BMJ Open

Placebo effects in trials evaluating 12 selected minimally invasive interventions: a systematic review and meta-analysis

Robin Holtedahl,[1] Jens Ivar Brox,[2] Ole Tjomsland[3]

To cite: 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**: e007331. doi:10.1136/bmjopen-2014-007331

▸ Prepublication history and additional material is available. To view please visit the journal (http://dx.doi.org/10.1136/bmjopen-2014-007331).

Received 29 November 2014
Revised 18 December 2014
Accepted 22 December 2014

For numbered affiliations see end of article.

Correspondence to
Dr Robin Holtedahl;
robi-hol@online

ABSTRACT

Objectives: To analyse the impact of placebo effects on outcome in trials of selected minimally invasive procedures and to assess reported adverse events in both trial arms.

Design: A systematic review and meta-analysis.

Data sources and study selection: We searched MEDLINE and Cochrane library to identify systematic reviews of musculoskeletal, neurological and cardiac conditions published between January 2009 and January 2014 comparing selected minimally invasive with placebo (sham) procedures. We searched MEDLINE for additional randomised controlled trials published between January 2000 and January 2014.

Data synthesis: Effect sizes (ES) in the active and placebo arms in the trials' primary and pooled secondary end points were calculated. Linear regression was used to analyse the association between end points in the active and sham groups. Reported adverse events in both trial arms were registered.

Results: We included 21 trials involving 2519 adult participants. For primary end points, there was a large clinical effect (ES≥0.8) after active treatment in 12 trials and after sham procedures in 11 trials. For secondary end points, 7 and 5 trials showed a large clinical effect. Three trials showed a moderate difference in ES between active treatment and sham on primary end points (ES ≥0.5) but no trials reported a large difference. No trials showed large or moderate differences in ES on pooled secondary end points. Regression analysis of end points in active treatment and sham arms estimated an R^2 of 0.78 for primary and 0.84 for secondary end points. Adverse events after sham were in most cases minor and of short duration.

Conclusions: The generally small differences in ES between active treatment and sham suggest that non-specific mechanisms, including placebo, are major predictors of the observed effects. Adverse events related to sham procedures were mainly minor and short-lived. Ethical arguments frequently raised against sham-controlled trials were generally not substantiated.

Strengths and limitations of this study

- Selection of trials with low risk of bias.
- Calculation of effect sizes on primary and pooled secondary end points in active treatment as well as sham arms.
- Heterogeneous interventions, outcome measures and timing of assessment.

INTRODUCTION

It is normally assumed that medical practices are based on firm clinical evidence and that new practices or techniques are introduced when superiority, or at least non-inferiority, has been demonstrated compared to established treatments. However, medical history reveals numerous examples contradicting this assumption. Forty-two per cent of 146 medical practices were found to be reversed in a recent review analysing 10 years of publication in a high-impact medical journal.[1] Large effects of an intervention in initial reports are often spurious findings, while the vast majority may represent substantial overestimations.[2]

Even though surgical and other invasive techniques generally have reached a high degree of sophistication through the last decades, not all invasive procedures have lived up to expectations. Promising results in initial observational studies have in some cases led to widespread clinical implementation, in spite of lack of documented effectiveness.[3] The reluctance to abandon contradicted medical practice is commonly ascribed to both culturally embedded medical practices and different forms of vested interests.[4 5] The continuation of unnecessary and potentially harmful interventions leads to major costs for both patients and society.

The randomised placebo-controlled trial is considered the gold standard for evaluating the effects of pharmacological treatments. However, there are relatively few controlled studies in peer-reviewed surgical journals and even fewer placebo (sham)-controlled studies.[6–8] Ethical concerns raised by the potential for harm to participants are usually cited as the main obstacle to sham-controlled

studies.[9] Problems of a practical nature relate to patient blinding, differing technical expertise, the heterogeneity of the interventional techniques and variable outcome specifications, making standardisation difficult to achieve.[10]

A meaningful effect in clinical trials may result from a large effect in the active treatment group, a small effect in the placebo group or a combination of both. Even though a placebo effect has been documented in a range of clinical conditions, there are few studies assessing the magnitude of the placebo effect in surgical procedures. In the present study, we analysed placebo-controlled trials of minimally invasive interventions in musculoskeletal, neurological and cardiac conditions. The aims were threefold: (1) to assess the magnitude of change in outcome from baseline to trial end point in the active treatment as well as placebo (sham) arms; (2) to explore the contribution of non-specific factors, including placebo, to the outcome of active treatment and (3) to assess the level of reported adverse effects in both trial arms.

