

RESEARCH

Use of placebo controls in the evaluation of surgery: systematic review

 OPEN ACCESS

Karolina Wartolowska *NDORMS research fellow*^{1,2}, Andrew Judge *university research lecturer*^{1,2,3}, Sally Hopewell *senior research fellow*^{2,4}, Gary S Collins *NDORMS senior research fellow*^{2,4}, Benjamin J F Dean *DPhil student*^{1,2}, Ines Rombach *statistician*^{1,2}, David Brindley *DPhil student*^{1,2,5,6}, Julian Savulescu *Uehiro chair in practical ethics*⁷, David J Beard *professor of musculoskeletal sciences*^{1,2,8}, Andrew J Carr *professor of orthopaedic surgery*^{1,2,8}

¹National Institute of Health Research Musculoskeletal Biomedical Research Unit, Oxford, UK; ²Botnar Institute of Musculoskeletal Sciences, Nuffield Department of Orthopaedics, Rheumatology and Musculoskeletal Sciences, University of Oxford, Oxford OX3 7LD, UK; ³MRC Lifecourse Epidemiology Unit, University of Southampton, Southampton General Hospital, Southampton, UK; ⁴Centre for Statistics in Medicine, University of Oxford, Oxford, UK; ⁵The Oxford—UCL Centre for the Advancement of Sustainable Medical Innovation, University of Oxford, Oxford, UK; ⁶Harvard Stem Cell Institute, Holyoke Centre, Cambridge, MA, USA; ⁷Oxford Uehiro Centre for Practical Ethics, University of Oxford, Oxford, UK; ⁸Surgical Intervention Trials Unit (SITU), RCS Surgical Trials Unit, Botnar Institute of Musculoskeletal Sciences, Oxford, UK

Abstract

Objective To investigate whether placebo controls should be used in the evaluation of surgical interventions.

Design Systematic review.

Data sources We searched Medline, Embase, and the Cochrane Controlled Trials Register from their inception to November 2013.

Study selection Randomised clinical trials comparing any surgical intervention with placebo. Surgery was defined as any procedure that both changes the anatomy and requires a skin incision or use of endoscopic techniques.

Data extraction Three reviewers (KW, BJFD, IR) independently identified the relevant trials and extracted data on study details, outcomes, and harms from included studies.

Results In 39 out of 53 (74%) trials there was improvement in the placebo arm and in 27 (51%) trials the effect of placebo did not differ from that of surgery. In 26 (49%) trials, surgery was superior to placebo but the magnitude of the effect of the surgical intervention over that of the placebo was generally small. Serious adverse events were reported in the placebo arm in 18 trials (34%) and in the surgical arm in 22 trials (41.5%); in four trials authors did not specify in which arm the events occurred. However, in many studies adverse events were unrelated to the intervention or associated with the severity of the condition. The existing placebo controlled trials investigated only less invasive

procedures that did not involve laparotomy, thoracotomy, craniotomy, or extensive tissue dissection.

Conclusions Placebo controlled trial is a powerful, feasible way of showing the efficacy of surgical procedures. The risks of adverse effects associated with the placebo are small. In half of the studies, the results provide evidence against continued use of the investigated surgical procedures. Without well designed placebo controlled trials of surgery, ineffective treatment may continue unchallenged.

Introduction

Modern surgery is changing rapidly. Surgical interventions can now be offered to improve function and quality of life not just to save life. The improvement in the safety of surgical procedures and anaesthesia has facilitated this change.¹ The mortality associated with anaesthesia has decreased from between 64 and 100 in 100 000 in the 1940s to between 0.4 and 1 in 100 000 at present.^{1,2} The prevalence of serious adverse events related to surgical interventions has remained relatively constant over the past 10 years, despite an increase in the number of surgical procedures performed each year.^{3,4} The postoperative death rate is between 1.9% and 4% and in most cases is due to the primary disease.^{3,5}

The increase in the applications for surgical procedures has been driven by a greater involvement of technology in surgical procedures. Such technological advances have made many

Correspondence to: K Wartolowska karolina.wartolowska@ndorms.ox.ac.uk

Extra material supplied by the author (see <http://www.bmj.com/content/348/bmj.g3253?tab=related#webextra>)

Terms used to search databases
Risk of bias of included studies

interventions less invasive, more likely to be endoscopic, and less resembling typical open surgery, such as laparotomy. However, these new procedures are often introduced into surgical practice without any formal evaluation of safety and efficacy, such as using randomised clinical trials. This is because, unlike drug products, such verification is currently not mandated by regulatory authorities.⁶ Furthermore, there is generally a scarcity of information reported on the surgical learning curves or the iterative development of a new technique. Both existing and innovative surgical practice clearly needs to be evaluated, and any evaluative method should take account of the unique idiosyncrasies and challenges presented by surgical interventions.

In considering any scientific evaluation, it is important to remember that an outcome of a surgical treatment is a cumulative effect of the three main elements: critical surgical element, placebo effects, and non-specific effects.⁷ The critical or crucial surgical element is the component of the surgical procedure that is believed to provide the therapeutic effect and is distinct from aspects of the procedures that are diagnostic or required to access the disease being treated.⁸ The placebo effects are related to the patients' expectation and the "meaning of surgery," whereas the non-specific effects are caused by fluctuations in symptoms, the clinical course of the disease, regression to the mean, report bias, and consequences of taking part in the trial, including interaction with the surgeons, nurses, and medical staff.^{7 9 10} It is reasonable to assume that surgery is associated with a placebo effect.¹¹⁻¹³ Firstly, because invasive procedures have a stronger placebo effect than non-invasive ones¹² and, secondly, because a confident diagnosis and a decisive approach to treatment, typical for surgery, usually results in a strong placebo effect.¹⁴

Placebo controlled randomised clinical trials of surgical interventions are relatively uncommon.¹⁵⁻¹⁷ Studies published so far have often led to fierce debates on the ethics, feasibility, and role of placebo in surgery.¹⁸⁻²¹ One reason for the poor uptake is that many surgeons, as well as ethicists, have voiced concerns about the safety of patients in the placebo group. Many of the concerns are based on personal opinion, with little supporting evidence.¹⁸⁻²¹ In the absence of any comprehensive information on the use of placebo controls in surgery, and the lack of evidence for harm or benefit of incorporating a placebo intervention, a systematic review of placebo use in surgical trials is warranted.

We assessed whether placebo controls should be used in the evaluation of surgical interventions by systematically reviewing all clinical trials in which the efficacy of surgery was compared with a placebo control.

Methods

Selection criteria

We performed a systematic review adhering to published guidance from the Cochrane Collaboration.²² Studies were eligible if they were randomised clinical trials in which the efficacy of surgery was compared with placebo. We defined surgery as any interventional procedure that changes the anatomy and requires a skin incision or the use of endoscopic techniques; dental studies were excluded. We used the term placebo to refer to a surgical placebo, a sham surgery, or an imitation procedure intended to mimic the active intervention; including the scenario when a scope was inserted and nothing was done but patients were sedated or under general anaesthesia and could not distinguish whether or not they underwent the actual surgery. We did not limit the inclusion criteria to any

particular condition, patient group, intervention, or type of outcome. We excluded studies investigating anaesthesia or other drug substances used perioperatively.

Search strategy

We developed search strategies for three electronic databases: Medline (Ovid), Embase (Ovid), and the Cochrane Central Register of Controlled Trials. We searched the databases from the date of their inception to 14 November 2013, with no restriction on language. (See supplementary appendix 1 for details of the search terms.) We did not systematically search for studies reported only as conference abstracts.

Three reviewers (KW, IR, BJFD) independently screened the initial set of records identified from the search and then screened the full text of any potentially relevant articles. Each reviewer independently assessed the eligibility of each study, and the final list of included studies was agreed by consensus.

We also screened the references of the relevant articles. Furthermore, we searched ClinicalTrials.gov (on 14 November 2013), a database of registered randomised clinical trials, to identify any recently completed or ongoing studies.

Dealing with duplicate publications

If several articles reported outcomes from a single trial (that is, with the same authors, location, patient population, and recruitment dates), we only included the article reporting the main outcome for the trial.

Data extraction

We used a standardised data extraction form to collect information about the characteristics of each study as well as the clinical improvement and superiority of the surgical intervention compared with the placebo for the main study outcome; as reported by the authors in the published article. For each study we extracted the year of publication; study population; condition; intervention; outcomes; sample size; number of participants; number of events, as well as mean and standard deviation for continuous outcomes; and serious adverse events and whether they were related to the procedure. To reduce the chance of errors, the three review authors extracted data separately, checked the entries for consistency, and agreed on a single set of data.

