BMJ Open

To what extent are surgery and invasive procedures effective beyond a placebo response? A systematic review with meta-analysis of randomised, sham controlled trials

Wayne B Jonas,[1] Cindy Crawford,[1] Luana Colloca,[2,3] Ted J Kaptchuk,[4] Bruce Moseley,[5] Franklin G Miller,[6] Levente Kriston,[7] Klaus Linde,[8] Karin Meissner[9]

To cite: Jonas WB, Crawford C, Colloca L, *et al*. To what extent are surgery and invasive procedures effective beyond a placebo response? A systematic review with meta-analysis of randomised, sham controlled trials. *BMJ Open* 2015;**5**: e009655. doi:10.1136/bmjopen-2015-009655

► Prepublication history and additional material is available. To view please visit the journal (http://dx.doi.org/10.1136/bmjopen-2015-009655).

Received 10 August 2015
Revised 28 October 2015
Accepted 5 November 2015

For numbered affiliations see end of article.

Correspondence to
Dr Wayne B Jonas;
wjonas@siib.org

ABSTRACT

Objectives: To assess the quantity and quality of randomised, sham-controlled studies of surgery and invasive procedures and estimate the treatment-specific and non-specific effects of those procedures.

Design: Systematic review and meta-analysis.

Data sources: We searched PubMed, EMBASE, CINAHL, CENTRAL (Cochrane Library), PILOTS, PsycInfo, DoD Biomedical Research, clinicaltrials.gov, NLM catalog and NIH Grantee Publications Database from their inception through January 2015.

Study selection: We included randomised controlled trials of surgery and invasive procedures that penetrated the skin or an orifice and had a parallel sham procedure for comparison.

Data extraction and analysis: Three authors independently extracted data and assessed risk of bias. Studies reporting continuous outcomes were pooled and the standardised mean difference (SMD) with 95% CIs was calculated using a random effects model for difference between true and sham groups.

Results: 55 studies (3574 patients) were identified meeting inclusion criteria; 39 provided sufficient data for inclusion in the main analysis (2902 patients). The overall SMD of the continuous primary outcome between treatment/sham-control groups was 0.34 (95% CI 0.20 to 0.49; p<0.00001; I^2=67%). The SMD for surgery versus sham surgery was non-significant for pain-related conditions (n=15, SMD=0.13, p=0.08), marginally significant for studies on weight loss (n=10, SMD=0.52, p=0.05) and significant for gastroesophageal reflux disorder (GERD) studies (n=5, SMD=0.65, p<0.001) and for other conditions (n=8, SMD=0.44, p=0.004). Mean improvement in sham groups relative to active treatment was larger in pain-related conditions (78%) and obesity (71%) than in GERD (57%) and other conditions (57%), and was smaller in classical-surgery trials (21%) than in endoscopic trials (73%) and those using percutaneous procedures (64%).

Conclusions: The non-specific effects of surgery and other invasive procedures are generally large. Particularly in the field of pain-related conditions, more evidence from randomised placebo-controlled trials is needed to avoid continuation of ineffective treatments.

Strengths and limitations of this study

- This is the first systematic review using a meta-analysis approach to estimate both specific and non-specific components in sham-controlled surgical trials, and to what extent those effects differ among conditions and procedures.
- All sensitivity analyses showed similar results as the main analysis, except one, namely the sensitivity analysis for large studies (≥100 patients), which showed a smaller non-significant effect size.
- Our results have implications for clinical research and practice by arguing against the continued use of ineffective invasive treatments, especially in the field of chronic pain.
- One limitation might be that the conclusions from our meta-analysis are restricted to available published data on surgical interventions that have been tested in sham-controlled clinical trials.

INTRODUCTION

Surgery and other invasive procedures such as endoscopy and percutaneous procedures are widely used in medicine but their specific efficacy and risk-benefit profile are rarely assessed in rigorous and systematic ways. The development of minimally invasive procedures has expanded the use of such interventions for treating a variety of conditions such as low-back pain,[1] arthritis,[2] endometriosis,[3] Parkinson's disease,[4] gastro-oesophageal reflux[5] and obesity.[6]

BMJ Open: first published as 10.1136/bmjopen-2015-009655 on 11 December 2015. Downloaded from http://bmjopen.bmj.com/ on September 26, 2024 by guest. Protected by copyright.

Rarely are these procedures evaluated using rigorous research designs involving randomisation, allocation concealment and blinding or placebo controls, which are considered gold standards for medical interventions. In the absence of controls for common sources of bias, studies on these procedures may give a false impression of their true efficacy. Is it possible to test invasive procedures using rigorous methods? Blinding of outcome assessment is challenging since mimicking a complex, invasive procedure such as surgery, or insertion of a scope or a needle, requires an elaborate sham procedure. Moreover, there is significant controversy over the ethics of using sham procedures, even with carefully informed patients, further restricting the number of such studies being carried out.[7 8] However, can we justify widespread use of these procedures without rigorous testing?

The use of blinded, sham procedures permits rigorous assessment of treatment efficacy by comparing the outcome in the treatment and sham groups. Specifically, sham procedures control for a variety of observed outcomes in the sham group that are distinct from the specific efficacy of the surgery or invasive procedure under investigation. These 'non-specific' outcomes include placebo responses (also sometimes called placebo effects), which we define here as the observed outcome changes in the sham groups. These changes are due to the natural history of the patient's condition or regression to the mean and a response to the ritual of medical treatments. Such rituals include the type of procedure (pill, needle, knife or touch), the status, authority and communication style of the provider, the setting and context of the treatment and the patient's and practitioners expectation about the outcome.[9]

Yet, invasive procedures are thought to incorporate many factors that may contribute to the placebo responses including use of a hospital-like setting; multiple, authoritative providers; frequent and repeated suggestions about expected outcomes; a physical invasion of the body; and an elaborate ritual of treatment delivery and recovery.[10] Thus, one would expect a significant contribution from surgical ritual and other non-specific factors to the observed outcomes during invasive procedures in clinical practice and in randomised trials without sham control groups. Several high profile studies support this hypothesis in which sham procedures involving only superficial anaesthesia were compared to the more invasive true procedure.[11–13] For example, Moseley *et al*[11] reported no greater pain improvement in patients with osteoarthritis of the knee that underwent arthroscopic knee surgery compared to a sham procedure in which a cut was made over the knee without introducing the arthroscope. Two more recent controlled studies of vertebroplasty for painful osteoporotic vertebral fractures reported similar degrees of pain relief from sham procedures involving only superficial anaesthesia compared to the more invasive active procedures.[12 13] In contrast, a systematic review comparing surgical with non-surgical treatments for painful osteoporotic vertebral fractures came to the conclusion that vertebroplasty and kyphoplasty are superior to non-surgical treatments.[14] Since invasive interventions frequently go along with larger non-specific effects than non-invasive treatments[15 16] surgical trials that do not include a sham surgery arm may give biased results. Thus, the efficacy of invasive procedures, for example, for chronic pain conditions, remains controversial.[17] In addition, many invasive procedures involve the risk of anaesthesia and high cost.[17] Therefore, it is important to estimate to what degree the observed outcomes from invasive procedures are due to specific efficacy of the treatments or to other factors.

To better understand these issues we conducted a systematic review and meta-analysis of studies on surgery and invasive procedures in which a parallel sham procedure was included for comparison. Our study aims were to: (1) assess the quantity and quality of such studies; (2) estimate the magnitude of specific effects over sham procedures; and, (3) estimate the contribution of the surgical ritual and other non-specific factors to outcomes from these procedures.

METHODS

Identification of studies

The following online databases were searched from their inception through January 2015: PubMed, EMBASE, CINAHL, CENTRAL (Cochrane Library), PILOTS, PsycInfo, DoD Biomedical Research, clinicaltrials.gov, NLM catalog, as well as NIH Grantee Publications Database. We used as our initial search terms: 'Diagnostic Techniques, Surgical' OR 'Orthopedic Procedures' OR 'Specialties, Surgical' OR 'Surgical Procedures, Operative' OR 'surgery' (Subheading) or surgery) AND ('Placebos' OR 'Placebo Effect' or sham surg* or placebo surg* or mock surg* or simulated surg* or placebo proc* or sham proc* or mock proc* or simulated proc*). We restricted our search to humans and randomised controlled trials. Variations of these search terms were made for MESH terms, where necessary, and are available on request from the first author.

The 'Grey literature' was searched by looking for relevant dissertations, conference proceedings, Google Scholar and searching the internet using the keyword scheme as well as searching all relevant reference lists of identified articles and related reviews. We also contacted and consulted with leading experts in the fields of surgery and placebo, and shared databases that these experts have collected over the years relating to placebo to make sure we captured all the relevant literature.

Eligibility criteria

Studies were included in the systematic review if they: (1) were randomised controlled trials; (2) involved a population for which there was a symptom-driven medical condition for which an invasive procedure or

BMJ Open: first published as 10.1136/bmjopen-2015-009655 on 11 December 2015. Downloaded from http://bmjopen.bmj.com/ on September 26, 2024 by guest. Protected by copyright.

classical surgery as defined below was being performed; and (3) had a comparison group that used a sham procedure to mimic the real procedure.

Classical surgery was defined as a procedure that followed the typical surgical experience that uses preoperative preparation, anaesthesia, an incisional trauma (usually through muscle and fascia and into the peritoneum) and a postoperative recovery process. Invasive procedures were defined as when an instrument was inserted into the body (either endoscopically or percutaneously) for the purpose of manipulating tissue or changing anatomy. In all cases we selected studies where when these procedures were compared to a sham procedure that used the same surgical or invasive procedure, instrument and ritual, but eliminated the hypothesised active component of tissue manipulation. We excluded studies in which the procedure was used simply as a delivery mechanism for another ongoing active treatment such as a pacemaker, brain or cardiac stimulation, or delivery of a drug or biological product. Studies where an invasive procedure was implemented for prevention of a medical condition or there was no symptom-driven condition were also excluded.

Four investigators (CC, LC, KL and KM) screened titles and abstracts for relevance in two phases based on the inclusion criteria: phase one eliminated all clearly irrelevant studies, phase two applied all inclusion/exclusion criteria listed above for the remaining studies. Any disagreements about including a study were resolved through discussion and consensus, and approved by the first author (WJ). All reviewers were fully trained in systematic review methodology. At least two reviewers had to review each citation in order for it to progress to the next phase of the review. A Cohen's κ on agreement was attained for both phases above 88%.