METHODS

Search strategy and selection criteria

We conducted electronic searches for randomised placebo-controlled trials of minimally invasive interventions for cardiac, neurological and musculoskeletal conditions. We defined minimally invasive procedures as interventions involving the introduction of a medical device, substance or other foreign material into the body through a cannula, catheter or arthroscope, thereby minimising damage to biological tissues at the point of entrance. We first used MEDLINE and Cochrane library to identify systematic reviews published between January 2009 and January 2014. The following key words were used in our search strategies: "randomi* controlled trial", "placebo OR sham" in combination with "low back pain", "neck OR cervical pain", "radiofrequency denervation", "facet joint AND "nerve block" OR injection", "intradiscal OR annular AND thermal", "epidural AND corticosteroid* AND sciatica OR radic*", "hyaluron* OR viscosuppl* AND knee AND osteoarthritis", "vertebroplast*", "arthroscop*", "debridement AND lavage AND knee AND osteoarthr*", "meniscectomy AND knee", "myocardial laser revascularization", "transplantation OR gene OR stem cell OR deep brain stimulation AND Parkinson* OR dystonia", "spinal cord stimulation", and "foramen ovale AND migraine". We used the "core clinical journals" filter in PubMed, which is an index of journals particularly relevant to practicing physicians.

From the most recently published systematic review of each procedure, we selected randomised placebo-controlled trials published later than January 2000. We excluded trials published before January 2000 because our primary aim was to assess interventions that are currently, or until recently have been, in common use. We selected trials that according to the review fulfilled at least four of the following methodological criteria: random allocation, allocation concealment, blinding of participant, blinding of assessor and intention-to-treat analysis. We chose these criteria because they were the most commonly used in the selected reviews, and because use of scales for assessing quality or risk of bias is explicitly discouraged in Cochrane reviews.[11] Two of the authors (RH and JIB) independently assessed the five methodological criteria in the RCTs included from systematic reviews.

We next searched MEDLINE for additional randomised placebo-controlled trials published between January 2000 and January 2014. Two of the reviewers (OT and JIB) independently assessed the five criteria aforementioned in the additional RCTs that were identified from this search.

Only English language journals were included. We excluded crossover trials, trials that did not report results as means, SD, SE or CIs in active and sham groups, as well as trials with only graphic representation of data. This review is reported in accordance with the PRISMA statement.[12]

Data extraction

We registered all continuous primary end points. In trials without continuous primary end points, with multiple end points or no defined primary end point, we selected an outcome related to pain or condition-specific end point. The heterogeneity of trials did not allow for use of pain as a primary outcome. We used the RCTs' defined primary outcome to avoid bias introduced by choosing our own end point. We also registered secondary end points in order to avoid potential bias from selective reporting in the included trials. End points describing medication, radiographic or physiological variables, social or psychological functions were not included. For the Parkinson's trials, only end points in the off-medication state were registered. Results from the last follow-up until 12 months were extracted. The trials' protocol registration, funding source, description of sham intervention, sample size, disease duration, length of follow-up and reported adverse events (AEs) in both trial arms were registered.

Data synthesis

To assess clinically important change, we calculated effect size (ES, Cohen's d) based on the means and SDs. We calculated ES for both the active and sham interventions to obtain information about the pre-to-post treatment change in both arms. Without first calculating ES of change in each trial arm, we would not be able to discern the relative contribution of placebo, which was one of the objectives of the study. Subtracting the average score after treatment from the average score before treatment and dividing the result by the average of the SDs before and after treatment calculates ES. An ES of 0.8 or more is assumed to be large, while an ES of 0.5–0.8 is considered moderate.[13] In trials with multiple secondary end points we calculated the pooled mean ES, without weighting.

Owing to the small sample sizes in most of the included trials, we calculated an adjusted ES in accordance with a recommended procedure.[14] Unadjusted linear regression analyses were used to explore the association between outcomes in the active and sham groups, both for primary and pooled secondary end points. For this analysis, we used Medcalc Statistical Software V.12.7.4.0.[15]

RESULTS

Selection of interventions and trials

The searches provided sham-controlled trials for the following interventions: percutaneous laser revascularisation of myocardium for angina pectoris (n=2), closure of foramen ovale for migraine (n=1), arthroscopic meniscectomy for meniscal tears (n=1), debridement (n=1) and injection of hyaluronic acid (n=3) for symptomatic osteoarthritis of the knee, injection or transplantation of biologically active material for Parkinson's disease (human retinal pigmental cells (n=1), fetal nigral cells (n=1) and neurturin (n=2)), epidural injections of corticosteroids for sciatica (caudal (n=1), interlaminar (n=2) and transforaminal (n=1)) routes, percutaneous heating of the intervertebral disc for chronic low back pain (intradiscal radiofrequency thermocoagulation (n=1), intradiscal electrothermal therapy (n=2)) and vertebroplasty for vertebral body fractures (n=2). We give a short description of each procedure's introduction, therapeutic rationale and history in web appendix table 1.

