



Original Investigation | Surgery

Use of Placebo and Nonoperative Control Groups in Surgical Trials

A Systematic Review and Meta-analysis

Teemu Karjalainen, MBBS, PhD; Juuso Heikkinen, PhD; Ljoudmila Busija, PhD; Jarkko Jokihaara, PhD; Adriane M. Lewin, MSc; Justine M. Naylor, PhD; Laura Harris, MPH; Ian A. Harris, MBBS, MMed, PhD; Rachelle Buchbinder, MBBS, MSc, PhD; Sam Adie, MBBS, MSpMed, MPH, PhD

Abstract

IMPORTANCE Nonspecific effects, particularly placebo effects, are thought to contribute significantly to the observed effect in surgical trials.

OBJECTIVE To estimate the proportion of the observed effect of surgical treatment that is due to nonspecific effects (including the placebo effect).

DATA SOURCES Published Cochrane reviews and updated, extended search of MEDLINE, Embase, and CENTRAL until March 2019.

STUDY SELECTION Published randomized placebo-controlled surgical trials and trials comparing the effect of the same surgical interventions with nonoperative controls (ie, no treatment, usual care, or exercise program).

DATA EXTRACTION AND SYNTHESIS Pairs of authors independently screened the search results, assessed full texts to identify eligible studies and the risk of bias of included studies, and extracted data. The proportion of all nonspecific effects was calculated as the change in the placebo control divided by the change in the active surgery and pooled in a random-effect meta-analysis. To estimate the magnitude of the placebo effect, we pooled the difference in outcome between placebo and nonoperative controls and used metaregression to estimate the association between the type of control group and the treatment effect (difference between the groups), adjusting for risk of bias, sample size, and type of outcome.

MAIN OUTCOMES AND MEASURES Between- and within-group effect sizes expressed as Hedges *g*.

RESULTS In this review, 100 trials were included comprising data from 62 trials with placebo controls (3 also included nonoperative controls), and 38 trials with nonoperative controls (32 interventions; 10 699 participants). Risk of bias across trials was comparable except for performance and detection bias, which was high in trials with nonoperative controls. The mean nonspecific effects accounted for 67% (95% CI, 61% to 73%) of the observed change after surgery; however, this varied widely between different procedures. The estimated surgical placebo effect had a standardized mean difference (SMD) of 0.13 (95% CI, -0.26 to 0.51). Trials with placebo and nonoperative controls found comparable treatment effects (SMD, -0.09 [95% CI, -0.35 to 0.18]; 15 interventions; 73 between-group effects; adjusted analysis: SMD, -0.11 [95% CI, -0.37 to 0.15]).

(continued)

Key Points

Question How large are nonspecific effects, including placebo effects, in surgical randomized trials?

Findings In this meta-analysis of 100 trials, those with nonoperative and nonblinded controls were found to yield results largely comparable with placebo-controlled trials. Nonspecific effects constituted a mean of two-thirds of the total effect of surgery.

Meaning This study found that improvements not caused by the surgery itself were mostly nonspecific effects other than placebo effects, suggesting that when surgical trials with nonoperative controls demonstrate no benefits, placebo-controlled studies may be redundant.

+ Supplemental content

Author affiliations and article information are listed at the end of this article.

Open Access. This is an open access article distributed under the terms of the CC-BY License.

Abstract (continued)

CONCLUSIONS AND RELEVANCE In this review, the change in health state after surgery was composed largely of nonspecific effects, but no evidence supported a large placebo effect. Placebo-controlled surgical trials may be redundant when trials with nonoperative controls consistently report no substantial association from surgery compared with nonoperative treatment.

JAMA Network Open. 2022;5(7):e2223903. doi:10.1001/jamanetworkopen.2022.23903

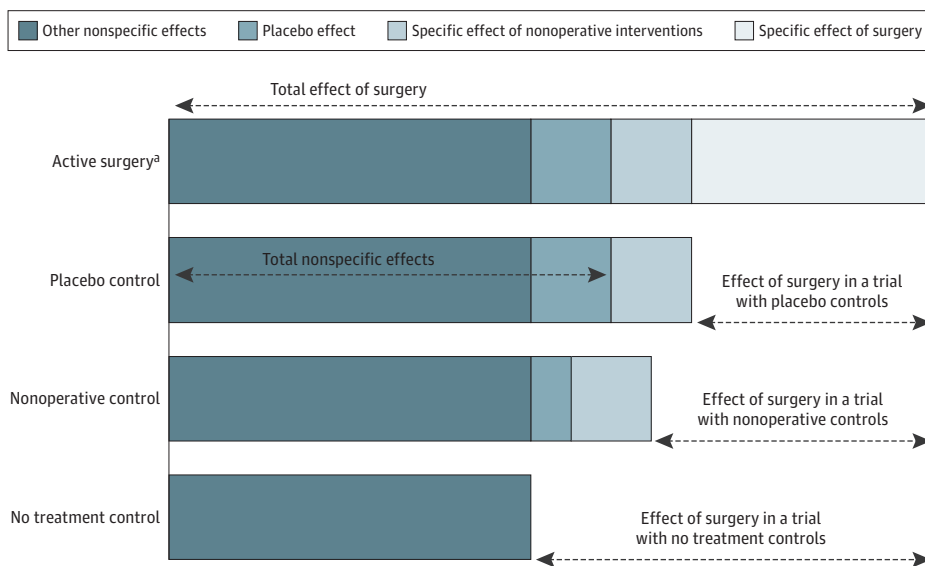
Introduction

Health changes after surgery comprise 2 parts—specific and nonspecific effects. Specific effects are a direct result of the physical changes caused by the key therapeutic element of the surgical procedure (eg, repair of a torn tendon). Nonspecific effects are not dependent on the key therapeutic step and can further be divided into placebo effects and other nonspecific effects, such as regression to the mean and the natural history of the disease.^{1,2}

To offset various nonspecific effects, trialists need to use a control group to measure the causal effect of surgery (ie, treatment effect). While the use of nonblinded and nonoperative controls enables the estimation of the effect of undergoing surgery compared with nonoperative treatment, use of a placebo control identifies the fundamental efficacy of the surgical intervention by accounting for all nonspecific effects, including the surgical placebo effect (**Figure 1**).