Risk of bias assessment

The three review authors also independently assessed the risk of bias in the included studies using the risk of bias tool criteria recommended by the Cochrane guidelines.^{22 23} In particular we assessed the method of random sequence generation; concealment of treatment; blinding of participants, care providers, and assessors; success of blinding; and use of intention to treat analysis.

Data synthesis

We assessed the beneficial effect of the surgical intervention on the basis of the original conclusions as any improvement in the main outcome of the trial and as superiority of the surgical treatment over the placebo—that is, the additional benefits of the critical surgical element. Moreover, we calculated statistically significant difference between the surgical intervention and placebo using the information reported in the results section of each study. We calculated the odds ratio for binary outcomes and the effect size for continuous outcomes using the effect estimate from analysis of covariance, the

difference in change score, or the difference in postoperative score, depending on the method of analysis and data reported within each individual study. We included only the primary outcome measure, whenever it was explicitly specified. If two primary outcomes were reported, we used both; however, when there were several main outcome measures, we chose those reported in the abstract or those used in other studies, so that the forest plots present similar outcomes. Where necessary, we changed the direction of effect so that the improvement was consistently presented in the same way in the forest plots.

As a measure of harms, we examined serious adverse events and whether they were attributable to the surgical or the placebo intervention. An optimal strategy to identify reports of adverse events does not exist.²⁴ We defined serious adverse events as harmful events that occurred during the trial, such as prolonged hospital stay, and events that required admission to hospital or resulted in death. We summarised the serious adverse events data using a grading system according to their severity, as definitely, likely or unlikely to be serious, and a grading system according to relation between the event and the procedure as definitely, likely, unlikely to be related to the procedure. Wherever possible, we used the results from the intention to treat analysis. Most studies did not provide sufficient information about harms to enable a formal statistical analysis.

Owing to the considerable heterogeneity of conditions, interventions, and outcomes it was not feasible to combine the results of individual studies in a meta-analysis. We present a descriptive analysis of the results of each individual study and present data in tables and figures. All analyses were carried out in Stata (version 12.1).

Results

Study selection

The search of Medline, Embase, and Cochrane databases identified a total of 4543 records. We found an additional 23 articles using a hand search of relevant literature and the references of the articles identified using the database search. Out of these 4566 records, 1597 duplicates were discarded, leaving 2969 records; of these, 2860 did not meet inclusion criteria, and a further 56 studies were excluded after reviewing the full text of the article. Among them were 12 articles reporting additional outcomes of a trial, seven follow-up papers, and seven studies that reported results in a way that made the comparison between the surgical and placebo arm impossible. This resulted in 53 full text articles that met the inclusion criteria and were included in this review. Figure 1⇓ details the study identification and reasons for exclusion.

A keyword search of the ClinicalTrials.gov database identified 14 relevant trials, including two studies already found using the electronic search^{25 26} and 12 studies that were not yet available as full text articles in November 2013.

Risk of bias assessment

All included studies were assessed for risk of bias (see supplementary appendix 2). The risk of bias related to sequence generation, concealment of treatment allocation, and blinding during the procedures was generally low. In some trials the measures undertaken to maintain blinding were not explicitly described; however, these trials used endoscopic techniques in patients under general anaesthesia or sedation. Patients and assessors were blinded in almost all trials; in the trials in which assessors were not blinded, the outcomes were objective. Only 12/53 (23%) trials assessed whether the blinding was successful,

including one in which the patients were likely to guess the correct allocation.⁶⁰

Characteristics of the reviewed trials

Overall, 39 of the 53 (74%) included studies were published after 2000. Most of the trials investigated minor and not directly life threatening conditions, such as severe obesity (n=7; 13%) or gastro-oesophageal reflux (n=6; 11%). The most common type of intervention was endoscopy, with 23 trials (43%) using this technique as a part of the investigated procedure. Thirteen trials (25%) used some exogenous material, implant, or tissue, and a further six used balloons. Most studies reported subjective outcomes such as pain (n=13; 25%), improvement in symptoms or function (n=17; 32%), or quality of life (n=8; 15%). Less than half of the trials (n=22; 42%) reported an objective primary outcome—that is, measures that did not depend on judgment of patients or assessors. The majority of trials were small; the number of randomised participants ranged between 10 and 298, with a median of 60. No placebo controlled surgical trials investigating more invasive surgical procedures such as laparotomy, thoracotomy, craniotomy, or extensive tissue dissection were identified. Tables 1⇓ and 2⇓ list the characteristics of each trial.

Clinical benefit of improvement as reported by the trials' authors

In around three quarters (n=38; 72%) of the studies the authors reported an improvement in both the surgical group and the placebo group (table 3⇓). In a further seven trials,³⁰⁻³⁶ the clinical improvement was observed in the surgical group but not in the placebo group; however, in five of these studies,^{30-33 36} the outcome measures were objective and did not depend on patients' ratings. Finally, in six studies no improvement was reported in either group,³⁷⁻⁴² and one study could not be interpreted in terms of improvement as the main outcome was failure of treatment defined as new or continued bleeding.⁴³

Superiority of surgery over placebo as reported by the trials' authors

In half of the included studies (n=26; 49%) the authors reported superiority of the surgical procedure over the placebo intervention, and in the remaining trials (n=27; 51%) the active surgical procedure was not statistically different from that of the placebo intervention (table 3).

Statistically estimated clinical benefit of critical surgical element in comparison to placebo

Overall, the magnitude of the treatment effect when the active surgical intervention group was compared with the placebo group was small but generally favoured the surgical treatment. The forest plots in figures 2⇓ and 3⇓ present the effect sizes and odds ratios for individual trials.

In 12 studies, the surgical intervention was significantly superior to placebo (tables 1 and 2 and figs 2 and 3).^{28-30 33 34 36 44-49} In 11 studies, the effect of the surgical intervention was significantly better for only some of the reported primary outcome measures.^{31 32 50-58} In 24 studies (45%), there was no statistically significant benefit of the active surgical intervention over the placebo (tables 1 and 2 and figs 2 and 3). In six studies we could not calculate the statistical estimated clinical benefit because the results were reported as median values and could not be used to calculate an effect size.^{27 35 42 59-61}

Harms

In just under half of the trials ($n=23/53$; 43%) the authors stated that there were no serious adverse events, although sometimes they reported minor adverse events (table 2). Three out of 53 studies (6%) did not report any information about adverse events. In the remaining 27 studies (51%), serious adverse events occurred in at least one of the study arms (table 1). In 17 of these 27 studies, serious adverse events were observed in both the surgical and the placebo arms, in five studies serious adverse events were only present in the active group, whereas in four studies the authors did not specify in which group the serious adverse events occurred. Not all serious adverse events were related to the procedure. For example, in six trials on gastrointestinal bleeding, deaths, rebleeds, or continuous bleeding were the main outcome of the study and were a result of the investigated condition and of the procedure being ineffective rather than it being harmful. In only two of these trials, adverse events, such as a perforation, were directly related to the intervention. In several trials, adverse events were rare (<5% of patients) or were unrelated to the procedure—for example, death from other causes. In general, the placebo arm was reported to be safer and adverse events were more serious and more common in the active group.

The interventions in the placebo arm were overall associated with less serious adverse events compared with the active arm, as the main surgical element was omitted and the authors made an effort to minimise risks by withholding part of the intervention—for example, partial burr holes rather than full trepanation⁴⁰ or not administering heparin.³⁹ Often the type of serious adverse events in the placebo group depended on the severity of the investigated condition and the invasiveness of the chosen procedure. Moreover, the serious adverse events in the placebo group were more likely if the procedures involved exogenous material; out of 13 trials using implanted exogenous substances, materials, or tissue, eight reported serious adverse events.^{28 37-40 62-64}

Of the 27 trials in which serious adverse events occurred, 17 trials reported events that were related or likely to be related to the procedure in the surgical group (table 1). Most of the studies did not specify whether the serious adverse events were directly or potentially associated with critical surgical element or other elements of surgical interventions. Complications in the placebo group related or likely to be related to some element of the procedure were reported in nine studies. Harms definitely directly related to the surgical placebo were reported in two studies, and included infections⁶⁴ as well as complications related to the device and the investigated condition itself³⁷; both trials were stopped early because of concerns about safety. In one of the trials on gastrointestinal bleeding, aspiration occurred in two patients in the placebo group that could have been related to the procedure as well as to the condition.³¹

The only adverse event reported as related to anaesthesia was in a trial by Schwartz and colleagues,⁵⁶ in which one patient in the placebo group had a bruise as result of a misplaced intravenous line during sedation.