The searches provided no sham-controlled trials of cervical, thoracic or lumbar facet joint nerve blocks or joint injections, spinal cord stimulation for low back pain, cervical epidural injections, transmyocardial laser revascularisation for angina pectoris, deep brain stimulation for Parkinson's disease or dystonia or arthroscopic procedures other than knee conditions. We found six placebo-controlled trials of radiofrequency denervation for low back pain, but all were excluded as: SD not provided (n=1),[16] compound primary end point (n=1),[17] risk of false-positive response because of only one diagnostic block (n=4).[18–21]

The study selection process is summarised in figure 1. The search provided five systematic reviews, all identified through searches in MEDLINE; none were commercially funded.[22–26] It identified a total of 71 clinical trials, 12 of these were not identified from the systematic reviews. Forty-four trials were excluded for methodological reasons, principally risk of bias. Six additional trials were excluded because ES could not be calculated.[27–32] Web appendix table 2 shows the excluded trials and the reasons for exclusion. Finally, 21 clinical trials with a total of 2519 participants were included in the present review (table 1). Trial interventions in active treatment and sham arms are also shown.

Fourteen trials from the systematic reviews fulfilled at least four of the five methodological criteria.[33–46] Seven trials provided through searches in MEDLINE fulfilled the same criteria.[47–53] The included and excluded secondary end points are shown in web appendix table 3. All trials reported approval of study protocol prior to patient enrolment (table 1). Seven trials were commercially funded.[36 37 45 50–53] Most of the trials had few participants, ranging from 20 to 346 (median 80).

Clinical outcomes after active treatment and sham

Twelve of the 21 trials showed a large ES on primary end points after active treatment, while 11 trials showed a similar ES after the sham procedure (figure 2 and table 2).

ES on primary end points was moderate in three of the active treatment groups and in two of the sham groups.

On pooled secondary end points, a large ES was estimated in seven trials after active treatment and in five trials after sham, while a moderate ES was reported in four and three trials, respectively (table 2).

In none of the trials did the actively treated group show a deterioration of primary end point during treatment, while this was the case for two of the sham groups (not reported to be related to the procedure). On secondary end points, deterioration occurred in two active treatment and two sham groups (table 2).

Differences in outcome between active treatment and sham

Better results on primary end points were reported with active treatment compared to sham in 14 of the 21 trials, but the differences were small. Three trials (1 epidural study,[41] 1 discogenic pain study[44] and 1 Parkinson's study[52]) reported a moderate effect but none showed a large effect (figure 3 and table 2). Seven trials reported a better primary end point outcome after sham than after active treatment.

Nineteen trials reported secondary end points, 11 of these reported better outcome after active treatment than after sham, but in no case did the differences reach a moderate ES (figure 3 and table 2). In 12 trials, the outcome was better for primary than for pooled secondary end points. This bore no relation to funding source.

On regression analyses, ESs in the sham groups predicted about 80% of the variance of ES in the active treatment groups, both on primary and pooled secondary end points (figures 4 and 5).

Adverse events

Eighteen studies provided information about AEs (table 1). Three of these trials reported no procedural AEs in any of the groups.[33 39 47] Major AEs were reported after active treatment in four trials[34 50 51 53] including one death in one of the Parkinson's studies.[51] In the sham groups, one trial[53] listed three major AEs possibly or probably related to the procedure, all presumed to be caused by antiplatelet medication, none of them life-threatening. Apart from this trial, there were no major AEs in the sham groups. The reported minor AEs were all of limited duration.