Rigorous placebo-controlled studies increase the certainty about the true therapeutic effect or efficacy of surgery, thereby providing best evidence for decision-making in clinical practice.³ It is not always possible or necessary (eg, for objective outcomes, such as death) to assess the effect of surgery with a placebo control. However, when surgery is used to improve pain and quality of life for non-life-threatening or self-limiting conditions, nonblinded trials could overestimate the benefits of surgery. The use of placebo controls has been considered particularly important in these cases. Surgery may elicit a pronounced placebo effect because of the imposing therapeutic envelope of inherent risk, the technical aspects of the intervention, and the higher costs compared with nonoperative treatment.⁴⁻⁶ However, there is no systematic evidence to support this hypothesis—previous studies suggesting a large placebo effect measured improvement after receiving placebo intervention that includes both placebo and other nonspecific effects combined.⁷⁻¹² The aims of this

Figure 1. Observed Effects in a Hypothetical Study



study were to estimate (1) the mean proportion of the observed effect of surgery that can be attributed to total nonspecific factors; (2) the magnitude of the surgical placebo effect; and (3) the difference in the treatment effect between placebo- and nonoperative-controlled surgical trials testing the same surgical intervention.

Methods

Study Design and Definitions Used in This Study

We designed a meta-epidemiological study that included meta-analysis and metaregression. This systematic review and meta-analysis followed the Preferred Reporting Items for Systematic Reviews and Meta-analyses (PRISMA) reporting guideline where applicable.¹³ Key concepts are detailed in **Table 1**.

Inclusion Criteria and Identification of Eligible Studies

We included randomized clinical trials that assessed the efficacy of any surgical intervention by including a placebo control. We also included comparative effectiveness trials where the same surgical intervention (as in the identified placebo trials) was compared with either no treatment, usual care, or exercise programs. Trials where the nonoperative control received cointerventions not provided to the surgical group were excluded because these could not estimate the effect of having surgery.

We defined surgery as any intervention that changes anatomy (including use of an implant), requires a skin or mucosal incision, or suturing to achieve the therapeutic effect. Surgical placebo was defined as an imitation procedure that cannot be differentiated by the patient but lacks the key therapeutic element.

We identified eligible placebo surgical studies from a published systematic review² and extended the search to November 21, 2018. The reference lists of included studies were also checked for eligible studies. To identify relevant effectiveness trials, we first identified relevant Cochrane reviews assessing the index surgical procedure and extended their literature searches to March 13 to March 15, 2019. When no relevant Cochrane review was identified, we devised a search algorithm and searched from the inception of databases until the date of search (eAppendix 1 in the [Supplement](#)).

Table 1. Definitions Used in This Study

Term	Definition
Effect of placebo surgery	The observed effect in participants who receive a placebo intervention. This is the total sum of all nonspecific effects (ie, placebo effect and other nonspecific effects).
Effect of surgery	The difference between active surgery and control group in a trial, ie, the quantitative measure of the effect that surgery has on an outcome measure. The magnitude can vary based on the type of control group.
Placebo effect	The effect experienced because psychological processes (based on expectations, conditioning or patient-caregiver interaction) related to receiving an intervention.
Other nonspecific effects	These include (but are not restricted to) the natural course of the disease, regression to the mean and the effect of other interventions (which includes effects and placebo effects of nonoperative cointerventions). This is the observed effect in the nonoperative control group.
Placebo surgery	A placebo surgical procedure that follows all the steps of a normal procedure, except (at minimum) the key therapeutic step is omitted. The fidelity of placebo surgery (extent of similarity to the active intervention) may vary from skin incision only (low fidelity) to including all other steps of the procedure except the presumed active component (high fidelity). ⁸
Specific effect of surgery	The effect provided by the key therapeutic step of a surgical procedure. This is the difference between active surgery and placebo control group.
Surgery	Any intervention that changes anatomy (including use of an implant to achieve that) or requires a skin or mucosal (epithelial) incision or suturing to achieve its effect.
Total effect of surgery	The sum of all effects, nonspecific and specific. This is the observed effect in the active surgery group.

Searches were performed in the Cochrane Central Register of Controlled Trials (CENTRAL), MEDLINE and Embase databases. Pairs of authors (T.K., J.H., S.A., L.H., J.N., A.L., and J.J.) independently conducted the title and abstract screening, followed by full-text review to determine eligibility.

Data Extraction

Pairs of authors (T.K., J.H., S.A., L.H., J.N., A.L., and J.J.) independently extracted the data. We extracted the following characteristics from each study: (1) study population; (2) sample size; (3) condition for which surgery was performed; (4) type of surgery (open or minimally invasive); (5) primary outcome; and (6) whether success of blinding was considered. Pairs of authors assessed the risk of bias for each study using the Cochrane Risk of Bias Tool v. 1.0.¹⁴ All discrepancies were solved by discussion.

Choosing and Extracting Outcome Estimates

Depending on the study design, we extracted 2 or 3 within-group changes from each trial: (1) the observed change in the active surgery group (total effect of surgery); (2) observed change in the placebo surgery control group (total non-specific effects); and (3) observed change in the nonoperative control group (Figure 1 and Table 1).