Discussion

Surgical randomised clinical trials incorporating a placebo arm are rare but this review shows that the results of many of the trials provide clear evidence against continued use of the investigated surgical procedures and in well designed studies the risks of adverse effects are small and the placebo arm is safer than surgery. The identified surgical randomised clinical trials were heterogeneous. The existing placebo controlled trials

investigated only less invasive procedures that did not involve laparotomy, thoracotomy, craniotomy, or extensive tissue dissection. About a half of the reviewed trials showed superiority of the surgical procedure over placebo intervention, but the magnitude of the effect directly related to the crucial surgical element was generally small. The majority of the trials showed an improvement in the surgical group as well as in the placebo group, which would suggest that some surgical procedures may have a placebo effect and that some of the benefits of surgery are related to factors other than the crucial surgical element.

Serious adverse events were reported in half of the reviewed trials, and the severity of possible serious adverse events was often related to the seriousness of the investigated condition and the invasiveness of the chosen procedure. Generally, the incidence of complications in the placebo group was lower than that in the surgical arm.

Strengths and weaknesses of this study

Modern surgery involves not only open surgery but also minimally invasive procedures, implants, and transplants; therefore, the boundary between surgery and other medical procedures was not always clear. In addition, identifying unique studies was not always straightforward—that is, differentiating between papers reporting two similar trials or different outcomes of the same trial.

Many trials reported several outcomes, often without identifying the primary measure. As a consequence we had to report only the outcomes for which the study had been powered, those reported in the abstract, or those used in other similar studies. The lack of a clear primary outcome also meant that we could not report a single primary outcome per study.

The available data only allowed comparison of effects in the surgical arm versus the placebo arm and did not allow for any estimation of the magnitude of the placebo effect. Interestingly, the effects for surgery versus placebo were smaller than expected. Only one study included an observational group.⁵⁶ We specifically chose not to include surgical trials with waiting list controls. Analysis of the placebo effect without controlling for non-specific effects of being in the trial would be flawed, as was pointed out in the comments^{66 67} in the original paper on placebo effects in surgery by Beecher.⁶⁸

The quality of the included studies varied. Many studies were small and used several primary outcome measures or cumulative measures or used a subgroup analysis without comparisons between groups. Reporting of adverse events was poor and many authors did not specify the primary cause of the complications or in which group they occurred. In the recent trials, the adverse events were described in more detail, but further improvement is required.⁶⁹ The standardised way of reporting harms was published only in 2004.⁷⁰

Possible benefits and harms and implications for clinicians and policy makers

The main benefit and value of placebo controlled randomised clinical trials is their ability to show the real efficacy of a surgical intervention. If a procedure is effective and superior over the placebo, the case is strong for it to be commissioned and funded. The opposite is also true; if a surgical procedure has no benefit over placebo the argument is strong for stopping its use. The well designed placebo controlled trial of surgery is a useful tool to challenge the continued commissioning and use of ineffective treatments.

The trial participants usually gain direct benefits, the main one being the perceived improvement attributable to the intervention. Whether this is acceptable by doctors and patients is the subject of an ongoing debate.^{71 72} Indirect benefits may include confirming or disproving the primary diagnosis thereby allowing patients to be referred for a more appropriate treatment. A reduced waiting time or receiving treatment free of charge (for treatment that is ordinarily paid for by patients) is a further potential advantage of taking part in a trial.

In the modern clinical setting the distinction between the research and treatment is less clear and these two are often integrated. The balance between benefits and overall risk to a patient in a trial and in standard clinical care are usually not much different, as both the clinical outcomes and the additional risks and inconveniences caused by additional assessments are not great.^{73 74}

The negative consequences for the trial patients may be considered mainly in terms of harms. In most of the trials, the investigated conditions were non-life threatening and the main aim of surgery was to improve function, symptoms, and quality of life, to reduce pain, or to remove the need for drug use. In these trials the harms were either lack of improvement or complications arising from the procedure, such as perforation after endoscopy. In the trials on gastrointestinal bleeding, the investigated condition was potentially life threatening and the procedure was not elective. In these six trials the negative events, such as bleeding and deaths, were the primary outcome and an indicator that the intervention was ineffective.

Placebo controlled surgical trials raise important ethical concerns¹⁸⁻²⁰ but are justified when there is a genuine equipoise¹⁹; that is, a disagreement in the medical community about whether one treatment is superior to another, because standard treatment does not exist or its efficacy is questioned.

Such trials conform to the ethical principles of non-maleficence—that is, the duty not to inflict expected harm, and justice.²¹ Surgical intervention may be associated with greater risk than drug intervention; therefore, to be justified it must be associated with greater “pure” benefit to outweigh such risks. Surgical trials are justified by equipoise not only when there is uncertainty about pure benefit, but also when there is uncertainty about whether benefit outweighs harms.

If a standard medical treatment is available, it may be offered as a part of the study, as in the trials on tissue transplantation in Parkinson’s disease, in which the participants continued their L-dopa drug. In addition to that, if one of the treatments in the trial becomes recognised as effective or a new treatment becomes available during the conduct of the trial, equipoise is disturbed and ethically the study should be stopped, as was the case in one trial⁶¹ where the intervention was accepted by the insurer as a standard procedure and the trial was terminated early, before it could show the superiority of the active procedure.

Another ethical concern is deception and risk to the trust between patients and surgeons. In the majority of studies and always in the recent trials, patients were informed about taking part in a placebo controlled trial, and informed consent was obtained. In the study by Moseley and colleagues, participants were also asked to write a clause in their notes acknowledging the fact that they might receive placebo intervention and that placebo might not be effective.⁷⁵ If patients are fully informed and give proper consent, then from the ethical point of view the conditions of free consent and autonomy are fulfilled.

Surgery of any form, including placebo surgery, is associated with some level of risk, whereas a placebo tablet or drug is not.

Therefore the balance between risks and benefits in surgical placebo controlled randomised clinical trials is different from that in drug trials. In most of the studies that reported serious adverse events, such complications were expected in the investigated conditions; even in the study on Alzheimer’s disease, mortality was comparable with other trials.³⁷ What matters most is that risk is minimised and that actual harm is as small as possible.⁷⁶⁻⁷⁸ In the reviewed trials, the placebo arm was usually designed to pose as little risk to the participants as possible and to be significantly less risky than the active surgical procedure.

In a situation where there is no certainty that the surgery is effective, the balance between risks and benefits within a study actually may be better in the placebo arm. For example, in the trials on fetal nigral tissue transplantation for Parkinson’s disease, the active treatment was no better than placebo in terms of outcome and resulted in more severe side effects, such as dystonias and dyskinesias. Moreover, the risk of infection and other complications associated with the actual tissue implantation was much higher than the risk in the placebo group. Interestingly, in the study by Freed and colleagues,³⁸ patients who had been in the placebo group still opted for the surgical procedure after the trial, despite the fact that no clear benefit of fetal nigral tissue transplantation had been shown. This may be because patients believe that invasive,¹² new,⁷⁹ and expensive⁸⁰ procedures are actually more effective.

Effects of existing trials on change in practice

The results of the placebo controlled surgical trials performed so far have had a varied impact on clinical care. With the exception of the trials on debridement for osteoarthritis⁷⁵ or internal mammary artery ligation for angina,⁸¹ most of the trials did not result in a major change in practice. Moseley’s study on debridement for knee osteoarthritis was well received and resulted in limiting the recommendations for debridement and lavage for osteoarthritis of the knee to patients only with clear mechanical symptoms such as locking.⁸² The reaction of the medical community to the trials on tissue transplantation to treat Parkinson’s disease was also favourable. Although this treatment is not currently recommended,⁴⁰ the need for more studies on mechanisms of disease and on tissue transplantation is recognised.^{83 84} These studies also provoked a discussion about ethical aspects of randomised clinical trials and placebo.^{18 19 21 85}

In contrast, the results of the trials on the efficacy of vertebroplasty^{62 64} were challenged and their authors were criticised for undermining the evidence supporting this commonly used procedure. The critics acknowledged that the injection of cement might be associated with many side effects, some of them potentially dangerous, but they argued that the treatment was justified because earlier unblinded trials had shown superiority of vertebroplasty over medical treatment.^{62 64 86} This argument against the validity of placebo controlled trials neglected any potential placebo effect of vertebroplasty.

Unanswered questions and future research

The placebo component of surgery has to be explored and better understood. Future surgical trials need to be better designed and to include the observational control arm to allow the estimation of the magnitude of the placebo effect while controlling for non-specific factors such as spontaneous improvement or regression to the mean. We need to know the size of these effects and factors that influence their magnitude in order to properly interpret results of clinical trials in general, especially as many

surgical trials nowadays use an observational design or case registers.

Placebo controlled trials in surgery are as important as they are in medicine, and they are justified in the same way. They are powerful, feasible way of showing the efficacy of surgical procedures. They are necessary to protect the welfare of present and future patients as well as to conduct proper cost effectiveness analyses. Only then may publicly funded surgical interventions be distributed fairly and justly. Without such studies ineffective treatment may continue unchallenged.