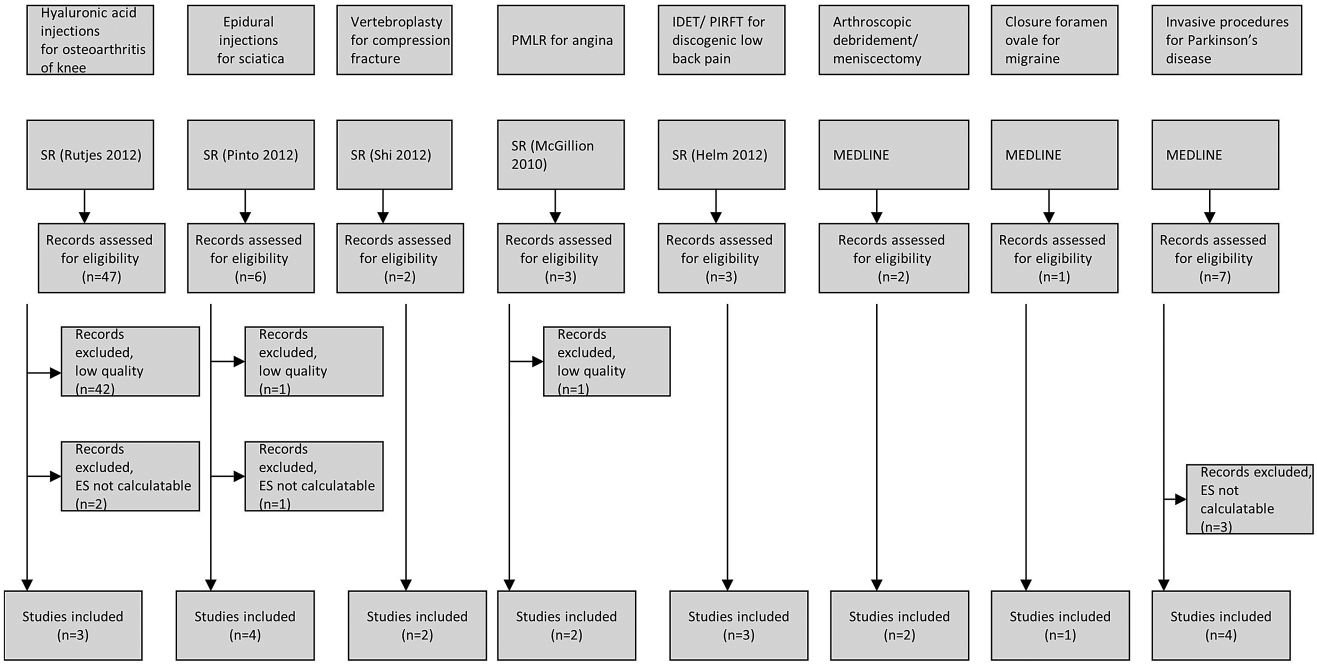

Figure 1 Flow chart of study selection in the present meta-analysis (ES, effect size; IDET, intradiscal electrothermal therapy; PIRFT, percutaneous intradiscal radiofrequency thermocoagulation; PMLR, percutaneous myocardial laser revascularisation; SR, systematic review).

DISCUSSION

Principal findings

Analysis of 21 sham-controlled trials of minimally invasive procedures showed that the ESs in the active treatment arms were predicted by the ESs in the sham arms. There was a large ES on primary end points in about half of both the active and sham interventions, but none of the trials showed a large difference in ES between active treatment and sham groups, either on primary or secondary end points.

The magnitude of the effect in each trial arm varied considerably, both between different procedures and between trials using the same procedure. For instance, in the active treatment groups, ES for primary end points varied from around 0 to almost 2 after active treatment and from about −0.4 to 1.5 after sham. Disparate outcomes were reported even between trials where technical parameters were similar. For instance, ES in the sham group in the three hyaluronic acid trials varied by a factor of 3 and in the epidural trials by a factor of 2. This variability is probably related to differences in study design, duration of disability before inclusion, contextual factors, including the doctor–patient relationship, as well as other factors. The close association between end points in the active treatment and sham groups on regression analyses suggests that a large part of the reported outcomes in the active treatment groups are due to placebo effects, statistical regression to the mean or the natural course of the condition.

Strengths and limitations of the study

It is our opinion that the calculation of ESs in active treatment as well as placebo arms is a strength of the present study. This made it possible to assess the magnitude of change in arms as well as the contribution of non-specific factors to change in the active treatment arms. The calculation of ESs provides an alternative assessment to probability estimates. Another strength of the study is the supplementary analyses of pooled secondary end points, enabling a more comprehensive evaluation than using primary end points alone. Reports of tactically motivated use of primary and secondary end points before publication in order to improve study results strengthen the argument for registering all relevant secondary end points.[54] Our finding that a majority of trials reported better results on primary than on secondary end points might lend support to such a hypothesis, although all trials, according to the authors, had sought and gained approval of the protocol from ethics committee and/or review board (table 1).