We used the same type of outcome from corresponding studies (ie, same condition and intervention) whenever possible, prioritizing primary outcomes and time point from placebo-controlled studies. Because the primary outcome sometimes varied in otherwise similar placebo-controlled studies, we created a preferential hierarchy a priori: (1) a measure of pain; (2) function; (3) disease-specific quality of life; or (4) generic quality of life. To mitigate possible bias related to our choice of outcome, we performed sensitivity analyses with the prespecified primary outcomes in the included trials. When the time point was not defined or was irregular across studies, we selected a time point that likely reflected the timing of the maximum benefit of the surgical intervention being assessed.

Data Handling and Statistical Analyses

We extracted the mean (SD) change within each group. If change scores were not reported, we calculated change using baseline and follow-up means and SDs.¹⁵ We used the median (IQR) as an approximation of mean when it was singularly reported. When unavailable, we estimated SD using 95% CIs, reported *P* values, (IQR) divided by 1.35 or range (range divided by 4).¹⁶ If we could not identify any data to estimate SD, we imputed SD from corresponding studies using the same outcome.

For binary outcomes, we extracted data on the number of participants classified as improved by the study authors (ie, a good outcome). Odds ratios (ORs) for change in binary outcomes were calculated as a ratio of the number of patients with a good outcome at follow up, to the number of patients with a good outcome at baseline.¹⁷ To enable the calculation of an OR in the presence of a cell with 0 counts (patients with a good outcome at baseline), 0.5 was added to each cell. ORs were log-transformed and converted into effect sizes.¹⁸

We harmonized the direction of effect so that positive within-group change indicated improvement and positive between-group effect size indicated a better outcome for the active surgery group. The treatment effects (between-group differences) were based on the difference in the amount of change from baseline and expressed as Hedges *g*. Treatment effects were calculated for active vs placebo surgery; active surgery vs nonoperative; placebo vs nonoperative.

To estimate the proportion of non-specific effects of the total change after surgery (including the placebo effect), we calculated the proportion for each placebo-controlled study as within-group change in the placebo surgical group divided by within-group change in the active surgical group. In pooling, we used the log of the proportion.¹⁹ We also pooled the within-group improvements in placebo groups (estimate for the magnitude of total nonspecific effects) in a separate meta-analysis.

To estimate the magnitude of the surgical placebo effect, we pooled the differences between placebo and nonoperative control groups in studies that included both arms.

To estimate the association between type of the control group (placebo vs nonoperative) and the treatment effect, we performed metaregression using type of control as an effect modifier. Only those condition and interventions that had both placebo and nonoperative controlled studies were included. To see how other study characteristics impacted our analysis, we performed an additional analysis adjusting for possible differences in study characteristics, including sample size, subjective vs objective outcomes, and risk of bias (ie, high or unclear vs low for allocation concealment, attrition, and selective reporting).

Data were pooled using DerSimonian and Laird random effects meta-analyses. When 1 group contributed twice to the same meta-analysis, we halved the sample size for the active group.¹⁶

We assessed statistical heterogeneity using visual inspection and I^2 statistics. To assess heterogeneity of the metaregression analysis, we calculated the subgroup (placebo vs nonoperative control) differences in each condition and pooled these in a random effect meta-analysis. Significant statistical heterogeneity was defined as I^2 greater than 60% and $P < .05$. For publication bias, we performed an Egger test and assessed funnel plots when 10 or more studies were available for a condition and intervention group. Tests were 2-tailed, and analyses were conducted using STATA statistical software version 12.1 (StataCorp).

Results

Characteristics of Included Studies

We included data from 100 randomized controlled trials: 62 (62%) with placebo controls and 38 (38%) with nonoperative controls assessing the same surgical intervention for the same conditions as the placebo-controlled trials. Three trials included both placebo and nonoperative controls (eFigure 1 in the Supplement). Risk of bias was comparable across trials except for the risk of detection and performance bias which was high in the trials with nonoperative controls (Table 2; eTable 1 in the Supplement).

The included trials assessed 32 different interventions or conditions in 10 699 randomized participants. For 15 intervention and condition pairs, we had both placebo-controlled trials (n = 31) and nonoperative controlled trials (n = 38) with a total of 8089 participants (eTable 2 in the Supplement). Details of trials that contributed to each analysis are included in eAppendix 2 in the Supplement.

Table 2. Characteristics of Included Studies

Characteristics	Studies, No. (%)	
	Placebo-controlled	Nonoperative
Year of publication, median (IQR)	2011 (2005-2014)	2012 (2005-2014)
Studies ^a	62	41
Participants ^a	6122	4685
Participants/trial, median (IQR)	66 (43-130)	88 (53-139)
Proportion of females, %	68	60
Proportion of males, %	32	40
Self-reported outcome	47 (76)	35 (83) ^b
Proportion mini-invasive	55 (87)	37 (90)
Defined primary outcome	47 (76)	29 (71)
Low risk of allocation concealment	41 (65)	21 (51)
Low risk of attrition bias	52 (83)	28 (68)
Low risk of selective reporting	26 (41)	14 (34)
Measured success of blinding	13 (21)	NA

^a Studies that had both placebo and nonoperative controls are tallied in both columns. Studies that were excluded from the meta-analyses are not included in these numbers.

^b Alkatout et al (eAppendix 2 in the Supplement) were included in 2 condition-intervention pairs and measured both objective (ie, pregnancy) and subjective (ie, pain) outcomes.

Outcomes of included trials were pain (52 trials [52%]), function-related scores (32 [32%]); weight loss (6 [6%]); pregnancy (5 [5%]); and presence of stress urinary incontinence after pelvic prolapse surgery (4 [4%]) (eTable 2 in the [Supplement](#)). The median (IQR) time point of outcome data extraction was 12 (3-12) months. Log transformation of a binary outcome to calculate standardized mean difference (SMD) was performed in 27 (27%) trials. The remainder reported continuous outcomes.