As surgery is inherently associated with some risk, it is important that the surgical treatment is truly effective and that the benefits outweigh the risks. Placebo controlled surgical trials are not free from adverse events but risks are generally minimal in well designed trials and the control arm is much safer than the active treatment. However, this review highlighted the need for better reporting of trials, including serious adverse events and their relation to a particular element of surgical procedure.

A need exists to “demystify” and extend the use of the surgical placebo in clinical trials. These should result in a greater acceptance of this type of trial by the surgical community, ethics committees, funding bodies, and patients. In turn, this would lead to more studies, better guidelines on the design and reporting of studies, and a larger body of evidence about efficacy and the risks of surgical interventions. Placebo controlled surgical trials are highly informative and should be considered for selected procedures.

Contributors: KW, AJC, and DJB conceived and designed the study. KW, AJ, SH, and GSC designed the search strategy and statistical analysis. KW, IR, and BJFD collected and managed the data. KW drafted the manuscript and is the guarantor of the study. KW, AJC, DJB, IR, BJFD, DB, JS, SH, and GSC critically revised the manuscript for important intellectual content. All authors had full access to all of the data in the study and can take responsibility for the integrity of the data and the accuracy of the data analysis, and revised and approved the final version of the article.

Funding: This study was funded by the NIHR Oxford Musculoskeletal Biomedical Research Unit. The funding source had no role in the design and conduct of the study, in the collection, analysis and interpretation of the data, or in the preparation, review, or approval of the manuscript.

Competing interests: All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare: the authors are involved in a placebo controlled surgical trial on shoulder pain (NCT01623011); no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

Ethical approval: Not required.

Data sharing: No additional data available.

Transparency: The lead author (the manuscript’s guarantor) affirms that the manuscript is an honest, accurate, and transparent account of the study being reported; no important aspects of the study have been omitted.

- Li G, Warner M, Lange BH, Lin H, Sun LS. Epidemiology of anesthesia-related mortality in the United States, 1999-2005. *Anesthesiology* 2009;110:759-65.
- Gottschalk A, Van Aken H, Zenz M, Standl T. Is anesthesia dangerous? *Dtsch Arztebl Int* 2011;108:469-74.
- Kable AK, Gibberd RW, Spigelman AD. Adverse events in surgical patients in Australia. *Int J Qual Health Care* 2002;14:269-76.
- Zegers M, de Bruijn M, de Keizer B, Merten H, Groenewegen P, van der Wal G, et al. The incidence, root-causes, and outcomes of adverse events in surgical units: implication for potential prevention strategies. *Patient Saf Surg* 2011;5:13.

- Calland JF, Adams RB, Benjamin DK Jr, O'Connor MJ, Chandrasekhara V, Guerlain S, et al. Thirty-day postoperative death rate at an academic medical center. *Ann Surg* 2002;235:690-6; discussion 6-8.
- McCulloch P, Altman DG, Campbell WB, Flum DR, Glasziou P, Marshall JC, et al. No surgical innovation without evaluation: the IDEAL recommendations. *Lancet* 2009;374:1105-12.
- Ernst E, Resch KL. Concept of true and perceived placebo effects. *BMJ* 1995;311:551-3.
- Tenery R, Rakatansky H, Riddick FAJ, Goldrich MS, Morse LJ, O'Bannon JM 3rd, et al. Surgical “placebo” controls. *Ann Surg* 2002;235:303-7.
- Braunholtz DA, Edwards SJL, Lilford RJ. Are randomized clinical trials good for us (in the short term)? Evidence for a “trial effect”. *J Clin Epidemiol* 2001;54:217-24.
- Wood L, Egger M, Gluud LL, Schulz KF, Juni P, Altman DG, et al. Empirical evidence of bias in treatment effect estimates in controlled trials with different interventions and outcomes: meta-epidemiological study. *BMJ* 2008;336:601-5.
- Moerman DE, Jonas WB. Deconstructing the placebo effect and finding the meaning response. *Ann Intern Med* 2002 March 19, 2002;136(6):471-6.
- Kaptchuk TJ, Goldman P, Stone DA, Stason WB. Do medical devices have enhanced placebo effects? *J Clin Epidemiol* 2000;53:786-92.
- De Craen AJ, Tijssen JG, de Gans J, Kleijnen J. Placebo effect in the acute treatment of migraine: subcutaneous placebos are better than oral placebos. *J Neuro* 2000;247:183-8.
- Thomas KB. General practice consultations: is there any point in being positive? *BMJ* 1987;294:1200-2.
- Wenner DM, Brody BA, Jarman AF, Kolman JM, Wray NP, Ashton CM. Do surgical trials meet the scientific standards for clinical trials? *J Am Coll Surg* 2012;215:722-30.
- Gelijns AC, Ascheim DD, Parides MK, Kent KC, Moskowitz AJ. Randomized trials in surgery. *Surgery* 2009;145:581-7.
- Howes N, Chagla L, Thorpe M, McCulloch P. Surgical practice is evidence based. *Br J Surg* 1997;84:1220-3.
- Macklin R. The ethical problems with sham surgery in clinical research. *N Engl J Med* 1999;341:992-6.
- Miller FG. Sham surgery: an ethical analysis. *Am J Bioeth* 2003;3:41-8.
- Hornig SBA, Miller FG. Ethical framework for the use of sham procedures in clinical trials. *Crit Care Med* 2003;31(3 Suppl):S126-30.
- London AJ, Kadane JB. Placebos that harm: sham surgery controls in clinical trials. *Stat Methods Med Res* 2002;11:413-27.
- Van Tulder M, Furlan A, Bombardier C, Bouter L. Updated method guidelines for systematic reviews in the cochrane collaboration. *Spine (Phila Pa 1976)* 2003;28:1290-9.
- Higgins JPT, Altman DG, Gotzsche PC, Juni P, Moher D, Oxman AD, et al. The Cochrane Collaboration’s tool for assessing risk of bias in randomised trials. *BMJ* 2011;343:d5928.
- Loke Y, Price D, Herxheimer A, Group iCAEM. Systematic reviews of adverse effects: framework for a structured approach. *BMC Med Res Methodol* 2007;7:32.
- Gillespie MB, Wylie PE, Lee-Chiong T, Rapoport DM. Effect of palatal implants on continuous positive airway pressure and compliance. *Otolaryngol Head Neck Surg* 2011;144:230-6.
- Steward DL, Huntley TC, Woodson BT, Surdulescu V. Palate implants for obstructive sleep apnea: Multi-institution, randomized, placebo-controlled study. *Otolaryngol Head Neck Surg* 2008;139:506-10.
- Davys HJ, Turner DE, Helliwell PS, Conaghan PG, Emery P, Woodburn J. Debridement of plantar callosities in rheumatoid arthritis: a randomized controlled trial. *Rheumatology (Oxford, England)* 2005;44:207-10.
- Devriere J, Costamagna G, Neuhaus H, Voderholzer W, Louis H, Tringali A, et al. Nonresorbable copolymer implantation for gastroesophageal reflux disease: a randomized sham-controlled multicenter trial. *Gastroenterology* 2005;128:532-40.
- Genco A, Cipriano M, Bacci V, Cuzzolaro M, Materia A, Raparelli L, et al. BioEnterics[reg] IntraGastric Balloon (BIB[reg]): a short-term, double-blind, randomised, controlled, crossover study on weight reduction in morbidly obese patients. *Int J Obes Relat Metab Disord* 2006;30:129-33.
- Arts J, Bisschops R, Blondeau K, Farre R, Vos R, Holvoet L, et al. A double-blind sham-controlled study of the effect of radiofrequency energy on symptoms and distensibility of the gastro-oesophageal junction in GERD. *Am J Gastroenterol* 2012;107:222-30.
- Fleischer D. Endoscopic Nd:YAG laser therapy for active esophageal variceal bleeding. A randomized controlled study. *Gastrointest Endosc* 1985;31:4-9.
- Fullarton GM, Birnie GG, Macdonald A, Murray WR. Controlled trial of heater probe treatment in bleeding peptic ulcers. *Br J Surg* 1989;76:541-4.
- Friedman M, Schalch P, Lin HC, Kakodkar KA, Joseph NJ, Mazloom N. Palatal implants for the treatment of snoring and obstructive sleep apnea/hypopnea syndrome. *Otolaryngol Head Neck Surg* 2008;138:209-16.
- Bajbouj M, Becker V, Eckel F, Mielhke S, Pech O, Prinz C, et al. Argon plasma coagulation of cervical heterotopic gastric mucosa as an alternative treatment for globus sensations. *Gastroenterology* 2009;137:440-4.
- Van Schie CH, Whalley A, Vileikyte L, Wignall T, Hollis S, Boulton AJ. Efficacy of injected liquid silicone in the diabetic foot to reduce risk factors for ulceration: a randomized double-blind placebo-controlled trial. *Diabetes Care* 2000;23:634-8.
- Thompson C, Chand B, Chen Y, Demarco D, Miller L, Schweitzer M, et al. Endoscopic suturing for transoral outlet reduction increases weight loss after Roux-en-Y gastric bypass surgery. *Gastroenterology* 2013;145:129-37.
- Silverberg GD, Mayo M, Saul T, Fellmann J, Carvalho J, McGuire D. Continuous CSF drainage in AD: results of a double-blind, randomized, placebo-controlled study. *Neurology* 2008;71:202-9.
- Freed CR, Greene PE, Breeze RE, Tsai W-Y, DuMouchel W, Kao R, et al. Transplantation of embryonic dopamine neurons for severe Parkinson’s disease. *N Engl J Med* 2001;344:710-9.
- Dowson A, Mullen MJ, Peatfield R, Muir K, Khan AA, Wells C, et al. Migraine Intervention With STARFlex Technology (MIST) trial: a prospective, multicenter, double-blind, sham-controlled trial to evaluate the effectiveness of patent foramen ovale closure with STARFlex septal repair implant to resolve refractory migraine headache. *Circulation* 2008;117:1397-404.
- Olanow CW, Goetz CG, Kordower JH, Stoessl AJ, Sossi V, Brin MF, et al. A double-blind controlled trial of bilateral fetal nigral transplantation in Parkinson’s disease. *Ann Neurol* 2003;54:403-14.
- Freeman BJC, Fraser RD, Cain CMJ, Hall DJ, Chapple DCL. A randomized, double-blind, controlled trial: intradiscal electrothermal therapy versus placebo for the treatment of chronic discogenic low back pain. *Spine* 2009;30:2369-77.
- Baeck LJJ, Liukko T, Rantanen I, Peltola JS, Partinen M, Ylikoski J, et al. Radiofrequency surgery of the soft palate in the treatment of mild obstructive sleep apnea is not effective