The present review is limited to selected minimally invasive procedures in cardiology, neurology and musculoskeletal conditions. While some procedures are or have been in wide clinical use, some are still in the clinical trial phase. Other sources of heterogeneity are variable duration of disease before inclusion, selection of outcome measures and time to follow-up. Results cannot be generalised to minimally invasive procedures in all medical disciplines, but a similar methodology could be applied to more systematic analyses of the role of non-specific effects in other minimally invasive procedures.

We applied principles from guidelines for conducting systematic reviews and meta-analyses, and included an independent assessment of methodological trial quality by two of the authors. We cannot rule out that we have missed relevant trials because we limited our search to

Table 1 Studies, protocol approval and funding, interventions in the active treatment and sham arms, and adverse events

Author	Protocol approval/ funding (commercial, non-commercial)	Invasive procedure/ indication	Sham intervention	Adverse events related to procedure, active treatment	Adverse events related to procedure, sham
Leon *et al*[34]	Food and Drug Administration/NC	Percutaneous myocardial laser revascularisation/ intractable angina pectoris	Laser turned on but no procedure performed	MAE in hospital (high dose): 4.1%	MAE in hospital: 0
Salem *et al*[33]	Ethics committee/NC			No procedural AE	
Sihvonen *et al*[48]	Review board/NC	Arthroscopic partial meniscectomy/degenerate meniscal tear	Routine arthroscopy, simulation of meniscectomy by manipulation etc	No MAE mAE: 6.6%	 mAE: 2.9%
Moseley *et al*[47]	Review board/NC	Arthroscopic debridement/ knee osteoarthritis	Simulated arthroscopy preparation, intravenous anaesthesia, skin incisions, no instruments entered knee, knee manipulated	No procedural AE	
Pham *et al*[35]	Review board/NC	Hyaluronic acid/knee osteoarthritis	Intra-articular injection of saline solution	No MAE Any mAE: 81.7%	 Any mAE: 1.2%
Altman *et al*[36]	Ethics committee/C			No MAE mAE: 12.8%	 mAE: 8%
Chevalier *et al*[37]	ClinicalTrials.org/C			No MAE mAE: 35.8%	 mAE: 33.8%
Kallmes *et al*[38]	Review board/NC	Percutaneous vertebroplasty with PMMA cement injection/ vertebral compression fracture	Conscious sedation + local anaesthaesia, pressure put on spine, simulation of odour with mixing of PMMA to imitate the smell during the active procedure	No MAE mAE: 14%	 mAE: 16%
Buchbinder *et al*[39]	Ethics committee at each participating center/NC		Conscious sedation + local anaesthaesia, needle inserted to rest on the lamina, PMMA container opened to imitate the smell during the active procedure	No procedural AE	
Cohen *et al*[41]	Review board/NC	Epidural injection of corticosteroids/sciatica	2 mL sterile water at 1–2 injection sites, transforaminal approach	No MAE mAE:36%	 mAE: 20%
Arden *et al*[42]	Ethics committee/NC		2 mL saline into interspinous ligament	No MAE mAE: 9%	 mAE: 10%
Valat *et al*[43]	Ethics committee/NC		2 mL saline into epidural space, interlaminar approach	No MAE mAE: 6%	 mAE: 8%
Iversen *et al*[40]	Ethics committee/NC		Subcutaneous injection of 2 mL saline superficial to the sacral hiatus	Not reported	
Freeman *et al*[45]	Ethics committee/C	IDET/discogenic low back pain	17-gauge introducer needle inserted into disc under fluoroscopic guidance, catheter inserted but not connected to generator, both subject and surgeon blinded	No MAE mAE: 11%	 mAE: 5%
Pauza *et al*[44]	Review board/NC		17-gauge needle introduced onto the outer annulus, mock electrode passage	Not reported	