The Mean Proportion of Nonspecific Effects

The mean within group improvement in placebo arms (ie, total nonspecific effects) was SMD 0.74 (95% CI, 0.59 to 0.89; $I^2 = 80\%$). As a proportion of the total effect of surgery, total nonspecific effects contributed between 0.01 and 1, and the mean pooled proportion was 0.67 (95% CI, 0.61 to 0.74; median [IQR], 0.71 [0.35-0.94]) (**Figure 2**; eTable 3 in the [Supplement](#)). The sensitivity analysis showed a mean pooled proportion of 0.73 (95% CI, 0.67 to 0.80; median [IQR], 0.75 [0.41-0.95]) using the primary outcomes defined by the trialists. Heterogeneity varied greatly between various interventions ($I^2 = 0\%$ to 100%) (eTable 3 in the [Supplement](#)).

Magnitude of the Surgical Placebo Effect

In 3 trials with 556 participants that had both a placebo and nonoperative control,²⁰⁻²² the pooled difference between these groups was SMD 0.13 (95% CI, -0.26 to 0.51; $I^2 = 63\%$). In the sensitivity analysis, the pooled difference was similar (SMD, 0.10; 95% CI, -0.26 to 0.46).

Difference in the Effect of Surgery Between Placebo- and Nonoperative-Controlled Trials

Metaregression included 73 between-group (active vs control) effects from 31 placebo-controlled trials and 38 trials with nonoperative controls. Type of control was not associated with the treatment effect. The mean difference (MD) between treatment effect in trials with a placebo control compared with those with nonoperative controls was -0.09 (95% CI, -0.35 to 0.17). The negative value indicates that placebo-controlled studies reported smaller effects than trials with nonoperative controls (**Figure 3**; eTable 3 in the [Supplement](#)). In the sensitivity analysis, the MD was -0.14 (95% CI, -0.41 to 0.14), the adjusted metaregression showed an SMD of -0.11 (95% CI, -0.37 to 0.15), and the sensitivity data set SMD of -0.18 (95% CI, -0.47 to 0.12).

Substantial heterogeneity was observed for the effect of surgery ($I^2 = 89\%$), but the type of control group did not explain this heterogeneity. It was most likely explained by the different surgical interventions among the included studies. Meta-analysis of subgroup differences showed no statistical heterogeneity ($I^2 = 0\%$) (Figure 3).

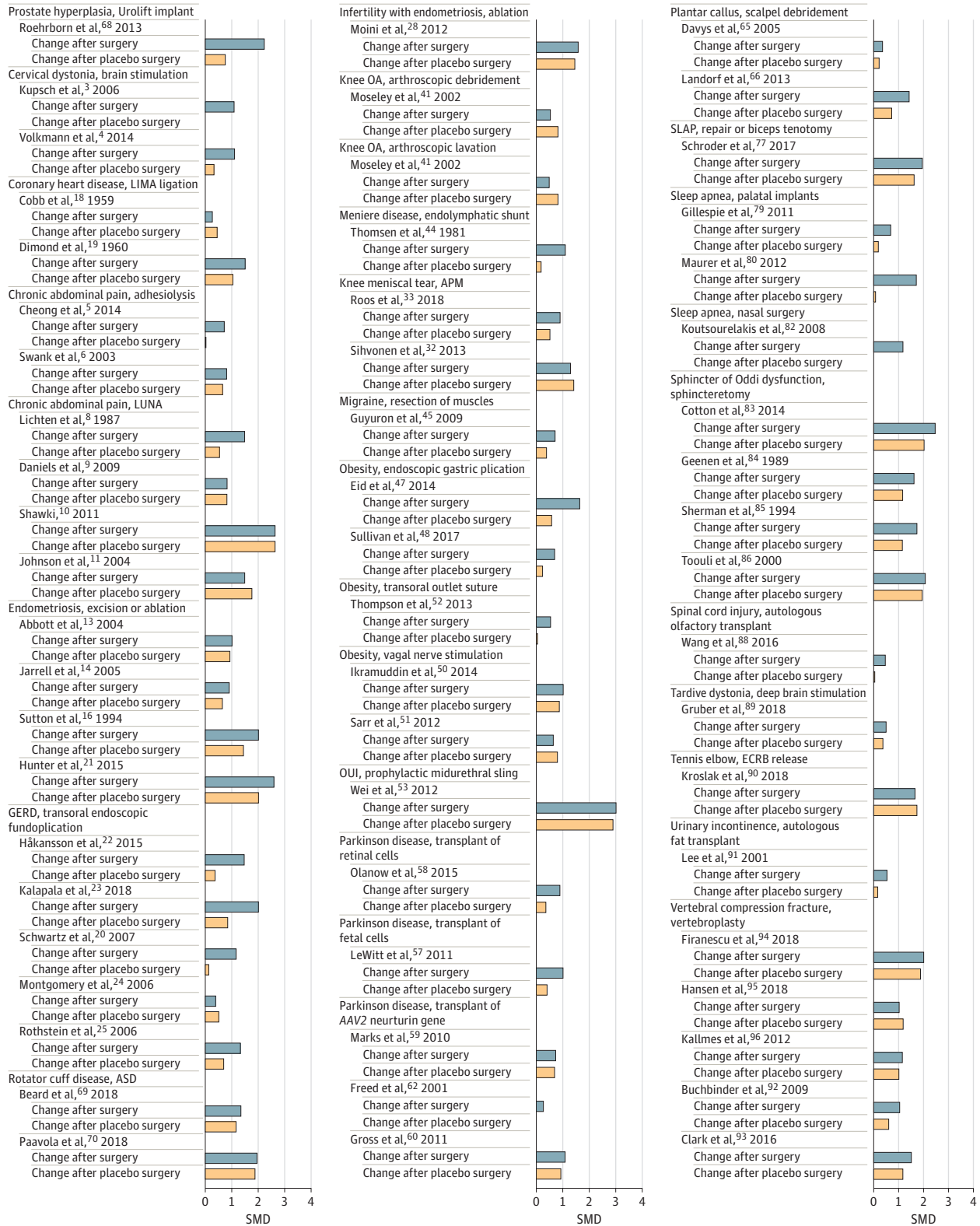
For 2 surgical interventions (vertebroplasty [12 trials] and prophylactic urethral sling [4 trials]), trials with nonoperative controls found larger benefits compared with placebo-controlled trials. Two conditions had the prerequisite number of studies (10 or more) for a separate metaregression. The SMD in the between-group effect size between placebo-controlled and nonoperative-controlled trials of subacromial decompression was 0 (95% CI, -0.8 to 0.8; 10 studies), and in vertebroplasty, it was 1.0 (95% CI, 0.18 to 1.9; 12 studies). The regression test for funnel plot asymmetry (performed in conditions with at least 10 trials) suggested no publication bias for studies within vertebroplasty ($P = .9$) or subacromial decompression trials ($P = .65$).