Quality assessment and data extraction

The methodological quality of the individual studies (sequence generation, allocation concealment, was assessed independently by three reviewers using the Cochrane Risk of Bias (ROB) tool.[18] Descriptive data was independently extracted on the following items: population; condition for which surgery was performed; sample (population) entered; dropout rate; informed consent details; whether a power calculation was performed and achieved; intervention and sham procedure used; primary and secondary outcomes and the statistical data associated with these; whether expectation was reported; author conclusions; adverse events reported; funding source, and reviewer comments. We also extracted from each study, if available, a continuous and a dichotomous main outcome at two time points (intermediate and late), and a continuous and a dichotomous pain outcome (when applicable). The most important outcome measure (miOM) was defined as either: (1) the primary main outcome measure (pMOM) at a time point as predefined in the trial; or (if not 1), (2) the only major outcome of a trial at the latest available time point; or (if neither 1 nor 2), (3) the clearly most relevant outcome determined by two independent reviewers at the latest available time point. Secondary outcomes were intermediate time points of the most important outcome measure; pain outcomes at the latest available time point; or, pain outcomes at the intermediate time point. All discrepancies were tracked by the review manager and were resolved by consensus and discussions during team meetings. Data were entered into a web-based, secure, systematic review management programme called Mobius Analytics SRS (Mobius Analytics Inc, Ottawa, Ontario, Canada).

Data synthesis and analysis

According to our analysis plan, the meta-analyses focused on continuous outcomes. The primary analysis was based on trials reporting a most important continuous outcome measure in sufficient detail to be included in the meta-analysis. Secondary analyses were based on trials reporting (1) a continuous outcome measure at an intermediate time point, (2) a pain measure at a late time point, (3) a pain measure at an intermediate time point. Trials reporting only a dichotomous outcome measure (responder data) are noted in online supplementary table 1, and a sensitivity analysis was computed for these outcomes (see below).

Within-group and between-group effect sizes were based on Cohen's[19] d for change within one group, and Cohen's d for between-group effect measures, respectively, correcting for small-sample bias.[20] In order to keep the effect size framework coherent for within-group and between-group designs, change from baseline was used throughout. When SD was not reported, it was calculated from pre-SD and post-SD,[21] using r=0.5 for the product-moment correlation between pre and post measures.

Analyses of continuous data were performed with the generic inverse variance module of the Cochrane Collaboration's Review Manager software (V.5.1), using standardised mean difference (SMD) as the effect size measure. As we expected heterogeneity, a random effects model was used. Within-group effect sizes were pooled in such a way that positive values indicate improvement, while positive values of between-group effect sizes indicate superiority (more pronounced improvement) of the intervention group over the control (sham) group. To estimate the relative contribution of non-specific outcomes to treatment effects, the per cent ratio of the pooled within-group treatment effects in the sham and the treatment groups was calculated. We used Cochrane's Q test and calculated I^2 to examine statistical heterogeneity, with low, moderate and high I^2 values of 25%, 50%, and 75%.[22] Egger's test was used to assess funnel plot asymmetry.[23] A p value of less than 0.05 was set as the level of significance.[24 25]

Subgroup analyses were performed according to predefined categories of target diseases and types of

BMJ Open: first published as 10.1136/bmjopen-2015-009655 on 11 December 2015. Downloaded from http://bmjopen.bmj.com/ on September 26, 2024 by guest. Protected by copyright.

surgery. To check the robustness of results, we performed sensitivity analyses with four criteria: (1) studies specifying a primary main outcome measure (pMOM); (2) imputing 0.3 and 0.7 for pre–post correlation coefficient r, when missing; (3) studies with total sample sizes ≥100; and, (4) studies with low risk of allocation concealment. An additional sensitivity analysis was performed for dichotomous outcomes of 12 studies that provided no continuous outcome (see online supplementary figure 1).

RESULTS

Eligible studies

Our search identified a total of 7360 citations. After excluding clearly irrelevant references the full text of 113 publications were obtained. Of these, 46 were excluded, mainly for not including an instrumental or surgical intervention or a sham procedure as defined above. A total of 55 studies (in 67 publications) involving a total of 3574 enrolled patients met our inclusion criteria for systematic review (figure 1).[26]

Characteristics and quality of included studies

Characteristics of the included studies are summarised in online supplementary table 1. About half (25) of the studies were carried out on pain-related conditions with back pain (7) being the most frequent[11 12 27–31] followed by arthritis (4),[13 32–34] angina from coronary artery disease (4),[35–39]) abdominal pain (3),[40–42] endometriosis (3),[43–47] cholia (2)[48 49] and migraine (2).[50 51]) The most frequently studied non-pain condition was obesity, especially when using balloon insertion (11).[52–62]) Other conditions that had more than one study included gastroesophageal reflux disease (GERD) (5),[63–67] Parkinson's Disease (2),[68–74] sleep apnoea (2),[75 76] dry eye (2)[77 78] and asthma (2).[79–81] Some other conditions were also studied (see online supplementary table 1).[80–90] Many (22) of the studies involved endoscopic or percutaneous procedures in which tissue was removed or altered or a material (eg, dye, cement, balloon) was inserted.[11–13 28 31 34 38 40–43 52 54 56 61 63 65 67 77–79 90] Some of these procedures used a catheter to reach an internal organ (such as the heart or gall bladder) or a needle to inject a material or cell (often into the lumbar spine or

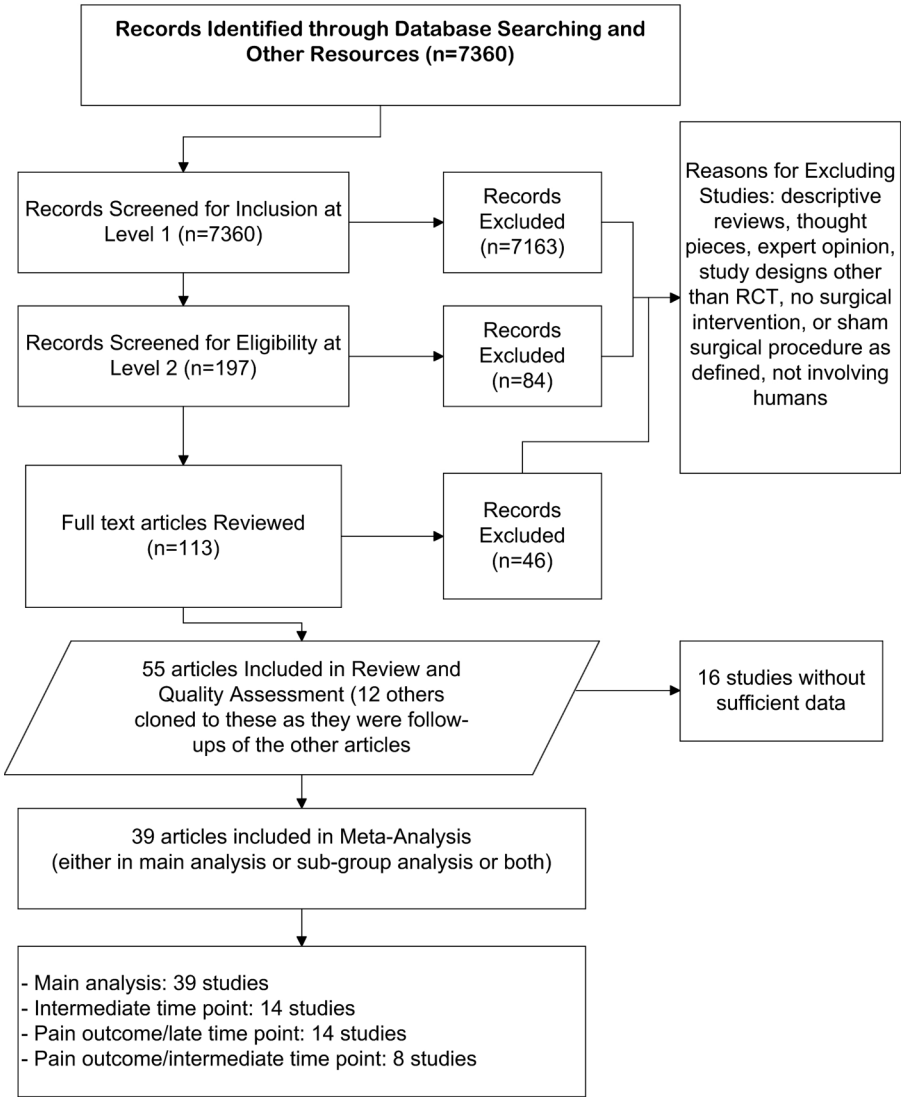


Figure 1 Flow chart of included studies. RCT, randomised controlled trial.

BMJ Open: first published as 10.1136/bmjopen-2015-009655 on 11 December 2015. Downloaded from http://bmjopen.bmj.com/ on September 26, 2024 by guest. Protected by copyright.

brain).[27 29 30 32 53 55 57 59 60 62 64 66 85 89] Five studies evaluated more classical surgical procedures in which the body was opened with a scalpel or drill.[50 51 74–76]

In most studies, blinding was achieved using elaborate sham procedures. Those mimicking classical surgical procedures usually cut the body, leaving a scar but causing less damage than the real surgery. Sham percutaneous and endoscopic procedures often involved superficial insertion of a needle or a scope. For example, in the Parkinson's studies on surgical interventions on the brain, sham procedures involved placing burr holes without penetration of the skull.[68–74] Sham surgery for endometriosis would often involve 'diagnostic laparoscopy' with no internal tissue destruction. Sham balloon insertion for obesity treatment usually involved inserting the balloon but not inflating it.[52–62]

Overall, the risk of bias was low in these studies, with some exceptions. Of the 55 studies (67 publications) included in the systematic review, 34 studies (62%) reported an adequate method for generating the allocation sequence, however only 23 (42%) had adequate concealment of allocation. Blinding of the patients and outcome assessors was adequate in 48 (87%) studies and incomplete data was adequately addressed in 52 (95%). Fifty-two (95%) of the studies were free from suggestion of selective outcome reporting and 53 studies were judged to be free of other sources of bias.