What is already known on this topic

Placebo controlled randomised clinical trials on efficacy of surgical procedures are rare and the main concerns are related to the benefits and harms related to such trials

What this study adds

In many of the reviewed trials improvement was present in both arms suggesting that the clinical effect may not be a result of the surgery; moreover, statistical superiority of the surgical procedure over the placebo was small

The absence of an observational or non-treatment control group in most studies precludes any evaluation of the magnitude of the placebo effect

Placebo controlled surgical trials are not free from adverse events, but the harms can be minimised and the placebo arm is generally much safer than the surgery arm

Placebo controlled trials of surgery are a powerful, feasible way of showing the efficacy of surgical procedures and without them ineffective treatment may continue unchallenged

- as a single-stage procedure: a randomized single-blinded placebo-controlled trial. *Laryngoscope* 2009;119:1621-7.
- 43 Hartigan P. Sclerotherapy for male alcoholic cirrhotic patients who have bled from esophageal varices: results of a randomized, multicenter clinical trial. *Hepatology* 1994;20:618-25.
- 44 Stone GW, Teirstein PS, Rubenstein R, Schmidt D, Whitlow PL, Kosinski EJ, et al. A prospective, multicenter, randomized trial of percutaneous transmyocardial laser revascularization in patients with nonrevascularizable chronic total occlusions. *J Am Coll Cardiol* 2002;39:1581-7.
- 45 Salem M, Rotevatn S, Stavnes S, Brekke M, Vollset SE, Nordrehaug JE. Usefulness and safety of percutaneous myocardial laser revascularization for refractory angina pectoris. *Am J Cardiol* 2004;93:1086-91.
- 46 Geenen JE, Hogan WJ, Dodds WJ, Toouli J, Venu RP. The efficacy of endoscopic sphincterotomy after cholecystectomy in patients with sphincter-of-Oddi dysfunction. *N Engl J Med* 1989;320:82-7.
- 47 Rothstein R, Filipi C, Caca K, Pruitt R, Mergener K, Torquati A, et al. Endoscopic full-thickness plication for the treatment of gastroesophageal reflux disease: a randomized, sham-controlled trial. *Gastroenterology* . 2006;131:704-12.
- 48 Shaheen NJ, Sharma P, Overholt BF, Wolfsen HC, Sampliner RE, Wang KK, et al. Radiofrequency ablation in Barrett's esophagus with dysplasia. *N Engl J Med* 2009;360:2277-88.
- 49 Sutton CJ, Ewen SP, Whitelaw N, Haines P. Prospective, randomized, double-blind, controlled trial of laser laparoscopy in the treatment of pelvic pain associated with minimal, mild, and moderate endometriosis. *Fertil Steril* 1994;62:696-700.
- 50 Castro M, Rubin AS, Lavolette M, Fiterman J, De Andrade Lima M, Shah PL, et al. Effectiveness and safety of bronchial thermoplasty in the treatment of severe asthma. *Am J Respir Crit Care Med* 2010;181:116-24.
- 51 Abbott J, Hawe J, Hunter D, Holmes M, Finn P, Garry R. Laparoscopic excision of endometriosis: a randomized, placebo-controlled trial. *Fertil Steril* 2004;82:878-84.
- 52 Laine L. Multipolar electrocoagulation in the treatment of active upper gastrointestinal tract hemorrhage. A prospective controlled trial. *N Engl J Med* 1987;316:1613-7.
- 53 Swank DJ, Swank-Bordewijk SCG, Hop WCJ, van Erp WFM, Janssen IMC, Bonjer HJ, et al. Laparoscopic adhesiolysis in patients with chronic abdominal pain: a blinded randomized controlled multi-centre trial. *Lancet* 2003;361:1247-51.
- 54 Bradley JD, Heilman DK, Katz BP, Gsell P, Wallick JE, Brandt KD. Tidal irrigation as treatment for knee osteoarthritis: a sham-controlled, randomized, double-blinded evaluation. *Arthritis Rheum* 2002;46:100-8.
- 55 Guyuron B, Reed D, Krieger JS, Davis J, Pashmini N, Amiri S. A placebo-controlled surgical trial of the treatment of migraine headaches. *Plast Reconstr Surg* 2009;124:461-8.
- 56 Schwartz MP, Wellink H, Gooszen HG, Conchillo JM, Samsom M, Smout AJPM. Endoscopic gastroplication for the treatment of gastro-oesophageal reflux disease: a randomised, sham-controlled trial. *Gut* 2007;56:20-8.
- 57 Pauza KJ, Howell S, Dreyfuss P, Pelozo JH, Dawson K, Bogduk N. A randomized, placebo-controlled trial of intradiscal electrothermal therapy for the treatment of discogenic low back pain. *Spine J* 2004;4:27-35.
- 58 Stuck BA, Sauter A, Hormann K, Verse T, Maurer JT. Radiofrequency surgery of the soft palate in the treatment of snoring. A placebo-controlled trial. *Sleep* 2005;28:847-50.
- 59 Montgomery M, Hakanson B, Ljungqvist O, Ahlman B, Thorell A. Twelve months' follow-up after treatment with the EndoCinch endoscopic technique for gastro-oesophageal reflux disease: a randomized, placebo-controlled study. *Scand J Gastroenterol* 1382;41:1382-9.
- 60 Hogan RB, Johnston JH, Long BW, Sones JQ, Ardell Hinton L, Bunge J, et al. A double-blind, randomized, sham-controlled trial of the gastric bubble for obesity. *Gastrointest Endosc* 1989;35:381-5.
- 61 Jarrell J, Mohindra R, Ross S, Taenzer P, Brant R. Laparoscopy and reported pain among patients with endometriosis. *J Obstet Gynaecol Can* 2005;27:477-85.
- 62 Kallmes DF, Comstock BA, Heagerty PJ, Turner JA, Wilson DJ, Diamond TH, et al. A randomized trial of vertebroplasty for osteoporotic spinal fractures. *N Engl J Med* 2009;361:569-79.
- 63 Gross RE, Watts RL, Hauser RA, Bakay RAE, Reichmann H, von Kummer R, et al. Intralesional 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* 2011;10:509-19.
- 64 Buchbinder R, Osborne RH, Ebeling PR, Wark JD, Mitchell P, Wriedit C, et al. A randomized trial of vertebroplasty for painful osteoporotic vertebral fractures. *N Engl J Med* 2009;361:557-68.
- 65 Lee PE, Kung RC, Drutz HP. Periurethral autologous fat injection as treatment for female stress urinary incontinence: a randomized double-blind controlled trial. *J Urol* 2001;165:153-8.
- 66 De Craen AJ, Kaptchuk TJ, Tijssen JG, Kleijnen J. Placebos and placebo effects in medicine: historical overview. *J R Soc Med* 1999;92:511-5.
- 67 Kaptchuk TJ. Powerful placebo: the dark side of the randomised controlled trial. *Lancet* 1998;351:1722-5.
- 68 Beecher HK. The powerful placebo. *JAMA* 1955;159:1602-6.
- 69 Haidich A-B, Birtsou C, Dardavessis T, Tirodimos I, Arvanitidou M. The quality of safety reporting in trials is still suboptimal: survey of major general medical journals. *J Clin Epidemiol* 2011;64:124-35.
- 70 Ioannidis JPA, Evans SJW, Gøtzsche PC, O'Neill RT, Altman DG, Schulz K, et al. Better reporting of harms in randomized trials: an extension of the CONSORT statement. *Ann Intern Med* 2004;141:781-8.
- 71 Fassler M, Meissner K, Schneider A, Linde K. Frequency and circumstances of placebo use in clinical practice - a systematic review of empirical studies. *BMC Med* 2010;8:15.
- 72 Hull SC, Colloca L, Avins A, Gordon NP, Somkin CP, Kaptchuk TJ, et al. Patients' attitudes about the use of placebo treatments: telephone survey. *BMJ* 2013;347:t2757.
- 73 Kass NE, Faden RR, Goodman SN, Pronovost P, Tunis S, Beauchamp TL. The research-treatment distinction: a problematic approach for determining which activities should have ethical oversight. *Hastings Cent Rep* 2013;43(s1):S4-15.
- 74 Birkmeyer JD, Reames BN, McCulloch P, Carr AJ, Campbell WB, Wennberg JE. Understanding of regional variation in the use of surgery. *Lancet* 2013;382:1121-9.
- 75 Moseley JB, O'Malley K, Petersen NJ, Menke TJ, Brody BA, Kuykendall DH, et al. A controlled trial of arthroscopic surgery for osteoarthritis of the knee. *N Engl J Med* 2002;347:81-8.
- 76 Savulescu J, Hope T. Ethics of research In: J Skrupski, ed. The Routledge companion to ethics. Routledge, 2010:781-95.
- 77 Savulescu J. Harm, ethics committees and the gene therapy death. *J Med Ethics* 2001;27:148-50.
- 78 Savulescu J. Commentary: safety of participants in non-therapeutic research must be ensured. *BMJ* 1998;316:891-2; discussion 3-4.
- 79 Walsh BT, Seidman SN, Sysko R, Gould M. Placebo response in studies of major depression. *JAMA* 2002;287:1840-7.
- 80 Waber RL, Shiv B, Carmon Z, Ariely D. Commercial Features of Placebo and Therapeutic Efficacy. *JAMA* . 2008 March 5, 2008;299(9):1016-7.
- 81 Cobb LA, Thomas GI, Dillard DH, Merendino KA, Bruce RA. An evaluation of internal-mammary-artery ligation by a double-blind technic. *N Engl J Med* 1959;260:1115-8.
- 82 Felson DT, Buckwalter J. Débridement and lavage for osteoarthritis of the knee. *N Engl J Med* 2002;347:132-3.
- 83 Fischbach GD, McKhann GM. Cell therapy for Parkinson's disease. *N Engl J Med* 2001;344:763-5.
- 84 Grossman RG. Cell transplantation in Parkinson's disease: implications for human clinical trials. *Neurosurgery* 2001;49:580-2.
- 85 Freeman TB, Vawter DE, Leaverton PE, Godbold JH, Hauser RA, Goetz CG, et al. Use of placebo surgery in controlled trials of a cellular-based therapy for Parkinson's disease. *N Engl J Med* 1999;341:988-92.
- 86 Weinstein JN. Balancing science and informed choice in decisions about vertebroplasty. *N Engl J Med* 2009;361:619-21.