Discussion

Summary of Principal Findings

In this review, we did not find evidence supporting the hypothesis that surgery is associated with a large placebo effect. Trials with nonoperative controls were not associated with systematically larger effects of surgery compared with those using placebo controls. In studies that include both types of control, placebo groups did not improve more when compared with nonoperative controls.

Figure 2. Observed Within Group Effect Sizes (Hedges g) in the Included Placebo-Controlled Trials in the Active Surgery and Placebo Surgery Groups



APM indicates arthroscopic partial meniscectomy; ASD, arthroscopic subacromial decompression; ECRB, extensor carpi radialis brevis; GERD, gastroesophageal reflux disease; LIMA, left internal mammary artery; LUNA, laparoscopic uterosacral nerve

ablation; OA, osteoarthritis; OUI, occult urinary incontinence; SLAP, superior labrum tear anterior to posterior. The reference numbers refer to eAppendix 2 in the Supplement full list of included studies.

Two-thirds of the total effects of surgery within the included trials comprised nonspecific effects. The estimated magnitude of the surgical placebo effect was only one-seventh of the total nonspecific effects; the metaregression (an indirect estimate of surgical placebo effect) suggested a similarly small placebo effect. This indicates that, in surgical placebo-controlled trials conducted to date, the large changes seen after placebo surgery²³ occur mainly because of other nonspecific effects (ie, regression to the mean and natural course of disease).

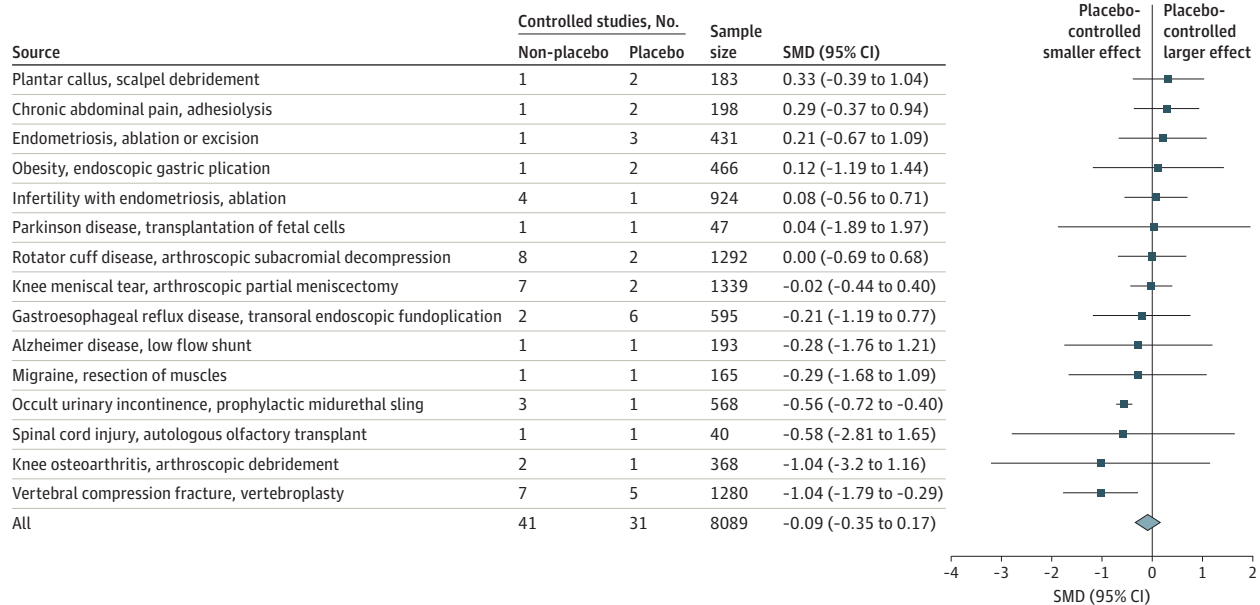
It is difficult to define or measure the placebo effect precisely,²⁴ but the important question is whether trialists should choose nonoperative or placebo controls for trials assessing the effect of surgery. Trials with nonoperative controls did not systematically overestimate effects of surgery (ie, the surgical placebo effect is not large enough to substantially bias the results). We found significant discrepancies in results between placebo and nonoperative controlled trials for only 2 interventions: vertebroplasty for osteoporotic vertebral fractures, where the primary outcomes were measured relatively early following the procedure (1 month), and urethral sling for occult urinary incontinence.