Overall analyses

Thirty-nine studies (2902 patients) with continuous data were included in the main analysis. The overall effect of surgery compared to sham surgery was highly significant (SMD 0.34, 95% CI 0.20 to 0.49; p<0.00001), while heterogeneity was large (I^2=67%, p<0.00001). Excluding one outlier[52] reduced I^2 to 57% (SMD, 0.30, 95% CI 0.17 to 0.43; p<0.00001), indicating moderate heterogeneity. Sensitivity analyses provided comparable effect sizes (figure 2), except for studies with overall sample sizes of 100 participants or more, for which the SMD was non-significant at 0.15 (n=10; 95% CI −0.02 to 0.32; p=0.09; I^2=66%). Inspection of the funnel plot suggests the presence of biases in the meta-analysis, such as small study bias or publication bias (figure 3). Asymmetry in the funnel plot was confirmed by the Egger's test (asymmetry coefficient 1.7, p=0.017).

Non-significant SMD were found when combining available data for the most important continuous outcome measure at an intermediate time point (n=14; SMD 0.12, 95% CI −0.05 to 0.29; p=0.17; I^2=54%) as well as for specific pain outcomes at a late (n=14; SMD 0.12, 95% CI −0.03 to 0.27; p=0.11; I^2=29%;) or an intermediate time point (n=8; SMD 0.07, 95% CI −0.06 to 0.20; p=0.31; I^2=0%).

Subgroup analyses of most important outcome measures

Subgroups by condition

Figure 4 summarises the SMD and subgroup means for between-group changes and the 95% CIs for each condition. The overall test for subgroup differences was significant (χ^2=10.26, p=0.04), indicating significant heterogeneity of SMD between subgroups. Fifteen studies (analysing 1584 patients) included in the meta-analysis investigated pain-related conditions, the overall SMD was non-significant at 0.13 (95% CI −0.01 to 0.28; p=0.08; I^2=46%). Ten studies (287 patients) reported on weight loss, the SMD was marginally significant at 0.52 (95% CI 0.01 to 1.03; p=0.05; I^2=76%). Excluding one outlier[52] reduced I^2 to 14% (SMD 0.27, 95% CI 0.00 to 0.55; p=0.05). Most (nine) of these studies involved balloon and sham balloon insertion. Five studies (342 patients) involved GERD. They showed a significant SMD of 0.65 (95% CI 0.31 to 1.00; p=0.0002; I^2=55%). One study on Parkinson's (34 patients) showed an SMD of 0.36 (95% CI −0.37 to 1.09). Eight studies (655 patients) on other diseases yielded a pooled SMD of 0.44 (95% CI 0.14 to 0.74, p=0.004; I^2=57%).

Subgroups by type of procedure

Between-group SMD did not differ significantly between classical surgery, endoscopic surgery and percutaneous procedures (χ^2=1.10, p=0.58; results not shown).

Dichotomous outcomes

Twelve studies provided only a dichotomous outcome measure (see online supplementary table 1). Sensitivity analyses showed an overall effect of surgery compared to sham surgery (risk ratio 1.54, 95% CI 1.11 to 2.15; p=0.01), while heterogeneity was large (I^2=59%, p=0.005). Subgroup analyses according to condition revealed a significant effect of surgery versus sham surgery for pain studies (n=9; risk ratio 1.60, 95% CI 1.11 to 2.30; p=0.01; I^2=59%, p=0.01) but not for other studies (n=3; risk ratio 2.19, 95% CI 0.44 to 10.84; p=0.33; I^2=60%, p=0.08; see online supplementary figure 1).

Changes from baseline within sham and active groups

The pooled SMD for changes from baseline was 0.61 in the sham groups (95% CI 0.47 to 0.75, p<0.00001, n=39, I^2=76%) and 0.92 (95% CI 0.74 to 1.09, p<0.00001, n=39, I^2=86%) in the treatment groups. Thus, on average, the changes in the sham groups accounted for 65% of the overall improvement from the treatments. This proportion of specific to non-specific treatment effects was larger in pain-related conditions (78%) and obesity (71%) than in GERD (57%) and other conditions (57%), and was considerably smaller in classical surgery trials (21%) than in endoscopic trials (73%) and those using percutaneous procedures (64%; figure 4). Changes in the sham groups accounted for 89% and 82% of overall improvement in intermediate and late pain outcomes.

DISCUSSION

This is the first comprehensive systematic review with meta-analysis estimating the magnitude of the specific

BMJ Open: first published as 10.1136/bmjopen-2015-009655 on 11 December 2015. Downloaded from http://bmjopen.bmj.com/ on September 26, 2024 by guest. Protected by copyright.

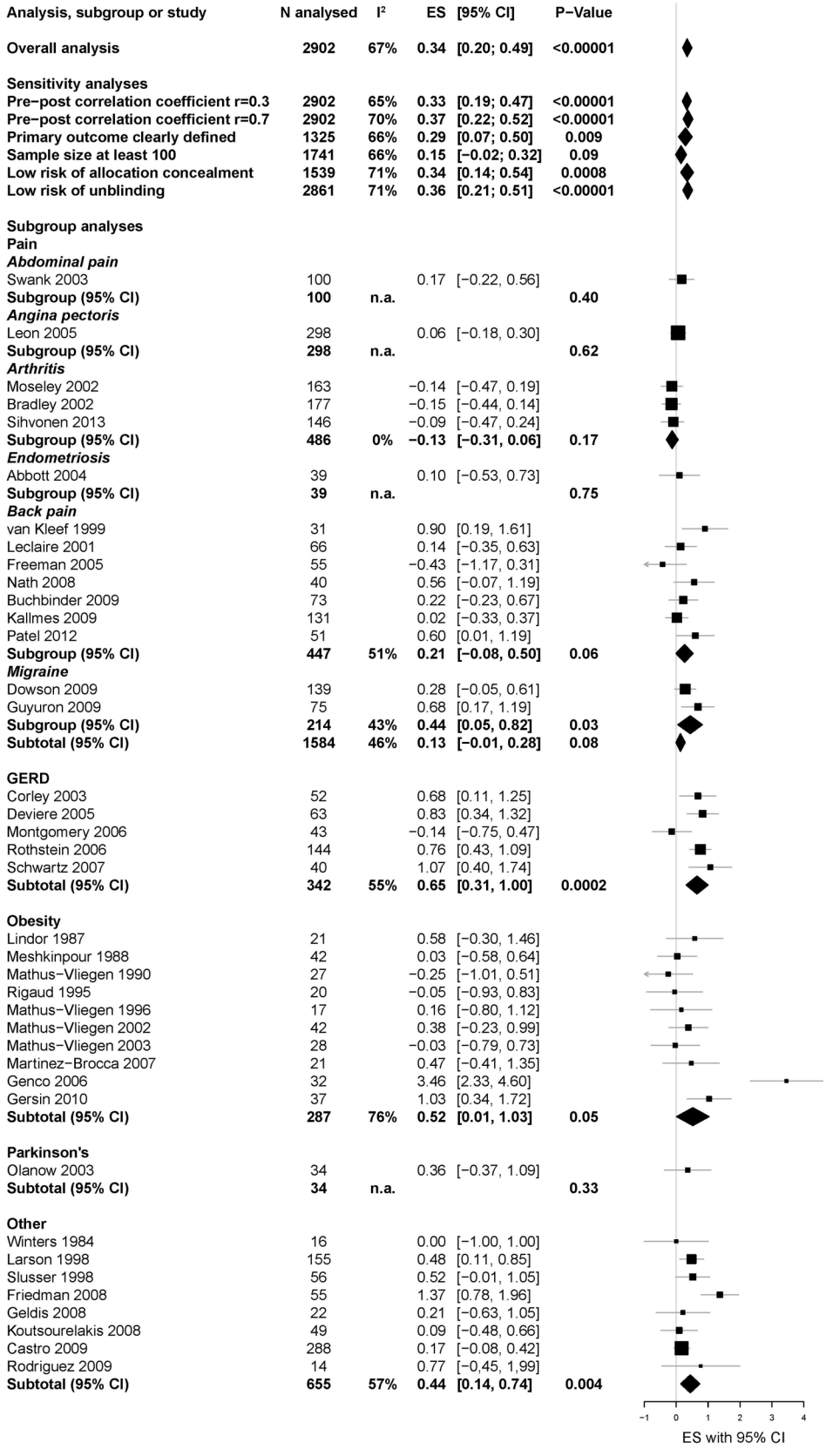


Figure 2 The specific effect of invasive procedures and surgery.

effects of surgery and invasive procedures for various conditions. While some high profile studies have reported no difference between treatment and sham procedures, we found a positive though modest overall effect size (Cohen's d) from the invasive procedures included in the analysis. When only larger studies (≥100 participants) are taken, the specific effects invasive procedures disappears, indicating the current evidence is

BMJ Open: first published as 10.1136/bmjopen-2015-009655 on 11 December 2015. Downloaded from http://bmjopen.bmj.com/ on September 26, 2024 by guest. Protected by copyright.

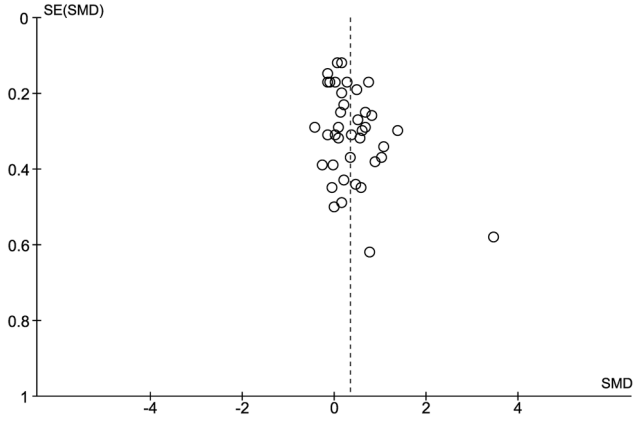


Figure 3 Funnel plot using continuous outcomes (effects of active vs sham treatment) of the 39 studies included in the main meta-analysis.

not strong and could be changed with more and better research. In addition, the contribution of non-specific effects is even more substantial for certain conditions and procedures. While non-specific effects accounted for approximately 65% of the effects from all invasive procedures, they made up to 78% of the active treatment effects in chronic pain conditions and 71% of the active treatment effects in obesity. These percentages are substantially higher than those observed in non-surgical trials, namely 40% for chronic pain conditions and 33% for obesity.[91] The higher contribution of non-specific effects in surgical trials could well be the result of higher placebo effects. However, the lack of no-treatment groups in our data set (and other data set)[92] allows no firm conclusion.[91] Our subgroup analyses indicate that the current evidence does not support the specific efficacy of invasive procedures for chronic pain conditions (p=0.08) and was borderline for obesity (p=0.05), but does support these procedures for GERD (p=0.0002). However, please note that the analysis of dichotomous outcomes showed a somewhat larger specific effect for pain studies (see online supplementary figure 1). There is insufficient data to make recommendations about the other conditions examined.