Continued

Table 1 Continued

Author	Protocol approval/ funding (commercial, non-commercial)	Invasive procedure/ indication	Sham intervention	Adverse events related to procedure, active treatment	Adverse events related to procedure, sham
			shown on monitor, generator noises produced		
Kvarstein *et al*[46]		Ethics committee/NC	PIRFT/discogenic low back pain	17-gauge canula and RF probe inserted into annulus, no RF current applied	Not reported
Olanow *et al*[49]	Review board/NC	Fetal nigral transplantation, 4 donors/Parkinson's disease	Scalp incisions, partial thickness burr holes, no cell transplantation, 6 months low-dose cyclosporine	No MAE mAE (rate/patient day: 0.66	mAE (rate/patient day: 0.39
Marks *et al*[50]	Review board/C	Gene delivery of AAV2-neurturin/Parkinson's disease	Scalp incisions, partial thickness burr holes, no intracranial injections	MAE: 4 Most frequent mAE: headache: 68%	MAE: 0 Most frequent mAE: headache: 50%
Gross *et al*[51]	Review board/C	Transplantation of human retinal pigmental cells/ Parkinson's disease	Scalp incisions, partial thickness burr holes, no cell transplantation	1 death MAE: 23%	0 deaths MAE: 0
LeWitt *et al*[52]	Review board/C	Insertion of AAV-GAD gene into subthalamic nucleus/ Parkinson's disease	Insertion of catheter caudal to nucleus, infusion of saline	No MAE mAE (probably related to procedure): 56%	mAE (probably related to procedure): 14%
Dowson *et al*[53]	Ethics committee/C	Patent foramen ovale closure with STARFlex Septal Repair Implant/migraine	General anaesthesia, skin incision in the groin	MAE (possibly or probably related to procedure): 11%	MAE (possibly or probably related to procedure): 4%

AAV2, adeno-associated; C, commercial; GAD, glutamic acid decarboxylase; IDET, intradiscal electrothermal therapy; MAE, major adverse events; mAE, minor adverse events; NC, non-commerical; PIRFT, percutaneous intradiscal radiofrequency thermocoagulation; PMMA, polymethylmethacrylate.

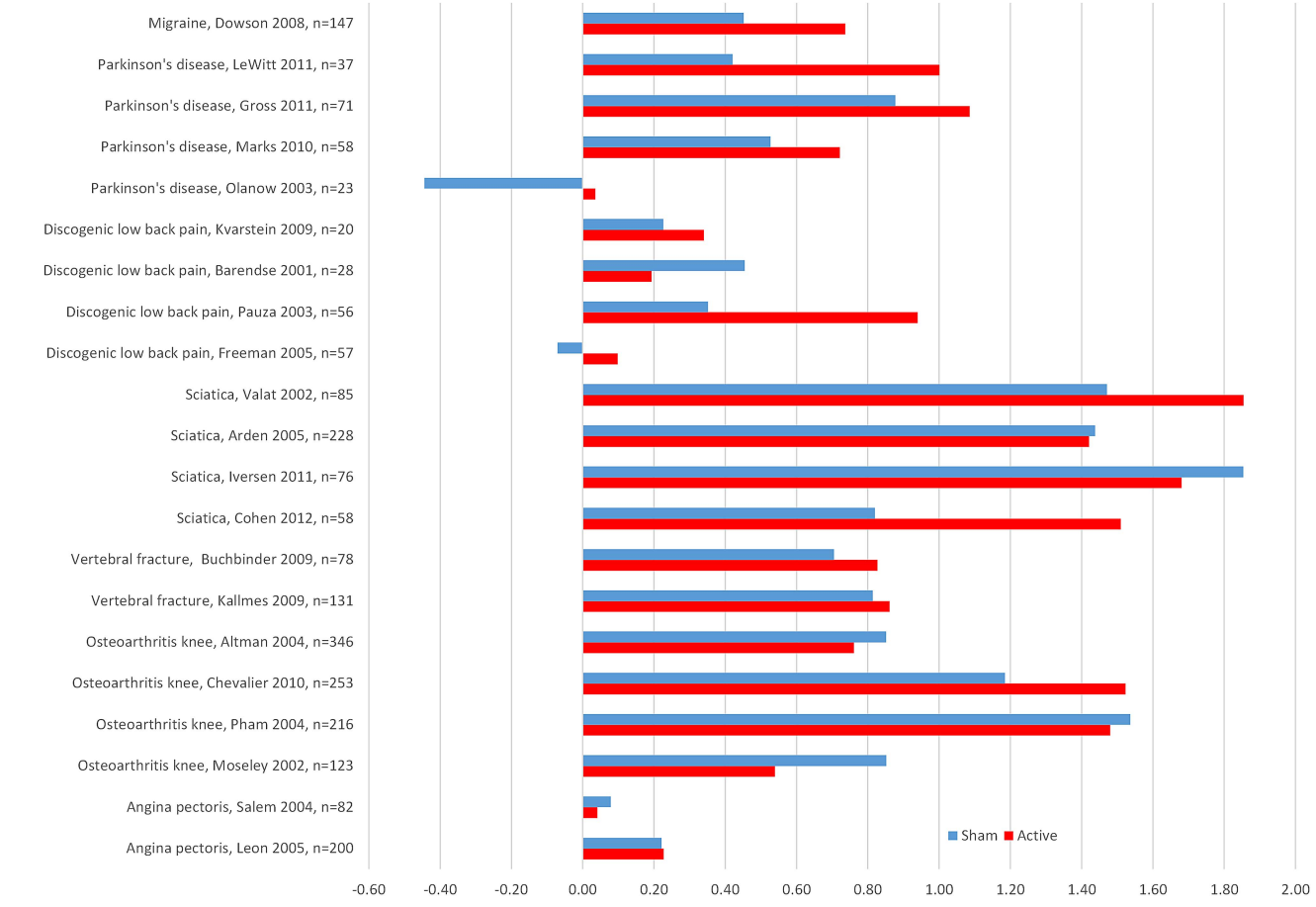

Figure 2 Effect sizes of active treatment and sham, primary end points.

