation in patients with angina. The symptoms, exercise tolerance and use of nitroglycerin of patients in the ligation and the sham groups improved to similar extents. The improvement was attributed to patient interaction with the surgeon, the surgeon’s personality, participation in the treatment, alleviation of anxiety and spontaneous improvement.

Use of placebo treatments during the 1980s and 1990s

These early trials in the 1950s studied few patients who were not informed that some of them would receive a placebo intervention. As was noted by Thomsen et al., ‘knowledge of the possibility of having a purely placebo operation probably reduces the efficacy also in the actively treated group, and the outcome of the trial will be equivocal’. In their trial of treatment for Menière’s disease, half of the patients underwent an endolymphatic sac shunt operation, while the other half underwent regular mastoidectomy. The patients were told that the trial was comparing two different operations that were believed to be equally effective. Thomsen et al. recognised the need to blind patients and investigators when subjective outcomes were being used to assess treatment effects:

In dealing with a disease or a syndrome like Menière’s disease, with its fluctuations in intensity of symptoms, spontaneous remission of symptoms, etc., open, uncontrolled studies are without any value in testing a treatment modality, be it medical or surgical. The double-blind technique therefore is an absolute necessity in such investigations, and any leak in the blind must be considered disastrous.

The advent of minimally invasive surgical techniques and endoscopes opened up new possibilities for blinding in surgical trials. A skin incision imitating open surgery was no longer necessary because using a scope meant that there was either no visible surgical wound in the case of endoscopy or bronchoscopy or only a small incision after laparoscopy or arthroscopy. For example, in trials of interventions to treat bleeding oesophageal varices, peptic ulcers and endometriosis, all patients underwent endoscopy or laparoscopy but only some patients had an additional active procedure to help secure haemostasis or remove the lesions. Other investigators used endoscopy to insert intragastric balloons in trials of surgical treatments for obesity.

Increased use of placebo-controlled surgical trials in the 21st century

By the beginning of the current century, the importance of blinding in assessing the effects of surgical treatments had become more widely accepted and there has been a rapid increase in the number of placebo-controlled surgical trials. Some of these trials have sparked heated debates about the role and ethical aspects of placebos in assessing efficacy of surgery, for example, the trials on the effects of dopaminergic neurone transplants in patients with advanced Parkinson’s disease. These studies used a skin incision and burr holes in the skull as well as several additional diagnostic and therapeutic procedures, some of them clinically necessary and some performed to maintain blinding. For example, patients in both study arms underwent magnetic
resonance imaging and positron emission tomography scans, or received drugs to prevent rejection of the transplanted cells or to prevent epilepsy after the transplantation procedure.

Placebo-controlled trials undertaken to assess the effects of arthroscopic debridement or tidal irrigation in patients with knee osteoarthritis have also been controversial. The blinded trial reported by Moseley et al. failed to detect the beneficial effects suggested by the results of earlier, unblinded trials of debridement and lavage.

Increasing use of implants, lasers or other energy-emitting devices to alter tissues has meant that, in some cases, it has been possible to blind surgeons. For example, in a study reported by Freeman et al., the surgeon inserted a catheter under fluoroscopic guidance and handed over the subsequent procedure to a technician, who delivered or did not deliver the radio-frequency energy. In other trials, the delivery system for palatal implants to reduce obstructive sleep apnoea was prepared by the manufacturer to either contain the implant or not, so that surgeons did not know whether they had performed an actual implantation or a sham procedure.

Increasingly sophisticated efforts to maintain blinding in surgical trials

Preparation for the placebo procedure has usually been done in the same way as for the active procedure. Steps have been taken to imitate visual, auditory and physical cues. To mimic sounds, surgeons have been required to talk through the steps of the procedure, to ask for the instruments, to use suction, or ask for a laser or other device to be activated, although it was not used in the placebo group.

Both the people performing interventions and endoscopic images have been screened from patients, either using a surgical drape or by arranging the operating room in such a way that the patient could not see the procedure. In one trial, patients were heavily sedated and required to wear dark goggles.

Surgeons also attempted to imitate sensory cues, for example, by manipulating the knee as if an actual arthroscopy was being performed, injecting saline subcutaneously to mimic tidal irrigation or by splashing saline on the knee to simulate lavage. In a trial of meniscectomy, surgeons used a mechanised shaver (without the blade) pushing it firmly against the patella to simulate the sensations the patient would experience during the surgery. In a trial of intragastric balloon insertion for obesity, operators manipulated the endoscope as it would be manipulated during the balloon insertion and created the sensation of resistance in the stomach.

Even smell during the surgery was imitated to make the placebo procedure indistinguishable from surgery. For example, in the trial by Deviere et al., there were concerns that patients could have known the allocation because the copolymer used in the active arm had a distinct smell. Therefore, in trials of vertebroplasty, a container with cement was opened during the placebo procedure to help with blinding.

Reports of several trials stated that the duration of procedures in the surgical and the control arms was matched, either by imitating the elements of the surgical procedure or by keeping all patients in the operation room for the same length of time. In some trials, however, the placebo procedure was shorter than the surgical intervention, because it was judged unethical to prolong placebo treatment.

Most trials blinded assessors but the surgeon and other staff in the operating room knew the group assignment; however, they did not participate in further treatment, post operative care or follow-up of the patient. In a trial by Thomsen et al., the post-operative care and assessment was done at a different hospital from the surgery. In the trial by Cotton et al., post-operative care was provided by physicians who were blinded when deciding on treatment, and when this was not sufficient, by an evaluating physician at the study site who was also blinded.

The importance of placebo-controlled surgical trials

Comparisons of surgical treatments believed to be active with placebo treatments are important because they reduce biased estimates of treatment effects. Although double-blind randomised trials are costly and often difficult to perform, their cost is a fraction of the costs to patients and health services of ineffective, invasive and harmful treatments.

Half a century after surgeons and others began to recognise the value of placebo-controlled trials of surgery, there are still no regulations requiring surgical procedures to be shown to be efficacious and safe in the same way there are for drugs. The consequence is much unjustified surgery being used at inevitable risk to the wellbeing of patients and expense to health services. Sixty-six placebo-controlled trials in surgery were reported in the first 15 years of the 21st century compared with only 19 during the whole of the second half of the 20th century. This recent increase in the use of placebo-controlled trials of surgery is an encouraging sign that biased estimates of the effects
of surgical treatment are being recognised and addressed.

Declarations
Competing Interests: None declared.

Funding: This work has been supported by funding from the NIHR Oxford Musculoskeletal Biomedical Research Unit. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

Ethics approval: Not applicable
Guarantor: KAW
Contributorship: KAW drafted the article, KAW, DJB, and AJC revised it critically for important intellectual content and approved the version to be published.

Acknowledgements: None.

Provenance: Invited article from the James Lind Library.

References
25. Ware JE. Measuring patients’ views: the optimum outcome measure. SF-36: a medical valid, reliable