Strengths and weakness of this study

This study has several limitations. First, both the central strength and limitation of our study is that we pooled effect estimates of the included studies. We consider this a strength at is allows us to: (1) make an estimate of the overall effects of invasive procedures in sham-controlled surgical studies, (2) estimate the strength of confidence in the currently available data as to the specific efficacy of those procedures; and, (3) empirically investigate to what extent results differ between conditions and procedures. Obviously, it is not reasonable to expect that surgery has similar specific effects across conditions and outcomes so our subgroup estimates should not be interpreted clinically without considering how the interventions and outcomes varied. This is also indicated by the moderate-to-large heterogeneity in our meta-analyses, indicating more variation of effect sizes than would be expected by chance. Second, it is difficult to fully double-blind invasive procedures. While most studies successfully blinded patients and outcome assessors, physicians doing these procedures could not be blinded. Thus, it is possible that they communicated information to patients that biased the studies. Price and others have shown that physician expectations can influence pain outcomes even when restrictions are placed on verbal communication.[93 94] Third, publication bias may play a role in the accuracy of our estimates. It is known that negative studies (in this case, studies showing no difference between real and sham procedures) are not published as frequently as positive studies. However, our search strategy was comprehensive and the study selection process was reliable. We also conducted a thorough search of the grey literature, as described above, and had input by experts in placebo research, increasing the likelihood of capturing all studies in this area. This activity allowed for a cross-check in the end to ensure we captured most of the relevant published randomised controlled trials for this review. We did not find any unpublished reports that met our inclusion criteria appropriate for this review, however there were some publications that were not readily accessible through the search engines commonly accessed that we were able to

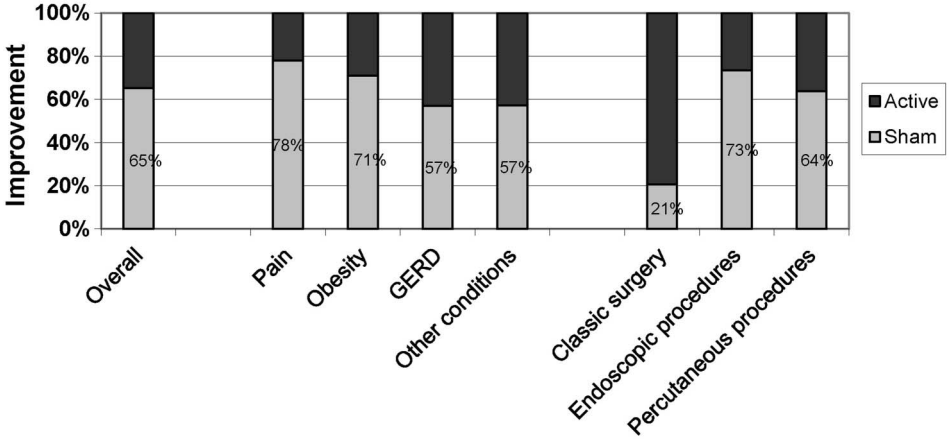


Figure 4 Relative contribution to improvement in the placebo and active treatment groups.

BMJ Open: first published as 10.1136/bmjopen-2015-009655 on 11 December 2015. Downloaded from http://bmjopen.bmj.com/ on September 26, 2024 by guest. Protected by copyright.

capture through these methods. Our sensitivity analyses on study quality factors did not change our primary findings, except restricting the analyses to large studies with 100 participants and above, revealed a considerably smaller, non-significant SMD at 0.15 (95% CI −0.02 to 0.32; p=0.09). Egger's test for funnel plot asymmetry, however, suggested a small study bias in our data set. While our combined estimates of effect size must be considered crude for the overall meta-analysis, they are reasonable estimates for the pain, GERD and obesity subgroups. Meta-analyses of placebo-controlled drug studies in pain, depression, hypertension, ulcer treatment and other areas often report a similar magnitude of specific treatment effects compared to non-specific effects.[95–98] Those studies, however, usually have much larger sample sizes, increasing confidence in their estimates. Finally, we found only one three-armed study that included no treatment, active and sham groups.[67] Therefore, it is not possible to estimate the contribution that the ritual and context make to outcomes in invasive procedures compared to no treatment. Especially in the field of pain and obesity such three-armed studies would seem to be essential for making good evidence-based decisions.

Our findings are consistent with a systematic review published in the *BMJ* in 2014.[92] That study, however, used vote count and reported that 74% of 55 trials showed improvement in the placebo arm with 51% reporting no difference between surgery and placebo and 49% reporting surgery was superior to placebo. We have built on that study by doing a more comprehensive literature search and meta-analysis which allowed us to estimate the magnitude of surgical effects, the confidence in the current findings and to examine that magnitude across various quality parameters, conditions, procedures and outcomes. We can now conclude that at least chronic pain conditions lack clear evidence for the efficacy of the explored surgical interventions (eg, classic surgery and endoscopic procedures. Since these conditions represent a high public health burden worldwide we need to obtain better evidence for the use of these procedures. In addition, it is clear that the evidence from placebo controlled trials in the field as a whole is poor.

Implications for practice, research and policy

These results have a number of implications for practice, research and policy. The evidence from available sham-controlled trials indicates that invasive procedures are not clearly more effective than sham procedures for various chronic pain conditions including endometriosis, back pain, arthritis, angina and migraine. There is evidence to support surgical interventions for GERD and limited evidence to support the use of balloon insertion for obesity.

Given the large number of invasive and surgical procedures being performed, it is noteworthy that we could identify only 55 sham-controlled studies in the literature. Certainly, not all invasive procedures warrant sham-controlled comparisons; for example, when results demonstrate indisputable changes in objective parameters the risks of sham procedures would be excessive. However, given that non-specific factors make a large contribution to the effects from invasive procedures for conditions like pain, more rigorous evaluation is needed before their widespread use is recommended for these conditions. A recent survey of surgeon's attitudes about sham surgery may provide an opportunity to conduct more such research. Surgeons generally agreed that a placebo component to surgical intervention might exist.[99] Furthermore, results of a recent systematic review indicate that the risks of adverse effects associated with sham surgical procedures are small.[92] Thus, more well-designed sham-controlled surgical trials are warranted to avoid the continued use of ineffective invasive treatments.

Author affiliations
[1]Samueli Institute, Alexandria, Virginia, USA
[2]Department of Pain and Translational Symptom Science, School of Nursing, University of Maryland, Baltimore, Maryland, USA
[3]Department of Anesthesiology, School of Medicine, University of Maryland, Baltimore, Maryland, USA
[4]Program in Placebo Studies, Beth Israel Deaconess Medical Center, Harvard Medical School Boston, Massachusetts, USA
[5]The Methodist Hospital, Houston, Texas, USA
[6]Department of Bioethics, Clinical Center, National Institutes of Health, Bethesda, Maryland, USA
[7]Department of Medical Psychology, University Medical Center Hamburg-Eppendorf, Hamburg-Eppendorf, Hamburg, Germany
[8]Institute of General Practice, Technische Universitat Munchen, Munich, Germany
[9]Institute of Medical Psychology, Ludwig-Maximilians-University Munich, Munich, Germany

Acknowledgements The authors would like to thank Ms LaDonna Johnson, Research Assistant at Samueli Institute for assistance with article retrieval and tracking, and Ms Viviane Enslein for assistance with manuscript preparation.

Contributors WBJ served as the PI on this project and was responsible for the conception and design of the project, obtaining funding, acquisition of data and interpretation of the data, drafting and final revision of the article, and final approval of the version submitted. CC served as the project manager and reviewer, and contributed to the conception and design of the systematic review, acquisition of data and analysis and interpretation of data, drafting the article, and approval of the version to be submitted. LC and KM served as study quality reviewers and contributed to the conception and design, acquisition of data, analysis and interpretation of data, drafting the manuscript and approval of the version to be submitted. In addition, KM led the meta-analysis section of the project. TJK and FGM served as subject matter experts, and were involved in the conception and design and interpretation of the data, revising the manuscript critically for important intellectual content and approval of the version to be submitted. LK served as the statistical expert on the project and was involved in the design and conduct of the meta-analysis, acquisition of data, analysis and interpretation of the data, contributing to the manuscript in statistics, meta-analysis techniques, the results section of the manuscript and approval of the version submitted. KL served in the conception and design, acquisition of the data, analysis and interpretation of the data, design of the meta-analysis technique for extracting data, assisted in drafting and revising the article for important intellectual content, and approval of the version to be submitted.

BMJ Open: first published as 10.1136/bmjopen-2015-009655 on 11 December 2015. Downloaded from http://bmjopen.bmj.com/ on September 26, 2024 by guest. Protected by copyright.

Funding This work is supported by the US Army Medical Research and Materiel Command under Award number W81XWH-08-1-0615. The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation. 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 None declared.

Provenance and peer review Not commissioned; externally peer reviewed.

Data sharing statement No additional data are available.