the Cochrane Library and MEDLINE, but the most relevant trials are likely to have been identified by our searches. By preferentially selecting core journals and trials that had previously been methodologically evaluated in systematic reviews, it was our intention to reduce the risk of bias by excluding studies of low quality. We realise that this selection process and the fact that we relied on previous methodological evaluations may have contributed to unrecognised selection bias.

The use of ES as a measure of clinical effect assumes a normal distribution of the data. This does not necessarily apply in the included trials because the majority of them are small. Including trials reporting non-parametric data would, however, necessitate other methods of statistical analysis. Small studies increase the likelihood of type-2 errors, though this is more relevant to probability estimates than analysis of ES.

Adequate blinding and lack of physiological effects

We cannot rule out that treatment-specific effects in the actively treated groups may have jeopardised blinding, leading to overestimation of treatment effects through positive expectations. However, all the included trials gave a detailed description of the sham procedure, and both participant and assessor blinding seems to have been adequate.

On a more general level, it has been argued that sham procedures are not inert and may have specific physiological effects, thereby underestimating a treatment effect.[55] Recently, Bickett *et al* hypothesised that epidural injection of small volumes of saline might have physiological effects.[56] However, it is to be noted that in the four selected epidural trials in the present study, improvements in the sham group were greater in the two trials using non-epidural saline than in those using epidural saline, making a physiological effect less likely. In our opinion, physiological effects of the sham interventions are also unlikely in the remaining procedures.

Surgery and other invasive procedures are commonly believed to be associated with enhanced placebo effects, a phenomenon coined mega-placebo.[57] In spite of their heterogeneous nature, the 21 selected trials share a medicotechnological context in which an a priori enhanced placebo response could be expected. If an ES >0.8 is considered as mega-placebo, half of the included sham interventions reached this level. Factors such as the level of enthusiasm and conviction conveyed by the therapist, the impression of advanced procedures and the extent to which these factors succeed in activating a placebo response are probably crucial in explaining the improvements after sham interventions and the correlation of end points in the active treatment and sham

Table 2 Effect size (ES) on primary and pooled secondary end points, showing differences between active treatment and sham arms.

Author/procedure	Limit disease duration/time to follow-up (months)	Trial arm/number of patients randomised	ES primary end point	ES pooled secondary end points (number of end points)
Leon 2005/percutaneous myocardial laser revascularisation	None/12		Exercise duration (s)	(10)
		Active/98	0.23	0.60
		Sham/102	0.22	0.54
ES active treatment vs sham			**0.01**	**0.07**
Salem 2004/percutaneous myocardial laser revascularisation	None/12		Exercise duration (s)	–
		Active/40	0.04	
		Sham/42	0.08	
ES active treatment vs sham			**−0.04**	
Sihvonen 2013/arthroscopic partial meniscectomy	>3/12		Lysholm knee score	(4)
		Active/70	0.86	0.58
		Sham/76	1.03	0.58
ES active treatment vs sham			**−0.17**	**0.00**
Moseley 2002/arthroscopic debridement	None/12		Knee Specific Pain Scale	(5)
		Active/59	0.54	0.11
		Sham/60	0.85	0.20
ES active treatment vs sham			**−0.31**	**−0.09**
Pham 2004/hyaluronic acid			VAS pain	(3)
	None/12	Active/131	1.48	1.35
		Sham/85	1.54	1.30
ES active treatment vs sham			**−0.06**	**0.05**
Chevalier 2010/hyaluronic acid			Womac A	Womac C function
	None/6	Active/124	1.52	1.13
		Sham/129	1.18	1.07
ES active treatment vs sham			**0.34**	**0.06**
Altman 2004/hyaluronic acid	None/6		Womac pain	(2)
		Active/172	0.76	0.38
		Sham/174	0.85	0.53
ES active treatment vs sham			**−0.09**	**−0.15**
Kallmes 2009/percutaneous vertebroplasty	<12/1		Roland-Morris Disability Questionnaire	(7)
		Active/68	0.86	0.72
		Sham/63	0.81	0.63
ES active treatment vs sham			**0.05**	**0.09**
Buchbinder 2009/percutaneous vertebroplasty	<12/6		Pain Score	(4)
		Active/38	0.83	0.46
		Sham/40	0.71	0.51