According to studies examining psychological mechanisms of the placebo effect, the placebo effect is based on expectations created by conditioning, verbal, or social cues.²⁵⁻²⁷ Conditioning plausibly does not occur in a single shot therapy, such as surgery, but regardless, conditioning may exist based on earlier experiences of receiving care. However, the theoretical model of conditioning predicts that the response will disappear if it is not maintained by new stimuli from receiving care.^{28,29} Pavlov's dogs stopped salivating without new stimuli, and we hypothesize that people do not keep feeling better because they received a single conditioned signal (eg, surgery) months or years ago. A washout of the placebo effect is also supported by empirical findings. In drug experiments, the placebo effect disappears just like the effect of a drug within hours after administration,³⁰ and in acupuncture studies, the placebo effect peaked at 12 weeks and disappeared by 52 weeks.³¹

Findings in Respect to Existing Evidence

Like previous studies, we found large improvements after placebo surgery.^{7,8,12,23} However, this observed improvement should not be considered the placebo effect because it includes all

Figure 3. Subgroup Differences Between Placebo-Controlled and Non-Placebo Controlled Trials



The effect of surgery was significantly smaller in placebo-controlled trials of vertebroplasty and prophylactic sling for urinary incontinence.

nonspecific effects. Our analyses suggest that these improvements are largely related to factors other than the placebo effect.

Previous evidence on participant blinding through placebo procedure is conflicting. A systematic review³² assessing the effect of participant blinding found a mean difference of 0.37 SMD for subjective outcomes but with 95% CIs consistent with a null effect (−0.77 to 0.04). A recent meta-epidemiological study³³ pooling data from 18 Cochrane reviews (138 trials) found no evidence that blinding affects the estimates (ratio of odd ratios [ROR], 0.98). Our study looking specifically at placebo controls in surgery supports these findings.

Contrary evidence includes the BRANDO study, which pooled data from 51 meta-analyses (530 trials), where the authors found a ROR of 0.78 (95% credible intervals, 0.65-0.92) for subjective outcomes suggesting that blinding attenuates the effect size.³⁴ The Risk of Bias in Evidence Synthesis study pooled data from 81 meta-analyses (1017 trials) and found a similar ROR of 0.83 (95% credible intervals, 0.73-0.93).³⁵ These studies suggest that bias may occur, but they did not assess surgery. Regarding our findings of nonspecific effects, our results are in line with a previous study assessing the nonspecific effects of 12 different mini-invasive procedures.¹²

Limitations

This study had limitations. The relatively small number of published placebo-controlled surgical trials and heterogeneity in the magnitude of nonspecific effects across the included trials limit the certainty of our estimates. Based on the results, we cannot reject the existence of clinically important placebo effects for all types of surgery. The observed large proportion of nonspecific benefits is likely related to both the conditions and the interventions being assessed using a placebo-controlled trial design. It is likely that surgical treatments of doubtful efficacy, coupled with conditions that have a self-limiting natural course, are the ones most likely to be evaluated in placebo-controlled trials and thus meet our inclusion criteria. For example, we did not find placebo-controlled trials investigating surgical excision of malignant skin tumors which would otherwise continue to grow.

Various nonoperative cointerventions in the control arms in nonoperative controlled studies may confound our analyses but do not alter our interpretation of the findings that nonoperative controlled studies should not be routinely considered biased because of the surgical placebo effect. Although nonoperative comparators often contained active cointerventions (eg, physical therapy, injections, etc), surgery is seldom used alone, and postoperative protocols included similar components to control arms. We specifically excluded studies in which the nonoperative interventions were not plausibly available for the surgical group to minimize bias arising from nonoperative interventions.

The small number of studies including both placebo and nonoperative controls limits the certainty of the estimate of magnitude of surgical placebo effect. In theory, this comparison may overestimate the placebo effect. Participants not receiving placebo surgery are aware of the allocation and may not similarly adhere to postoperative rehabilitation and may report worse outcomes because of having lower expectations.²⁴

By using the comprehensive searches of corresponding Cochrane reviews, we likely identified all relevant studies, but it is still possible we missed eligible trials. Our search was conducted in March 2019 but given our research question was related to the use of surgical placebo (rather than a clinical research question), it is unlikely additional recent trials will affect our conclusion.

Conclusions

The findings of this systematic review suggest that although the difference in the effect size between placebo-controlled and nonoperative controlled trials was small, placebo-controlled trials yield the least biased results. However, a difference of SMD equal to 0.09 translates to only 0.2 points in pain measured on a 0 to 10 scale (using a typical SD of 2); and the upper boundary of the 95% CI (0.35 corresponds with 0.7 of 10 points) is well below what is considered clinically meaningful. When

surgical trials with nonoperative controls have demonstrated consistently that a particular surgery does not provide clinically meaningful benefits, placebo-controlled trials are redundant. It is implausible that absence of a placebo control would hide a clinically important effect. However, as shown in the vertebroplasty and prophylactic sling for urinary incontinence trials, placebo-controlled trials are important when nonblinded trials show inconsistent results or the evidence suggests that the benefits may be small. Finally, the lack of a large placebo effect in surgery means that the use of ineffective surgical interventions for a placebo effect cannot be justified.

ARTICLE INFORMATION

Accepted for Publication: June 9, 2022.

Published: July 27, 2022. doi:10.1001/jamanetworkopen.2022.23903

Open Access: This is an open access article distributed under the terms of the [CC-BY License](#). © 2022 Karjalainen T et al. *JAMA Network Open*.

Corresponding Authors: Teemu Karjalainen, MBBS, PhD, Hospital Nova of Central Finland, Hoitajantie 3, Jyväskylä, 40620 Finland (teemukarjalainen@me.com); Sam Adie, MBBS, MSpMed, MPH, PhD, St George and Sutherland Clinical School, UNSW Sydney, Kogarah, NSW 2217, Australia (sam.adie@gmail.com).