Open Access This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) 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/4.0/

REFERENCES

1. Friedly J, Standaert C, Chan L. Epidemiology of spine care: the back pain dilemma. *Phys Med Rehabil Clin N Am* 2010;21:659–77.
2. Khanna A, Gougoulias N, Longo UG, *et al*. Minimally invasive total knee arthroplasty: a systematic review. *Orthop Clin North Am* 2009;40:479–89, viii.
3. Donnez J, Squifflet J, Donnez O. Minimally invasive gynecologic procedures. *Curr Opin Obstet Gynecol* 2011;23:289–95.
4. Politis M, Lindvall O. Clinical application of stem cell therapy in Parkinson's disease. *BMC Med* 2012;10:1.
5. Rosemurgy AS, Donn N, Paul H, *et al*. Gastroesophageal reflux disease. *Surg Clin North Am* 2011;91:1015–29.
6. Encinosa W, Bernard D, Steiner C, *et al*. Use and costs of bariatric surgery and prescription weight-loss medications. *Health Aff* 2005;24:1039–46.
7. Miller FG, Kaptchuk TJ. Sham procedures and the ethics of clinical trials. *J R Soc Med* 2004;97:576–8.
8. Miller FG, Wendler D. The ethics of sham invasive intervention trials. *Clin Trials* 2009;6:401–2.
9. Meissner K, Bingel U, Colloca L, *et al*. The placebo effect: advances from different methodological approaches. *J Neurosci* 2011;31:16117–24.
10. Johnson A. Surgery as placebo. *Lancet* 1994;344:1140–2.
11. Moseley JB Jr., O'Malley K, Petersen NJ, *et al*. A controlled trial of arthroscopic surgery for osteoarthritis of the knee. *New Engl J Med* 2002;347:81–8.
12. Buchbinder R, Osborne RH, Ebeling PR, *et al*. A randomized trial of vertebroplasty for painful osteoporotic vertebral fractures. *N Engl J Med* 2009;361:557–68.
13. Kallmes DF, Comstock BA, Heagerty PJ, *et al*. A randomized trial of vertebroplasty for osteoporotic spinal fractures. *N Engl J Med* 2009;361:569–79.
14. Papanastassiou ID, Phillips M, Van Meirhaeghe J, *et al*. Comparing effects of kyphoplasty, vertebroplasty, and non-surgical management in a systematic review of randomized and non-randomized controlled studies. *Eur Spine J* 2012;21:1826–43.
15. Meissner K, Fässler M, Rücker G, *et al*. Differential effectiveness of placebo treatments: a systematic review of migraine prophylaxis. *JAMA Intern Med* 2013;173:1941–51.
16. Bannuru RR, McAlindon TE, Sullivan MC, *et al*. Effectiveness and implications of alternative placebo treatments: a systematic review and network meta-analysis of osteoarthritis trials. *Ann Intern Med* 2015;163:365–72.
17. Schauer PR, Kashyap SR, Wolski K, *et al*. Bariatric surgery versus intensive medical therapy in obese patients with diabetes. *N Engl J Med* 2012;366:1567–76.
18. Higgins J. Chapter 8: assessing risk of bias of included studies. In: Higgins J, Green S, eds. *Cochrane handbook of systematic reviews of interventions*. Wiley Online Library, 2008. http://handbook.cochrane.org/chapter_8/8_assessing_risk_of_bias_in_included_studies.htm
19. Cohen J. *Statistical power analysis for the behavioral sciences*. 2nd edn. Hillsdale, NJ: Erlbaum, 1988.
20. Hedges L. Estimation of effect size from a series of independent experiments. *Psychol Bull* 1982;92:490–9.
21. Follmann D, Elliott P, Suh I, *et al*. Variance imputation for overviews of clinical trials with continuous response. *J Clin Epidemiol* 1992;45:769–73.
22. Higgins J, Thompson S, Deeks J, *et al*. Measuring inconsistency in meta-analyses. *BMJ* 2003;327:557–60.
23. Egger M, Smith GD, Schneider M, *et al*. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 1997;315:629–34.
24. DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986;7:177–88.
25. Higgins J, Green S, eds. *Cochrane handbook for systematic reviews of interventions*. Chichester: Wiley-Blackwell, 2008.
26. Liberati A, Altman DG, Tetzlaff J, *et al*. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate health care interventions: explanation and elaboration. *Ann Intern Med* 2009;151:W65–94.
27. Freeman BJ, Fraser RD, Cain CM, *et al*. A randomized, double-blind, controlled trial: intradiscal electrothermal therapy versus placebo for the treatment of chronic discogenic low back pain. *Spine (Phila Pa 1976)* 2005;30:2369–77; discussion 78.
28. Leclaire R, Fortin L, Lambert R, *et al*. Radiofrequency facet joint denervation in the treatment of low back pain: a placebo-controlled clinical trial to assess efficacy. *Spine* 2001;26:1411–16; discussion 17.
29. Nath S, Nath CA, Pettersson K. Percutaneous lumbar zygapophysial (facet) joint neurotomy using radiofrequency current, in the management of chronic low back pain: a randomized double-blind trial. *Spine* 2008;33:1291–7.
30. van Kleef M, Barendse GA, Kessels A, *et al*. Randomized trial of radiofrequency lumbar facet denervation for chronic low back pain. *Spine (Phila Pa 1976)* 1999;24:1937–42.
31. Patel N, Gross A, Brown L, *et al*. A randomized, placebo-controlled study to assess the efficacy of lateral branch neurotomy for chronic sacroiliac joint pain. *Pain Med* 2012;13:383–98.
32. Bradley JD, Heilman DK, Katz BP, *et al*. Tidal irrigation as treatment for knee osteoarthritis: a sham-controlled, randomized, double-blinded evaluation. *Arthritis Rheum* 2002;46:100–8.
33. Moseley JB Jr., Wray NP, Kuykendall D, *et al*. Arthroscopic treatment of osteoarthritis of the knee: a prospective, randomized, placebo-controlled trial. Results of a pilot study. *Am J Sports Med* 1996;24:28–34.
34. Sihvonen R, Paavola M, Malmivaara A, *et al*. Arthroscopic partial meniscectomy versus sham surgery for a degenerative meniscal tear. *N Engl J Med* 2013;369:2515–24.
35. Cobb LA, Thomas GI, Dillard DH, *et al*. An evaluation of internal-mammary-artery ligation by a double-blind technic. *N Engl J Med* 1959;260:1115–18.
36. Dimond EG, Kittle CF, Crockett JE. Comparison of internal mammary artery ligation and sham operation for angina pectoris. *Am J Cardiol* 1960;5:483–6.
37. Leon MB, Kornowski R, Downey WE, *et al*. A blinded, randomized, placebo-controlled trial of percutaneous laser myocardial revascularization to improve angina symptoms in patients with severe coronary disease. *J Am Coll Cardiol* 2005;46:1812–19.
38. Salem M, Rotevatn S, Stavnes S, *et al*. Release of cardiac biochemical markers after percutaneous myocardial laser or sham procedures. *Int J Cardiol* 2005;104:144–51.
39. Salem M, Rotevatn S, Stavnes S, *et al*. Usefulness and safety of percutaneous myocardial laser revascularization for refractory angina pectoris. *Am J Cardiol* 2004;93:1086–91.
40. Cote G, Imperiale T, Schmidt S, *et al*. Similar efficacies of biliary, with or without pancreatic, sphincterotomy in treatment of idiopathic recurrent acute pancreatitis. *Gastroenterology* 2012;143:1502–09 e01.
41. Boelens O, Assen TV, Houterman S, *et al*. A double-blind, randomized, controlled trial on surgery for chronic abdominal pain due to anterior cutaneous nerve entrapment syndrome. *Ann Surg* 2013;257:845–9.
42. Swank D, Swank-Bordewijk S, Hop W, *et al*. Laparoscopic adhesiolysis in patients with chronic abdominal pain: a blinded randomised controlled multi-centre trial. *Lancet* 2003;361:1247–51.
43. Abbott J, Hawe J, Hunter D, *et al*. Laparoscopic excision of endometriosis: a randomized, placebo-controlled trial. *Fertil Steril* 2004;82:878–84.
44. Jarrell J, Brant R, Leung W, *et al*. Women's pain experience predicts future surgery for pain associated with endometriosis. *J Obstet Gynaecol Can* 2007;29:988–91.
45. Jarrell J, Mohindra R, Ross S, *et al*. Laparoscopy and reported pain among patients with endometriosis. *J Obstet Gynaecol Can* 2005;27:477–85.

BMJ Open: first published as 10.1136/bmjopen-2015-009655 on 11 December 2015. Downloaded from http://bmjopen.bmj.com/ on September 26, 2024 by guest. Protected by copyright.