Accepted: 5 May 2014

Cite this as: *BMJ* 2014;348:g3253

This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 3.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: <http://creativecommons.org/licenses/by-nc/3.0/>.

Tables

Table 1 | Characteristics of reviewed trials reporting serious adverse events

Reference	Condition	Intervention	Ex	Primary outcomes	No	Improvement	Superior	Significant*	Serious adverse event††			
									Ser	Rel	Ser	Rel
Silverberg et al 2008	Alzheimer's disease	Ventriculoperitoneal shunt	Y	Clinical assessment	230	N	ND	N	Y	Y	Y	Y
Kallmes et al 2009	Osteoporotic vertebral fractures	Percutaneous vertebroplasty	Y	Pain, function	131	Both study arms	ND	N	Y	Y‡	Y	L‡
Castro et al 2010	Severe asthma	Radiofrequency bronchial thermoplasty	N	Quality of life	297	Both study arms	Surgical	Y/N	Y	Y‡	Y	L‡
Gross et al 2011	Parkinson's disease	Human cells transplantation	Y	Clinical assessment	71	Both study arms	ND	N	Y	Y	Y	U
Freed et al 2001	Parkinson's disease	Dopamine neurons transplantation	Y	Perception of change	40	N	ND	N	Y	L	Y	U‡
Stone et al 2002	Coronary disease	PMLR	N	Function	141	Both study arms	ND	Y	Y	L	Y	L‡
Salem et al 2004	Coronary disease	PMLR	N	Clinical assessment	82	Both study arms	Surgical	Y	Y	L	Y	L
Abbott et al 2004	Endometriosis	Laparoscopy+excision	N	Pain, quality of life	39	Both study arms	Surgical	Y/N	Y	L	Y	U‡
Lee et al 2001	Urinary stress incontinence	Autologous fat injection	N	Symptoms, function, objective test	68	Both study arms	ND	N	Y	Y	L	Y
Dowson et al 2008	Migraine	Patent foramen ovale closure	Y	Pain	147	N	ND	N	Y	Y	L	L
Hartigan et al 1994	Oesophageal varices§	Endoscopy+sclerotherapy	N	Bleeding, death	253	NR	Surgical	N	Y	†	Y	†
Fleischer et al 1985	Bleeding esophageal varices§	Endoscopy+laser	N	Bleeding	20	Surgical	Surgical	Y/N	Y	Y†	Y	Y†
Fullarton et al 1989	Bleeding peptic ulcers§	Endoscopy+heater probe	N	Bleeding	43	Surgical	Surgical	Y/N	Y	Y†	Y	†
Freitas et al 1985	Bleeding peptic ulcers§	Endoscopy+electrocoagulation	N	Bleeding, death	78	Both study arms	Surgical	N	Y	†	Y	†
MacLeod et al 1983	Bleeding peptic ulcers§	Endoscopy+laser	N	Bleeding, death	45	Both study arms	Surgical	N	Y	†	Y	†
Laine et al 1987	Upper gastrointestinal tract bleeding§	Endoscopy+electrocoagulation	N	Bleeding, death	44	Both study arms	Surgical	Y/N	Y	†	Y	†
Leon et al 2005	Coronary disease	PMLR	N	Function	298	Both study arms	ND	N	Y	L‡	Y	L‡
Corley et al 2003	GERD	Radiofrequency surgery	N	Symptoms, quality of life	64	Both study arms	Surgical	N	U	L	Y	L‡
Geenen et al 1989	Sphincter of Oddi dysfunction	Endoscopic sphincterotomy+ERCP	N	Symptoms, objective test	47	Both study arms	Surgical	Y	‡	Y	‡	Y
Buchbinder et al 2009	Osteoporotic vertebral fractures	Percutaneous vertebroplasty	Y	Pain	78	Both study arms	ND	N	‡	L	‡	L
Thompson et al 2013	Obesity	Transoral outlet reduction	N	Objective test	77	Surgical	Surgical	Y	‡	U	‡	U
Pauza et al 2004	Discogenic low back pain	Intradiscal electrothermal therapy	N	Pain, function, quality of life	64	Both study arms	Surgical	Y/N	‡	U	‡	U
Swank et al 2003	Chronic abdominal pain	Laparoscopy+adhesiolysis	N	Pain, quality of life, drugs	100	Both study arms	ND	Y/N	Y	Y	N	N

Table 1 (continued)

Reference	Condition	Intervention	Ex	Primary outcomes	No	Improvement	Superior	Significant*	Serious adverse event††			
									Ser	Rel	Ser	Rel
Deviere et al 2005	GERD	Endoscopy+copolymer	Y	Drugs	64	Both study arms	Surgical	Y	Y	Y‡	N	N
Rothstein et al 2007	GERD	Endoscopy+plication	N	Quality of life	159	Both study arms	Surgical	Y	Y	L	N	N
Shaheen et al 2009	Barrett's oesophagus	Radiofrequency ablation	N	Clinical assessment	127	Both study arms	Surgical	Y	Y	L‡	N	N
Olanow et al 2003	Parkinson's disease	Fetal tissue transplantation	Y	Clinical assessment	34	N	ND	N	Y	U	N	N

Ex=exogenous material used; Imp=improvement as originally reported by the trial's authors, Sup=superiority of one arm over other as originally reported by trials' authors; No=number of patients randomised into trial; Sig=statistical significance of one arm over the other as calculated by us; Ser=whether adverse events were serious; Rel=whether serious adverse events were relevant to the procedure; ND=no difference; Y=yes; N=no; Y/N=yes or no depending on outcome; L=likely to be serious or related to the procedure; U=unlikely to be serious or related to procedure; PMLR=percutaneous myocardial laser revascularisation; NR=not reported; GERD=gastro-oesophageal reflux disease; ERCP=endoscopic retrograde cholangiopancreatography.