Author Affiliations: Department of Surgery, Hospital Nova Central Finland, Jyväskylä, Finland (Karjalainen); Monash Department of Clinical Epidemiology, Cabrini Institute, Malvern, Australia (Karjalainen); School of Public Health and Preventive Medicine, Department of Epidemiology and Preventive Medicine, Monash University, Malvern, Australia (Karjalainen); Department of Orthopaedics, Oulu University Hospital, Oulu, Finland (Heikkinen); Biostatistical Consulting Platform, School of Public Health and Preventive Medicine, Monash University, Melbourne, Australia (Busija); Tampere University and Tampere University Hospital, Tampere, Finland (Jokihaara); Ingham Institute for Applied Medical Research, South Western Sydney Clinical School, UNSW Sydney, Liverpool, New South Wales, Australia (Lewin); South Western Sydney Clinical School, UNSW, Liverpool Hospital, New South Wales, Australia (Naylor); Sydney Orthopaedic Trauma and Reconstructive Surgery, Kogarah, New South Wales, Australia (L. Harris); Ingham Institute for Applied Medical Research, South Western Sydney Clinical School, UNSW Sydney, Liverpool, New South Wales, Australia (I. A. Harris); Department of Epidemiology and Preventive Medicine, School of Public Health and Preventive Medicine, Monash University (Buchbinder); Monash Department of Musculoskeletal Health and Clinical Epidemiology, Cabrini Health, Melbourne, Australia (Buchbinder); St George and Sutherland Clinical School, UNSW Sydney, Kogarah, New South Wales, Australia (Adie).

Author Contributions: Drs Karjalainen and Adie had full access to all of the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis.

Concept and design: Karjalainen, Busija, I. Harris, Buchbinder, Adie.

Acquisition, analysis, or interpretation of data: Karjalainen, Heikkinen, Busija, Jokihaara, Lewin, Naylor, L. Harris, Buchbinder, Adie.

Drafting of the manuscript: Karjalainen, Heikkinen, Jokihaara, Adie.

Critical revision of the manuscript for important intellectual content: Karjalainen, Busija, Jokihaara, Lewin, Naylor, L. Harris, I. Harris, Buchbinder, Adie.

Statistical analysis: Karjalainen, Busija, Adie.

Obtained funding: Adie.

Administrative, technical, or material support: Adie.

Supervision: I. Harris.

Data entry: Naylor.

Conflict of Interest Disclosures: Dr Karjalainen reported receiving grants from the Finnish Medical Foundation during the conduct of the study. Dr Buchbinder reported receiving support from an Australian National Health and Medical Research fellowship during the conduct of the study. Dr Busija reported receiving consulting fees from the Charité Universitätsmedizin Berlin, Jesuit Social Services Victoria, and Swinburne University of Technology outside the submitted work. No other disclosures were reported.

Additional Contributions: We thank Andrew Barnett, BEng, BSc (SitoWise, UNSW Australia) for his assistance in handling the data and Vieda Lusa, MBio, (Department of Surgery, Hospital Nova Central Finland, Jyväskylä, Finland) for help in preparing and formatting the figures and tables for the manuscript. Neither received compensation for their time.

REFERENCES

1. Ernst E, Resch KL. Concept of true and perceived placebo effects. *BMJ*. 1995;311(7004):551-553. doi:10.1136/bmj.311.7004.551
2. Wartolowska K, Judge A, Hopewell S, et al. Use of placebo controls in the evaluation of surgery: systematic review. *BMJ*. 2014;348:g3253. doi:10.1136/bmj.g3253
3. Schünemann H, Brozek J, Guyatt G, Oxman A; The GRADE Working Group. GRADE Handbook for grading quality of evidence and strength of recommendations. Accessed June 29, 2022. <https://gdt.gradepro.org/app/handbook/handbook.html>
4. Harris IA. In: Harris IA, ed. *Surgery, The Ultimate Placebo: A Surgeon Cuts through the Evidence*. University of New South Wales Press; 2016.
5. Kaptchuk TJ, Goldman P, Stone DA, Stason WB. Do medical devices have enhanced placebo effects? *J Clin Epidemiol*. 2000;53(8):786-792. doi:10.1016/S0895-4356(00)00206-7
6. Beard DJ, Campbell MK, Blazeby JM, et al. Considerations and methods for placebo controls in surgical trials (ASPIRE guidelines). *Lancet*. 2020;395(10226):828-838. doi:10.1016/S0140-6736(19)33137-X
7. 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*. 2015;5:e009655. doi:10.1136/bmjopen-2015-009655
8. Gu AP, Gu CN, Ahmed AT, et al. Sham surgical procedures for pain intervention result in significant improvements in pain: systematic review and meta-analysis. *J Clin Epidemiol*. 2017;83:18-23. doi:10.1016/j.jclinepi.2016.12.010
9. Beecher HK. Surgery as placebo: a quantitative study of bias. *JAMA*. 1961;176:1102-1107. doi:10.1001/jama.1961.63040260007008
10. Louw A, Diener I, Fernández-de-Las-Peñas C, Puentedura EJ. Sham surgery in orthopedics: a systematic review of the literature. *Pain Med*. 2017;18(4):736-750. doi:10.1093/pm/pnw164
11. Probst P, Grummich K, Harnoss JC, et al. Placebo-controlled trials in surgery: a systematic review and meta-analysis. *Medicine (Baltimore)*. 2016;95(17):e3516. doi:10.1097/MD.0000000000003516
12. Holtedahl R, Brox JI, Tjomsland O. Placebo effects in trials evaluating 12 selected minimally invasive interventions: a systematic review and meta-analysis. *BMJ Open*. 2015;5(1):e007331. doi:10.1136/bmjopen-2014-007331
13. Page M, McKenzie J, Bossuyt P, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. 2020. Accessed July 8, 2022. <https://www.equator-network.org/reporting-guidelines/prisma/>
14. Higgins JPT, Savović J, Page MJ, Elbers RGSJ. Assessing risk of bias in a randomized trial. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJWV, eds. *Cochrane Handbook for Systematic Reviews of Interventions Version 6.1 (Updated September 2020)*. Cochrane; 2020:8. Accessed June 29, 2022. <https://training.cochrane.org/handbook/current/chapter-08>
15. Higgins JPT, Li T, Deeks JJ. Choosing effect measures and computing estimates of effect. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJWV, eds. *Cochrane Handbook for Systematic Reviews of Interventions Version 6.1 (Updated September 2020)*. Cochrane; 2020:8. Accessed June 29, 2022. <https://training.cochrane.org/handbook/current/chapter-06>
16. Weir CJ, Butcher I, Assi V, et al. Dealing with missing standard deviation and mean values in meta-analysis of continuous outcomes: a systematic review. *BMC Med Res Methodol*. 2018;18(1):25. doi:10.1186/s12874-018-0483-0
17. Hirji KF, Fagerland MW. Calculating unreported confidence intervals for paired data. *BMC Med Res Methodol*. 2011;11:66. doi:10.1186/1471-2288-11-66
18. Chinn S. A simple method for converting an odds ratio to effect size for use in meta-analysis. *Stat Med*. 2000;19(22):3127-3131. doi:10.1002/1097-0258(20001130)19:22<3127::AID-SIM784>3.0.CO;2-M
19. Hedges LV, Gurevitch J, Curtis PS. The meta-analysis of response ratios in experimental ecology. *Ecology*. 1999;80(4):1150-1156. doi:10.1890/0012-9658(1999)080[1150:TMAORR]2.0.CO;2
20. Paavola M, Malmivaara A, Taimela S, et al; Finnish Subacromial Impingement Arthroscopy Controlled Trial (FIMPACT) Investigators. Subacromial decompression versus diagnostic arthroscopy for shoulder impingement: randomised, placebo surgery controlled clinical trial. *BMJ*. 2018;362:k2860. doi:10.1136/bmj.k2860
21. Beard DJ, Rees JL, Cook JA, et al; CSAW Study Group. Arthroscopic subacromial decompression for subacromial shoulder pain (CSAW): a multicentre, pragmatic, parallel group, placebo-controlled, three-group, randomised surgical trial. *Lancet*. 2018;391(10118):329-338. doi:10.1016/S0140-6736(17)32457-1