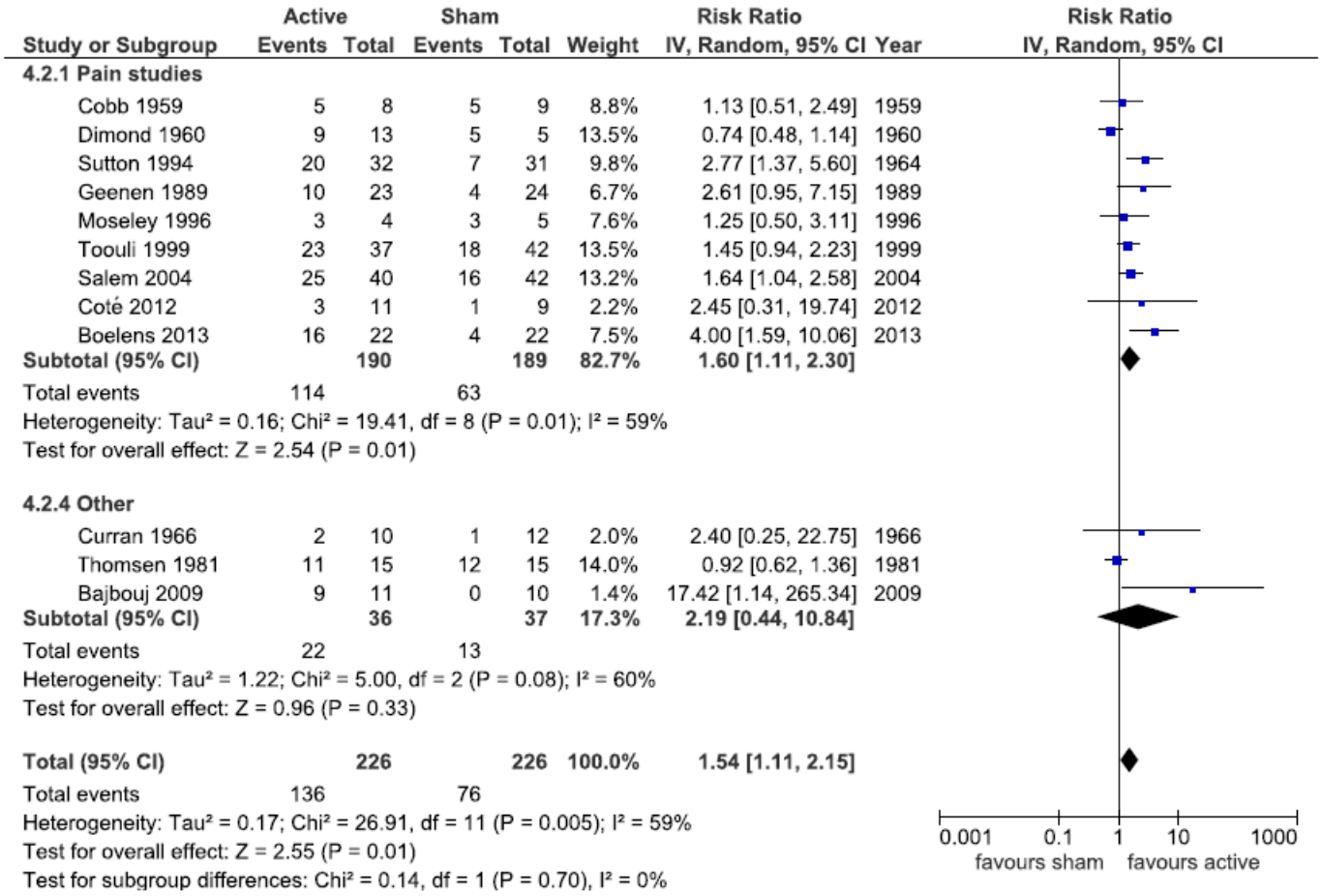
46. Sutton CJ, Pooley AS, Ewen SP, *et al*. Follow-up report on a randomized controlled trial of laser laparoscopy in the treatment of pelvic pain associated with minimal to moderate endometriosis. *Fertil Steril* 1997;68:1070–4.
47. Sutton CJ, Ewen SP, Whitelaw N, *et al*. Prospective, randomized, double-blind, controlled trial of laser laparoscopy in the treatment of pelvic pain associated with minimal, mild, and moderate endometriosis. *Fertil Steril* 1994;62:696–700.
48. Geenen JE, Hogan WJ, Dodds WJ, *et al*. The efficacy of endoscopic sphincterotomy after cholecystectomy in patients with sphincter-of-Oddi dysfunction. *N Engl J Med* 1989;320:82–7.
49. Toouli J, Roberts-Thomson IC, Kellow J, *et al*. Manometry based randomised trial of endoscopic sphincterotomy for sphincter of Oddi dysfunction. *Gut* 2000;46:98–102.
50. Dowson A, Mullen MJ, Peatfield R, *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.
51. Guyuron B, Reed D, Kriegler JS, *et al*. A placebo-controlled surgical trial of the treatment of migraine headaches. *Plast Reconstr Surg* 2009;124:461–8.
52. Genco A, Cipriano M, Bacci V, *et al*. BioEnterics Intragastric Balloon (BIB): a short-term, double-blind, randomised, controlled, crossover study on weight reduction in morbidly obese patients. *Int J Obes (Lond)* 2006;30:129–33.
53. Gersin KS, Rothstein RI, Rosenthal RJ, *et al*. Open-label, sham-controlled trial of an endoscopic duodenojejunal bypass liner for preoperative weight loss in bariatric surgery candidates. *Gastrointest Endosc* 2010;71:976–82.
54. Lindor KD, Hughes RW Jr., Ilstrup DM, *et al*. Intragastric balloons in comparison with standard therapy for obesity—a randomized, double-blind trial. *Mayo Clin Proc* 1987;62:992–6.
55. Martinez-Brocca MA, Belda O, Parejo J, *et al*. Intragastric balloon-induced satiety is not mediated by modification in fasting or postprandial plasma ghrelin levels in morbid obesity. *Obes Surg* 2007;17:649–57.
56. Mathus-Vliegen EM, Tytgat GN, Veldhuyzen-Offermans EA. Intragastric balloon in the treatment of super-morbid obesity. Double-blind, sham-controlled, crossover evaluation of 500-milliliter balloon. *Gastroenterology* 1990;99:362–9.
57. Mathus-Vliegen EMH, Tytgat GNJ. Gastro-oesophageal reflux in obese subjects: influence of overweight, weight loss and chronic gastric balloon distension. *Scand J Gastroenterol* 2002;37:1246–52.
58. Mathus-Vliegen EMH, Tytgat GNJ. Intragastric balloon for treatment-resistant obesity: safety, tolerance, and efficacy of 1-year balloon treatment followed by a 1-year balloon-free follow-up. *Gastrointest Endosc* 2005;61:19–27.
59. Mathus-Vliegen EMH, Van Weeren M, Van Eerten PV. LOS function and obesity: the impact of untreated obesity, weight loss, and chronic gastric balloon distension. *Digestion* 2003;68:161–8.
60. Mathus-Vliegen LMH, Tytgat GNJ. Twenty-four-hour pH measurements in morbid obesity: effects of massive overweight, weight loss and gastric distension. *Eur J Gastroenterol Hepatol* 1996;8:635–40.
61. Meshkinpour H, Hsu D, Farivar S. Effect of gastric bubble as a weight reduction device: a controlled, crossover study. *Gastroenterology* 1988;95:589–92.
62. Rigaud D, Trostler N, Rozen R, *et al*. Gastric distension, hunger and energy intake after balloon implantation in severe obesity. *Int J Obes Relat Metab Disord* 1995;19:489–95.
63. Corley DA, Katz P, Wo JM, *et al*. Improvement of gastroesophageal reflux symptoms after radiofrequency energy: a randomized, sham-controlled trial. *Gastroenterology* 2003;125:668–76.
64. Deviere J, Costamagna G, Neuhaus H, *et al*. Nonresorbable copolymer implantation for gastroesophageal reflux disease: a randomized sham-controlled multicenter trial. *Gastroenterology* 2005;128:532–40.
65. Montgomery M, Hakanson B, Ljungqvist O, *et al*. Twelve months' follow-up after treatment with the EndoCinch endoscopic technique for gastro-oesophageal reflux disease: a randomized, placebo-controlled study. *Scand J Gastroenterol* 2006;41:1382–9.
66. Rothstein R, Filipi C, Caca K, *et al*. Endoscopic full-thickness plication for the treatment of gastroesophageal reflux disease: a randomized, sham-controlled trial. *Gastroenterology* 2006;131:704–12.
67. Schwartz MP, Wellink H, Gooszen HG, *et al*. Endoscopic gastroplication for the treatment of gastro-oesophageal reflux disease: a randomised, sham-controlled trial. *Gut* 2007;56:20–8.
68. Freed CR, Greene PE, Breeze RE, *et al*. Transplantation of embryonic dopamine neurons for severe Parkinson's disease. *N Engl J Med* 2001;344:710–19.
69. Gordon PH, Yu Q, Qualls C, *et al*. Reaction time and movement time after embryonic cell implantation in Parkinson disease. *Arch Neurol* 2004;61:858–61.
70. Greene PE, Fahn S. Status of fetal tissue transplantation for the treatment of advanced Parkinson disease. *Neurosurg* 2002;13:e3.
71. McRae C, Cherin E, Diem G, *et al*. Does personality change as a result of fetal tissue transplantation in the brain? *J Neurol* 2003;250:282–6.
72. McRae C, Cherin E, Yamazaki TG, *et al*. Effects of perceived treatment on quality of life and medical outcomes in a double-blind placebo surgery trial. *Arch Gen Psychiatry* 2004;61:412–20.
73. Nakamura T, Dhawan V, Chaly T, *et al*. Blinded positron emission tomography study of dopamine cell implantation for Parkinson's disease. *Ann Neurol* 2001;50:181–7.
74. Olanow CW, Goetz CG, Kordower JH, *et al*. A double-blind controlled trial of bilateral fetal nigral transplantation in Parkinson's disease. *Ann Neurol* 2003;54:403–14.
75. Friedman M, Schalch P, Lin HC, *et al*. Palatal implants for the treatment of snoring and obstructive sleep apnea/hypopnea syndrome. *Otolaryngol Head Neck Surg* 2008;138:209–16.
76. Koutsourelakis I, Georgoulopoulos G, Perraki E, *et al*. Randomised trial of nasal surgery for fixed nasal obstruction in obstructive sleep apnoea. *Eur Respir J* 2008;31:110–17.
77. Geldis JR, Nichols JJ. The impact of punctal occlusion on soft contact lens wearing comfort and the tear film. *Eye Contact Lens* 2008;34:261–5.
78. Slusser TG, Lowther GE. Effects of lacrimal drainage occlusion with nondissolvable intracanalicular plugs on hydrogel contact lens wear. *Optom Vis Sci* 1998;75:330–8.
79. Castro M, Rubin AS, Laviolette M, *et al*. Effectiveness and safety of bronchial thermoplasty in the treatment of severe asthma: a multicenter, randomized, double-blind, sham-controlled clinical trial. *Am J Respir Crit Care Med* 2010;181:116–24.
80. Curran WS, Graham WG. Long term effects of glomectomy. Follow-up of a double-blind study. *Am Rev Respir Dis* 1971;103:566–8.
81. Curran WS, Oser JF, Longfield AN, *et al*. Glomectomy for severe bronchial asthma. A double-blind study. *Am Rev Respir Dis* 1966;93:84–9.
82. Bajbouj M, Becker V, Eckel F, *et al*. Argon plasma coagulation of cervical heterotopic gastric mucosa as an alternative treatment for globus sensations. *Gastroenterology* 2009;137:440–4.
83. Bretlau P, Thomsen J, Tos M, *et al*. Placebo effect in surgery for Meniere's disease: a three-year follow-up study of patients in a double blind placebo controlled study on endolymphatic sac shunt surgery. *Am J Otol* 1984;5:558–61.
84. Davys HJ, Turner DE, Helliwell PS, *et al*. Debridement of plantar callosities in rheumatoid arthritis: a randomized controlled trial. *Rheumatology (Oxford)* 2005;44:207–10.
85. Larson TR, Blute ML, Bruskewitz RC, *et al*. A high-efficiency microwave thermoablation system for the treatment of benign prostatic hyperplasia: results of a randomized, sham-controlled, prospective, double-blind, multicenter clinical trial. *Urology* 1998;51:731–42.
86. Thomsen J. Placebo effect in surgery for Meniere's disease. *Arch Otolaryngol* 1981;107:271–7.
87. Thomsen J, Bretlau P, Tos M, *et al*. Placebo effect in surgery for Meniere's disease: three-year follow-up. *Otolaryngology—Head & Neck Surgery* 1983;91:183–6.
88. Thomsen J, Bretlau P, Tos M, *et al*. Endolymphatic sac-mastoid shunt surgery. A nonspecific treatment modality? *Ann Otol Rhinol Laryngol* 1986;95:32–5.
89. Winters C, Artnak EJ, Benjamin SB, *et al*. Esophageal bougienage in symptomatic patients with the nutcracker esophagus. A primary esophageal motility disorder. *JAMA* 1984;252:363–6.
90. Rodriguez L, Reyes E, Fagalde P, *et al*. Pilot clinical study of an endoscopic, removable duodenal-jejunal bypass liner for the treatment of type 2 diabetes. *Diabetes Technol Ther* 2009;11:725–32.
91. Krogsboll LT, Hrobjartsson A, Gotzsche PC. Spontaneous improvement in randomised clinical trials: meta-analysis of three-armed trials comparing no treatment, placebo and active intervention. *BMC Med Res Methodol* 2009;9:1.
92. Wartolowska K, Judge A, Hopewell S, *et al*. Use of placebo controls in the evaluation of surgery: systematic review. *BMJ* 2014;348:g3253.
93. Robinson ME, Gagnon CM, Riley JL III, *et al*. Altering gender role expectations: effects on pain tolerance, pain threshold, and pain ratings. *J Pain* 2003;4:284–8.