*Statistical significance of effect as calculated by us using reported values.

†Serious adverse events were direct consequence of investigated conditions not interventions.

††Graded according to severity: definitely serious (Y), that is, reported as serious adverse events by trial authors; likely to be serious (L), that is, not specifically reported as serious but resulting in death or admission to hospital; and unlikely (U) to be serious or reported as minor adverse events. Adverse events were also graded according to causality as definitely related to the procedure (Y), that is, reported by trials' authors as such; likely to be related to the procedure (L) as classified by us on basis of details reported by trials' authors; and unlikely to be related to the procedure or reported by trials' authors as unrelated to the procedure (U).

‡Trials' authors did not specify arm in which adverse events occurred.

§Trials on upper gastrointestinal ulcers or varices in which bleeding or death were the primary outcome.

Table 2 | Characteristics of reviewed trials not reporting serious adverse events

Reference	Condition	Intervention	Ex	Primary outcomes	No	Impr	Sup	Significant*	Serious adverse event			
									Ser	Rel	Ser	Rel
Guyron et al 2009	Migraine	"Deactivation" of trigger points	N	Pain	76	Both study arms	Surgical	Y/N	N	N	N	N
Scolapio et al 2001	Dysphagia	Endoscopy+B dilation	N	Function	86	Both study arms	ND	N	N	N	N	N
Bajbouj et al 2009	Globus sensation	Endoscopy+ablation of inlet patches	N	Symptoms	21	Surgical	Surgical	Y	N	N	N	N
Arts et al 2012	GERD	Radiofrequency surgery	N	Objective test	22	Surgical	Surgical	Y	N	N	N	N
Montgomery et al 2006	GERD	Endoscopic gastroplasty	N	Objective test, quality of life, drugs	46	Both study arms	ND	NR	N	N	N	N
Schwartz et al 2007	GERD	Endoscopic gastroplasty	N	Symptoms, drugs	60	Both study arms	Surgical	Y/N	N	N	N	N
Geliebter et al 1990	Obesity	Endoscopy+B	B	Objective test	10	Both study arms	ND	N	N	N	N	N
Hogan et al 1989	Obesity	Endoscopy+B	B	Objective test	59	Both study arms	ND	NR	N	N	N	N
Lindor et al 1987	Obesity	Endoscopy+B	B	Objective test	22	Both study arms	ND	N	N	N	N	N
Mathus-Vliegen et al 1990	Obesity	Endoscopy+B	B	Objective test	28	Both study arms	ND	N	N	N	N	N
Meshkinpour et al 1988	Obesity	Endoscopy+B	B	Objective test	23	Both study arms	ND	N	N	N	N	N
Genco et al 2006	Obesity	Endoscopy+B	B	Objective test, death	32	Both study arms	Surgical	Y	N	N	N	N
Bradley et al 2002	Osteoarthritis	Tidal irrigation	N	Pain, function	180	Both study arms	ND	Y/N	N	N	N	N
Moseley et al 2002	Osteoarthritis	Arthroscopic debridement or lavage	N	Pain	180	Both study arms	ND	N	N	N	N	N
Freeman et al 2005	Discogenic low back pain	Intradiscal electrothermal therapy	N	Pain, quality of life, clinical assessment	57	N	ND	N	N	N	N	N
Baeck et al 2009	Sleep apnoea	Radiofrequency surgery	N	Objective test, symptoms, quality of life	32	N	ND	NR	N	N	N	N
Gillespie et al 2011	Sleep apnoea	Palatal implants	Y	Objective test	51	Placebo	ND	N	N	N	N	N
Maurer et al 2012	Sleep apnoea	Palatal implants	Y	Objective test	22	Both study arms	ND	N	N	N	N	N
Friedman et al 2008	Sleep apnoea	Palatal implants	Y	Objective test	62	Surgical	Surgical	Y	N	N	N	N
Steward et al 2008	Sleep apnoea	Palatal implants	Y	Objective test, function	100	Both study arms	Surgical	N	N	N	N	N
van Schie et al 2000	Diabetic foot	Silicone injection	Y	Clinical assessment	28	Surgical	Surgical	NR	N	N	N	N
Davys et al 2005	Plantar callosities in RA	Debridement of callosity	N	Pain	38	Both study arms	ND	NR	N	N	N	N
Sutton et al 1994	Endometriosis	Laparoscopy+laser	N	Pain, symptoms	74	Both study arms	Surgical	Y	N	N	N	N
Jarrell et al 2005	Endometriosis	Laparoscopy+excision	N	Pain	29	Both study arms	ND	NR	NR	NR	NR	NR
Thomsen et al 1981	Meniere's disease	Decompression	N	Symptoms, clinical assessment	30	Both study arms	ND	N	NR	NR	NR	NR

RESEARCH

Table 2 (continued)

Reference	Condition	Intervention	Ex	Primary outcomes	No	Impr	Sup	Significant*	Serious adverse event			
									Surgery		Placebo	
									Ser	Rel	Ser	Rel
Stuck et al 2005	Snoring	Radiofrequency surgery of soft palate	N	Symptoms	26	Both study arms	Surgical	Y/N	NR	NR	NR	NR

Ex=exogenous material used; Impr=improvement as originally reported by the trial’s authors, Sup=superiority of one arm over other as originally reported by trials’ authors; No=number of patients randomised into trial; Sig=statistical significance of one arm the over other as calculated by us; Ser=whether adverse events were serious; Rel=whether serious adverse events were related to the procedure; ND=no difference; Y=yes; N= no; Y/N=yes or no depending on outcome; L=likely to be serious or related to the procedure; B=balloon; U=unlikely to be serious or related to the procedure; GERD=gastro-oesophageal reflux disease; NR=not reported; RA=rheumatoid arthritis.

*Statistical significance of effect as calculated by us using reported values.

Table 3| Outcomes as reported by trials' authors

Intervention	Superiority		Total
	No difference	Surgery	
Surgery and placebo	20	18	38
Surgery only	0	7	7
Placebo only	1	0	1
Neither surgery nor placebo	6	0	6
NA	0	1	1
Total	27	26	53

NA=outcome was prevention of death so trial cannot be interpreted in terms of improvement.

Figures

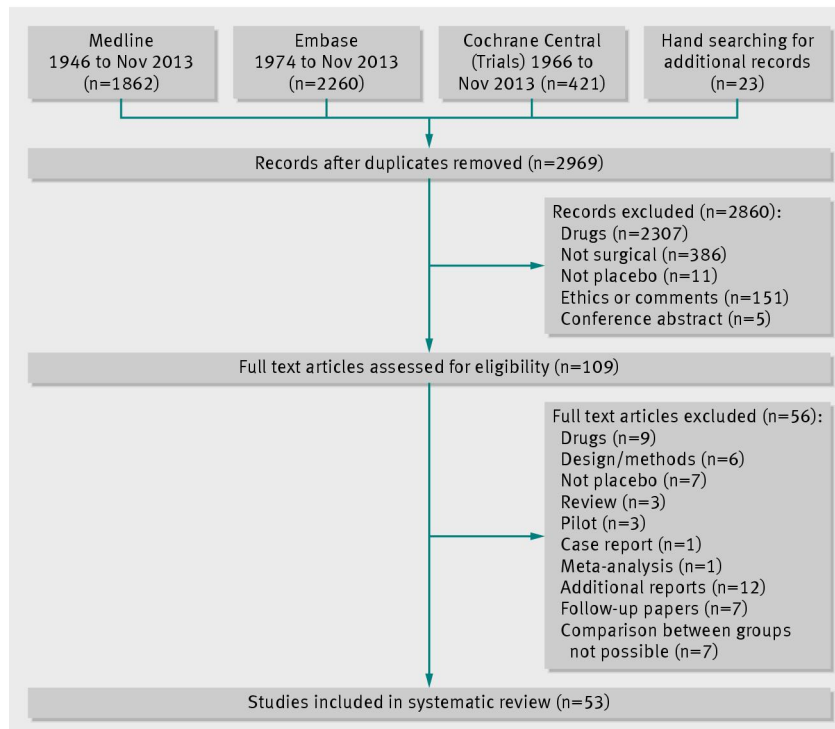


Fig 1 Flow chart of study identification, listing first reason for exclusion during review process