22. 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(1):20-28. doi:10.1136/gut.2006.096842
23. Wartolowska KA, Gerry S, Feakins BG, et al. A meta-analysis of temporal changes of response in the placebo arm of surgical randomized controlled trials: an update. *Trials*. 2017;18(1):323. doi:10.1186/s13063-017-2070-9
24. Hróbjartsson A. What are the main methodological problems in the estimation of placebo effects? *J Clin Epidemiol*. 2002;55(5):430-435. doi:10.1016/S0895-4356(01)00496-6
25. Colagiuri B, Schenk LA, Kessler MD, Dorsey SG, Colloca L. The placebo effect: From concepts to genes. *Neuroscience*. 2015;307:171-190. doi:10.1016/j.neuroscience.2015.08.017
26. Colloca L, Benedetti F. How prior experience shapes placebo analgesia. *Pain*. 2006;124(1-2):126-133. doi:10.1016/j.pain.2006.04.005
27. Colloca L, Miller FG. How placebo responses are formed: a learning perspective. *Philos Trans R Soc Lond B Biol Sci*. 2011;366(1572):1859-1869. doi:10.1098/rstb.2010.0398
28. Walach H, Maidhof C. Is the placebo effect dependent on time: a meta-analysis. In: Kirsch I, ed. *How Expectancies Shape Experience*. American Psychological Association; 2004:321-332. doi:10.1037/10332-013
29. André-Obadia N, Magnin M, Garcia-Larrea L. On the importance of placebo timing in rTMS studies for pain relief. *Pain*. 2011;152(6):1233-1237. doi:10.1016/j.pain.2010.12.027
30. Lasagna L, Laties VG, Dohan JL. Further studies on the pharmacology of placebo administration. *J Clin Invest*. 1958;37(4):533-537. doi:10.1172/JCI103635
31. Koog YH, Jung WY. Time course of placebo effect of acupuncture on pain: a systematic review. *ISRN Pain*. 2012;2013:204108. doi:10.1155/2013/204108
32. Page MJ, Higgins JPT, Clayton G, Sterne JAC, Hróbjartsson A, Savović J. Empirical evidence of study design biases in randomized trials: systematic review of meta-epidemiological studies. *PLoS One*. 2016;11(7):e0159267. doi:10.1371/journal.pone.0159267
33. Moustgaard H, Clayton GL, Jones HE, et al. Impact of blinding on estimated treatment effects in randomised clinical trials: meta-epidemiological study. *BMJ*. 2020;368:l6802. doi:10.1136/bmj.l6802
34. Savović J, Jones HE, Altman DG, et al. Influence of reported study design characteristics on intervention effect estimates from randomized, controlled trials. *Ann Intern Med*. 2012;157(6):429-438. doi:10.7326/0003-4819-157-6-201209180-00537
35. Savovic J, Turner RM, Mawdsley D, et al. Association between risk-of-bias assessments and results of randomized trials in Cochrane Reviews: the ROBES meta-epidemiologic study. *Am J Epidemiol*. 2018;187(5):1113-1122. doi:10.1093/aje/kwx344

SUPPLEMENT.

eFigure. PRISMA Flowchart

eTable 1. Risk of Bias Judgements

eTable 2. Conditions, Interventions, Number of Studies, the Extracted Outcomes, and Timepoints Where Outcomes Were Extracted

eTable 3. Proportion of Nonspecific Effects for Each Study and Condition or Intervention Subgroup

eTable 4. Effect Sizes in Studies Used in Meta-Regression Comparing Placebo and Nonplacebo Studies

eAppendix 1. Search Strategies

eAppendix 2. List of Included Trials