BMJ Open: first published as 10.1136/bmjopen-2015-009655 on 11 December 2015. Downloaded from http://bmjopen.bmj.com/ on September 26, 2024 by guest. Protected by copyright.

94. Vase L, Robinson ME, Verne GN, *et al*. The contributions of suggestion, desire, and expectation to placebo effects in irritable bowel syndrome patients. An empirical investigation. *Pain* 2003;105:17–25.
95. Meissner K, Distel H, Mitzdorf U. Evidence for placebo effects on physical but not on biochemical outcome parameters: a review of clinical trials. *BMC Med* 2007;5:3.
96. Mora M, Nestoriuc Y, Rief W. Lessons learned from placebo groups in antidepressant trials. *Philos Trans R Soc Lond B Biol Sci* 2011;366:1879–88.
97. Enck P, Klosterhalfen S. The placebo response in functional bowel disorders: perspectives and putative mechanisms. *Neurogastroenterol Motil* 2005;17:325–31.
98. Leucht S, Hierl S, Kissling W, *et al*. Putting the efficacy of psychiatric and general medicine medication into perspective: review of meta-analyses. *Br J Psychiatry* 2012;200:97–106.
99. Wartolowska K, Beard D, Carr A. Attitudes and beliefs about placebo surgery among orthopedic shoulder surgeons in the United Kingdom. *PLoS ONE* 2014;9:e91699.

BMJ Open: first published as 10.1136/bmjopen-2015-009655 on 11 December 2015. Downloaded from http://bmjopen.bmj.com/ on September 26, 2024 by guest. Protected by copyright.

Supplemental Figure 1. Meta-analysis of dichotomous outcomes from 12 studies that provided no continuous outcome



Online Supplemental Material

Supplemental Table 1. Characteristics of Included Studies

Source	Condition	N Total Entered (Treatment/ Control)	Treatment	Sham Treatment	Concealment	Outcome for Main Analysis	Author Conclusions
PAIN							
Sutton et al,1994[46] Sutton et al,1997[45]	endometriosis	T32/ C 31	laprasocopic laser treatment of endometriosis	diagnostic laparoscopy with three incisions in the same locations as the active treatment and the removal of the serosanguinous fluid from the Pouch of Douglas to perform an inspection of the pelvic peritoneum	Unclear	patients with pain improvement at 6 months ‡	(+)
Abbott et al, 2004[42]	endometriosis	T 20/ C 19	laparoscopic excision of endometriosis tissues	diagnostic laparoscopic procedure with no laparoscopic excision of endometriosis tissues -(delayed surgery)	Adequate	overall pain reduction at 6 months †	(+)
Jarrell et al,2005[44] Jarrell et al, 2007[43]	endometriosis	T 8/ C 7	laparoscopy biopsy and excision	three incisions and a biopsy but no tissue excision	Adequate	§	(+)
Leclaire et al, 2001[27]	low back pain (> 3 months)	T 36/ C 34	percutaneous radiofrequency articular denervation done under fluoroscopic guidance	the same procedure without denervation	Adequate	change of Roland-Morris questionnaire score at 12 weeks *	0

Van Kleef et al, 1999[29]	chronic low back pain	T 15/ C 16	radiofrequency lumbar facet denervation (60-second radiofrequency lesion at 80 C of the medial branch of the posterior primary ramus of the segmental nerves L3–L5 on one or on both sides)	sham denervation where electrodes were introduced as in the active treatment but no radiofrequency lesion was made	Unclear	VAS pain score at 8 weeks †	(+)
Nath et al, 2008[28]	chronic low back pain	T 20/ C 20	percutaneous radiofrequency neurotomy	identical procedure except no current was used and the electrode tip remained at body temperature	Unclear	global improvement for back pain according to VAS at 6 months †	(+)
Freeman et al, 2005[26]	chronic discogenic low back pain	T 38/ C 19	positioning of IDET catheter and delivering of electrothermal energy	positioning of the IDET catheter without delivering electrothermal energy	Unclear	Low back pain outcome score (LBOS) at 6 months †	0
Buchbinder et al, 2009[12]	painful osteoporotic vertebral fractures	T 38/ C 40	vertebroplasty	same procedures up to the insertion of the 13-gauge needle to rest on the lamina. The central sharp stylet was then replaced with a blunt stylet. To simulate vertebroplasty, the vertebral body was gently tapped, and polymethylmethacrylate was prepared so that its smell permeated the room.	Adequate	overall pain score at 3 months *	0

Kallmes et al, 2009[13]	painful osteoporotic vertebral fractures	T 68/ C 63	vertebroplasty	simulated procedure without cement where methacrylate monomer was opened to simulate the odor associated with PMMA mixing but needles were not placed nor PMMA infused	Adequate	RDQ (Roland-Morris disability score) at 1 month †	0
Patel et al, 2012[30]	chronic sacroiliac joint pain	T 34/ C 17	lateral branch neurotomy using cooled radiofrequency	sham neurotomy where the treatment was identical, except that radiofrequency energy was not delivered.	Unclear	comparison of mean change from baseline in NRS pain score between treatment and sham groups at the 3-month follow-up time-point *	(+)
Cobb, 1959[34]	angina pectoris	T 8/ C 9	ligation of the internal mammary arteries after they had been isolated	sham ligation of the internal mammary arteries after they had been isolated	Unclear	subjective improvement in angina >40% ‡	0
Dimond, 1960[35]	angina pectoris	not stated	ligation of internal mammary artery	sham ligation (skin incision with exposure of the internal mammary arteries with no ligation)	Unclear	>=50% subjective improvement ‡	0
Salem et al, 2004[38] Salem et al, 2005[37]	angina pectoris	T 40/ C 42	Percutaneous Myocardial Laser Revascularization (PMLR) plus optimal medical therapy	sham procedure involving the laser catheter being inserted but connected to a hidden lead box plus optimal medical therapy	Adequate	mean Canadian Cardiovascular Society angina (CCS) class at 12 months ‡	(+)

Leon,et al, 2005[36]	angina pectoris	T 196/ C 102	Biosense direct myocardial revascularization (DMR) with laser catheter introduced and advanced to the left ventricular (LV)	the laser (already in the room) was turned on but no further procedure was performed	Unclear	exercise treadmill duration at 6 months *	0
Moseley et al, 1996[32]	arthritis	Lavage: 3/ Debribement: 2/ C 5	arthroscopic débridement or arthroscopic lavage	patients undergoing placebo arthroscopy were prepared, draped, examined, and injected with local anesthetic in the same manner as the other two groups. Three stab wounds were made in the skin with a scalpel, but no instruments of any kind were placed into the knee	Unclear	a "strongly agree" response to question "Do you feel the operation was worthwhile?" ‡	(-)
Moseley et al, 2002[11]	arthritis	Lavage: 61/ Debribement: 59/ C 60	arthroscopic débridement or arthroscopic lavage	as described above	Adequate	KSPS (knee-specific pain scale) at 24 months *	0
Bradley et al, 2002[31]	arthritis	T 89/ C 91	arthroscopic knee irridation	sham irridation (needle was advanced to, but not through, the joint capsule)	Unclear	WOMAC function score (Western Ontario and McMaster Universities Osteoarthritis Index) at 52 weeks †	(-)
Sihvonen et al., 2013[33]	arthritis	T 70/ C 76	arthroscopic partial meniscectomy	a standard arthroscopic partial meniscectomy was simulated with sham surgery	Adequate	the Lysholm knee score at 12 mos after surgery *	0
Dowson et al, 2008[49]	migraine with aura	T 74/ C 73	Patent foramen ovale (PFO) closure with the STARFlex septal repair implant	a sham procedure (skin incision in the groin only)	Adequate	migraine attacks/month at 4-6 months †	0