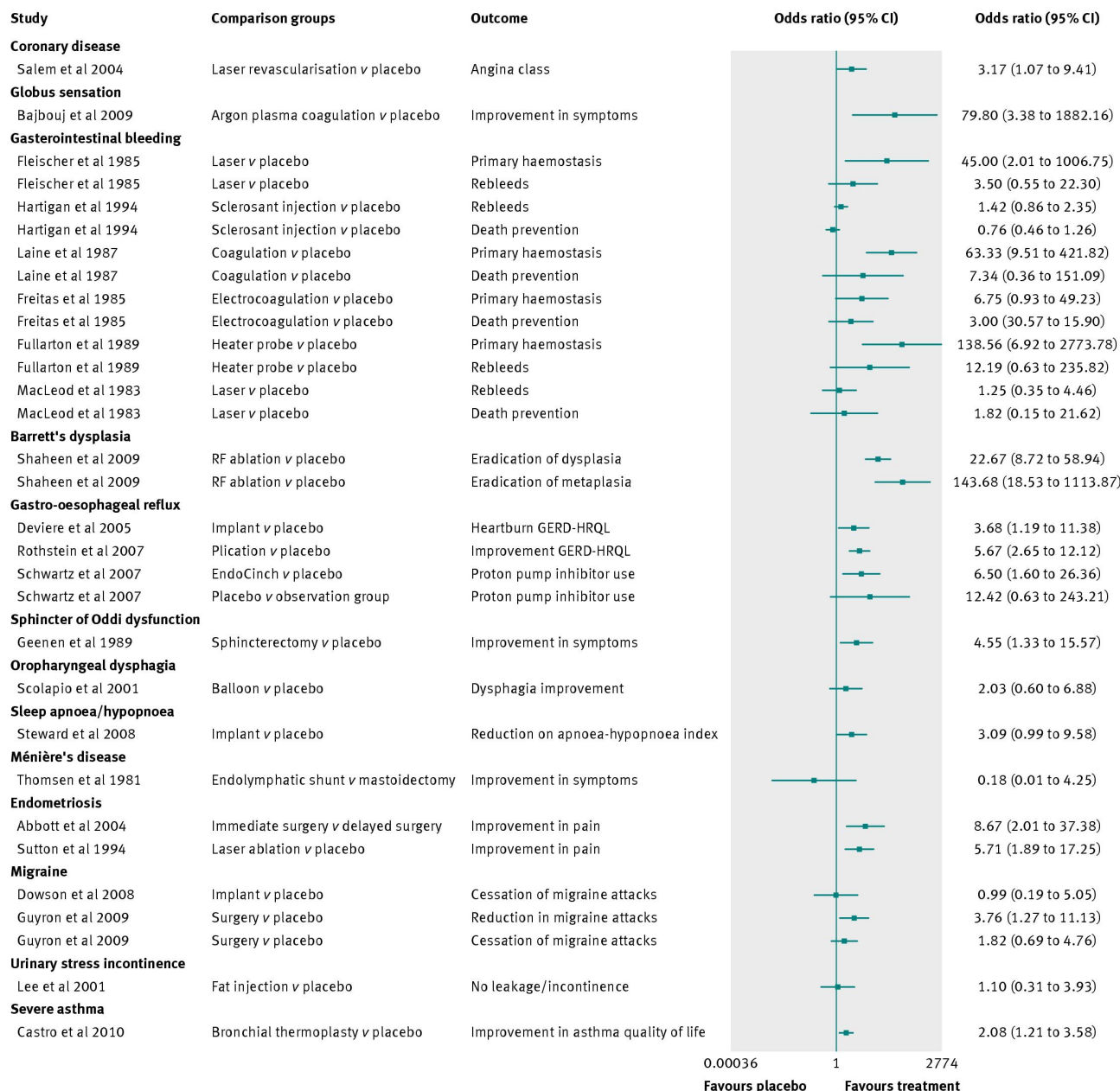


Fig 2 Forest plot of studies with binary outcome measures showing magnitude of effect (odds ratios) in active group compared with placebo group. GERD-HRQL=gastro-oesophageal reflux disease health related quality of life

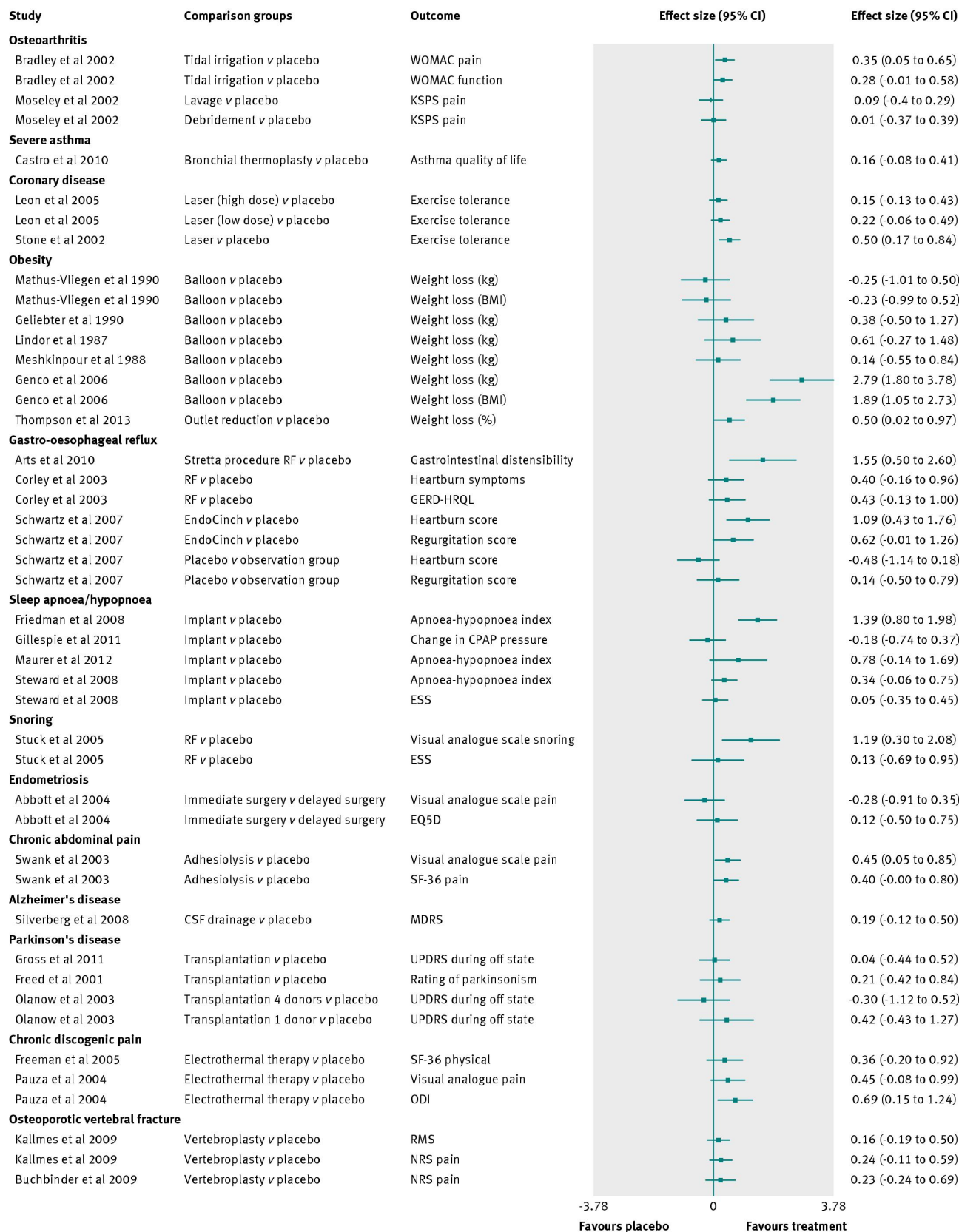


Fig 3 Forest plot of studies with continuous outcome measures showing magnitude of effect (effect sizes) in active group compared with placebo group. Outcome values in Stone and colleagues trial were not normally distributed; therefore, the effect size does not represent the true difference. WOMAC=Western Ontario and McMaster Universities arthritis index; KSPS=knee specific pain scale; QoL=quality of life; BMI=body mass index; GERD-HRQL=gastro-oesophageal reflux disease health related quality of life; NRS=numerical rating scale; RF=radiofrequency; SF-36=short form (36) health survey; CSF=cerebrospinal fluid; UPDRS=unified Parkinson's disease rating scale; EQ5D=EuroQol Group health questionnaire; RMS=modified Roland-Morris scale; ODI=Oswestry disability index; ESS=Epworth sleepiness scale; MDRS=Mattis dementia rating scale