Guyuron et al, 2009[50]	migraine headache	T 49/ C 26	Surgery in the predominant trigger sites (frontal (F), temporal (T) and occipital (O)) with endoscopic removal of the glabellar muscles encasing the supraorbital and supratrochlear nerves, removal of a segment of the zygomaticotemporal branch of the trigeminal nerve, and removal of the greater occipital nerve	sham surgery with exposure of the muscles and nerves through a similar incision but the integrity of the structures was maintained	Adequate	change in frequency of migraine attacks at 12 months †	(+)
Geenen et al, 1989[47]	cholia	T 23/ C 24	endoscopic sphincterotomy	sham sphincterotomy was performed exactly as was the true sphincterotomy except that the sphincterotome was positioned in the duodenal lumen during activation of the electrocautery unit	Unclear	patients with "good" symptom scores at 12 months ‡	(+)
Toouli et al, 2000[48]	cholia	SO stenosis: Sham 13/ Endoscopic sphinctectomy (ES) 13/ SO dyskinesia: sham 10/ ES 11/ Normal manometry: Sham 19/ ES 13	endoscopic sphinctectomy (ES)	endoscopy with introduction of a papillotome into the duodenum after which the "noises" of sphincterotomy were created	Unclear	improvement in abdominal pain at 24 months ‡	(+)

Cote et al, 2012[39]	painful pancreatic sphincter dysfunction	T 11/ C 9	biliary endoscopic sphincterotomy (BES)	sham endoscopy (not described)	Adequate	patients with two or more distinct episodes of acute pancreatitis at follow up evaluation ‡	0
Boelens et al, 2013[40]	painful anterior cutaneous nerve entrapment syndrome	T 22/ C 22	Neurectomy of the intercostal nerve endings at the level of the abdominal wall	Sham surgery with the exposure to the abdominal wall and then closed with no further surgical procedure	Adequate	patients achieving a minimal 50% improvement in pain perception measured using a VAS ‡	(+)
Swank et al, 2003[41]	chronic abdominal pain and adhesions	T 52/ C 48	Laparoscopic adhesiolysis	diagnostic laparoscopy only	Adequate	pain relief after one year using a VAS *	0
Davys et al, 2005[83]	plantar callosities in rheumatoid arthritis	T 19/ C 19	normal callus treatment (NCT) comprised of sharp scalpel debridement of the callosity	NCT was simulated by the podiatrist using a bluntedged scalpel such that no callus tissue was debrided	Adequate	forefoot pain §	0
Obesity							
Lindor et al, 1987[53]	obesity	T 11/ C11	intragastric balloon placed	an empty introducer tube was inserted, its position was confirmed fluoroscopically and 200 ml of air was pumped into the stomach	Unclear	weight loss at 2-3 months †	0
Meshkinpour et al, 1988[60]	obesity	T 21/ C 21	the Garren-Edwards gastric bubble inserted	insertion of an unloaded introducer tube after having modified the slit and simulation of the bubble inflation process	Unclear	weight loss at 12 weeks †	0

Mathus-Vliegen et al, 1990[55]	obesity	T 14/ C 14	four treatment combinations were compared: group A: balloon-sham; group B: sham-balloon; group C balloon-balloon; and group D: sham-sham. the treatment included two periods of each of balloon or sham therapy for 4 months	sham groups involved insertion of a balloon where the balloon was empty	Unclear	weight loss at 1-17 weeks †	0
Rigaud et al, 1995[61]	obesity	T 11/ C 9	endoscopic air filled balloon insertion	sham involved empty balloon insersion	Unclear	weight †	0
Mathus-Vliegen et al, 1996[59]	obesity	T 8/ C 9	endoscopic balloon insertion	empty balloon insertion	Unclear	weight at 4 months †	0
Mathus-Vliegen et al, 2002[56]	obesity	T 20/ C 23	endoscopic balloon insertion	empty balloon insertion	Unclear	weight at 13 weeks †	0
Mathus-Vliegen et al, 2003[58]	obesity	T 11/ C 17	endoscopic balloon insertion	empty balloon insertion	Unclear	weight at 13 weeks †	0
Genco et al, 2006[51]	obesity	T 16/ C 16	endoscopic balloon insertion	endoscopic examination in sedated patients but no balloon insertion in the first 3 months of treatment	Unclear	body mass index †	(+)
Gersin et al, 2010[52]	obesity	T 21/ C 26	endoscopic insertion of a duodenojejunal bypass liner (DJBL)	conscious sedation during which an EGD and a mock procedure were performed	Unclear	percentage of excess weight loss at 12 weeks *	(+)

Mathus-Vliegen et al, 2005[57]	treatment-resistant obesity	T 12/ C 21	the placement assembly consisted of a sheath with the collapsed balloon and a balloon fill tube	the collapsed balloon was not present in the assembly but otherwise mimicked the balloon procedure	Adequate	§	0
Martinez-Brocca et al, 2007[54]	treatment-resistant obesity	T 11/ C 11	endoscopic balloon insertion	insertion of a deflated balloon	Unclear	weight at 12 weeks †	(+)
GERD							
Corley et al, 2003[62]	GERD	T 35/ C 29	radiofrequency energy delivery to the gastroesophageal junction at each deployment position	balloon inflation at each deployment position without needle deployment or energy delivery	Adequate	heartburn score at 6 months †	(+)
Deviere et al, 2005[63]	GERD	T 32/ C 32	endoscopic implantation of a biocompatible nonresorbable copolymer	diagnosic upper endoscopy only endoscopy	Adequate	GERD-HRQL heartburn score reduction at 3 months †	(+)
Montgomery et al, 2006[64]	GERD	T 22/ C 24	gastric plication with endoscopical placement of two to four sutures	endoscopy with patient anaesthetized for approximately 40 min (corresponding to the time needed to perform the real procedure)	Adequate	GSRS gastrointestinal symptom rating scale at 3 months †	0

Rothstein et al, 2006[65]	GERD	T 72/ C 72	gastric plication with endoscopical placement of one sutures	The sham procedure mimicked the treatment procedure through the positioning of the plicator but with no sutures placed	Adequate	GERD-HRQL (health-related quality of life) at 3 months †	(+)
Schwartz et al, 2007[66]	GERD	T 20/ C 20	endoscopic gastroplication along the lesser curvature and greater curvatures	the sham procedure was carried out using the same equipment, but without the suturing needle and thread loaded	Adequate	heartburn score at 3 months †	(+)
Parkinson's							
Freed et al, 2001[67] Nakamura et al, 2001[72] Greene et al, 2002[69] McRae et al, 2003[70] McRae et al, 2004[71] Gordon et al, 2004[68]	parkinson's disease	T 20/ C 20	deep brain implantation of embryonic dopamine neurons	identical procedure except that the dura mater was not penetrated after the twist-drill holes had been made in the frontal bone	Unclear	§	0
Olanow et al, 2003[73]	parkinson's disease	T 23/ C 11	bilateral fetal nigral transplantation into the brain	treated in an identical manner except that partial burr holes that did not penetrate the inner table of the skull, no needles were inserted and no tissue was implanted	Unclear	UPDRS motor score (Unified Parkinson's Disease Rating Scale) at 24 months *	0
Other							

Curran, 1966[80] Curran, 1971[79]	asthma	T 10/ C 13	standard glomectomy procedure	sham glomectomy in which the carotid bifurcation was only exposed and subcutaneous tissue was sent to the pathologist	Unclear	physician's opinion of "definite" improvement at 6 months ‡	0
Thomsen et al, 1983[86] Bretlau et al, 1984[82] Thomsen,1981 [85] Thomsen,1986 [87]	meniere's disease	T 15/ C 15	regular endolymphatic sac shunt operation with insertion of silastic into the sac draining out into the mastoid cavity	mastoidectomy with care was taken not to remove the bone over the endolymphatic sac in order to avoid a decompression	Unclear	investigator's evaluation of operative effect: "good effect" at 12 months ‡	(-)
Winters, 1984[88]	dialated esophagus	T 8/ C 8	bougienage (dilatation of esophagus) with a 54 F bougie	sham bougienage using a 24 F bougie	Unclear	chest pain score at 2 weeks †	0
Larson et al, 1998[84]	Benign Prostatic Hyperplasia (BPH)	T 125/ C 44	high efficiency microwave thermoablation	same procedure except without microwave power and coolant temperature was increased in increments from 8 to 20°C over the same time period as in the microwave group	Unclear	AUA score (American urological association) at 6 months †	(+)
Slusser et al, 1998[77]	dry eye	T 28/ C 28	Herrick lacrimal plugs were inserted into both canaliculi of one previously randomized eye	manipulation of eye without insertion of Herrick lacrimal plugs	Unclear	subjective dryness of eye (VAS) at 5 weeks †	(+)
Geldis et al, 2008[76]	dry eye	T 19/ C 19	insertion of punctal plug	tissue manipulation producing the same sensation to the eye with no plug insersion	Unclear	contact lens dry eye questionnaire (CLDEQ) score at 6 weeks †	0

Friedman et al, 2008[74]	sleep apnea	T 31/ C 31	palatal implant insertion	identical procedure with no insertion of palatal implants	Adequate	reduction in AHI (apnoea/hypopnoea index) at 95 days *	(+)
Koutsourelakis et al, 2008[75]	sleep apnea	T 27/ C 22	standard submucosal resection of the nasal septum	nasal tissue manipulation simulating the operation but with no resection of the nasal septum	Unclear	reduction in AHI (apnoea/hypopnoea index) at 3-4 months †	0
Bajbouj et al, 2009[81]	globus	T 11/ C 10	endoscopy performed and APC therapy applied	after endoscopy, patients were connected with APC applicator but current flow was disconnected as soon as patients were asleep	Adequate	patients with overall symptom improvement (yes or no) at 3 months ‡	(+)
Castro et al., 2009[79]	asthma	T 190/ C98	bronchoscopy with bronchial thermoplasty	bronchoscopy, which mimicked thermal bronchoplasty	Adequate	asthma quality of life (AQLQ) score, average of 6, 9 and 12 weeks †	(+)
Rodriguez et al, 2009[89]	Type II diabetes	T 12/ C 6	endoscopic insertion of a duodenal-jejunal bypass liner (DJBL)	sham endoscopy	Unclear	improvement in glycemic control as measured by HbA1c change from baseline to week 12 *	(+)

* primary main outcome measure used; † most important outcome measure used; ‡ only responder data available, not used in main analysis; § insufficient data to include in meta-analysis

author conclusions: (+) favors intervention, (0) no benefit, (-) favors placebo or "placebo effect noted in author conclusions"

T: real treatment; C: sham control