Continued

Table 2 Continued

Author/procedure	Limit disease duration/time to follow-up (months)	Trial arm/number of patients randomised	ES primary end point	ES pooled secondary end points (number of end points)
ES active treatment vs sham			**0.12**	**−0.05**
Cohen 2012/epidural injection of corticosteroids	<6 /1		NRS leg pain	(2)
		Active/28	1.51	0.88
		Sham/30	0.82	0.39
ES active treatment vs sham			**0.69**	**0.49**
Iversen 2011/epidural injection of corticosteroids	>3/12		Oswestry Disability Index	–
		Active/36	1.68	
		Sham/40	1.85	
ES active treatment vs sham			**−0.17**	
Arden 2005/epidural injection of corticosteroids	>1<18/12		Oswestry Disability Index	(2)
		Active /120	1.42	1.14
		Sham/108	1.44	1.21
ES active treatment vs sham			**−0.02**	**−0.07**
Valat 2002/epidural injection of corticosteroids	<6/1		VAS Pain	(3)
		Active/42	1.85	1.10
		Sham/43	1.47	0.99
ES active treatment vs sham			**0.38**	**0.10**
Freeman 2005/intradiscal electrothermal therapy	≥3/6		Oswestry Disability Index	(6)
		Active/38	0.10	−0.03
		Sham/19	−0.07	0.12
ES active treatment vs sham			**0.17**	**−0.15**
Pauza 2003/intradiscal electrothermal therapy	>6/6		Oswestry Disability Index	(3)
		Active/32	0.94	0.90
		Sham/24	0.35	0.46
ES active treatment vs sham			**0.59**	**0.44**
Kvarstein 2009/percutaneous intradiscal radiofrequency thermocoagulation	>6/12		Brief Pain Inventory	(5)
		Active/10	0.34	0.54
		Sham/10	0.23	0.24
ES active treatment vs sham			**0.11**	**0.30**
Olanow 2003/fetal nigral transplantation	None/24		UPDRS 3 off	(5)
		Active/12	0.04	−0.24
		Sham/11	−0.44	−0.19
ES active treatment vs sham			**0.48**	**−0.06**
Marks 2010/gene delivery of AAV2-neurturin	≥60/12		UPDRS 3 off	(7)
		Active/38	0.72	0.23
		Sham/20	0.53	−0.05

Continued

Table 2 Continued

Author/procedure	Limit disease duration/time to follow-up (months)	Trial arm/number of patients randomised	ES primary end point		ES pooled secondary end points (number of end points)
ES active treatment vs sham			**0.19**		**0.28**
Gross 2011/transplantation of human retinal pigmental cells	≥60/12		UPDRS 3 off		(2)
		Active/35	1.09	0.08	
		Sham/36	0.88	0.06	
ES active treatment vs sham			**0.21**	**0.02**	
LeWitt 2011/AAV-GAD gene into subthalamic nucleus	≥60/6		UPDRS 3 off		(7)
		Active/16	1.00	0.30	
		Sham/21	0.42	0.21	
ES active treatment vs sham			**0.58**	**0.08**	
Dowson 2008/patent foramen ovale closure	None/6		Frequency migraine/month (per protocol)		Headache Impact Test
		Active/74	0.74	1.02	
		Sham/73	0.45	1.06	
ES active treatment vs sham			**0.28**	**0.04**	

NRS, Numerical Rating Scale; UPDRS, Unified Parkinson's Disease Rating Scale; VAS, visual analogue scale; Womac, Western Ontario and McMaster Universities Osteoarthritis Index.

groups. Participants' perception of whether they received active treatment or sham has been shown to contribute more to clinical improvement than the biological effects per se.[32 58]

Non-specific factors

The role of non-specific factors, primarily spontaneous remission or statistical regression to the mean in placebo-controlled studies is controversial.[59] A recent meta-analysis analysing 202 trials with an untreated group, spanning 60 different clinical conditions, found rather small differences between placebo and no treatment, with ESs in the range of 0.2–0.3.[60] Apart from acupuncture trials (mean ES 0.68), the authors did not include trials reporting the effectiveness of invasive procedures. Another meta-analysis studied the placebo effect of a range of treatments (pharmacological, non-pharmacological and surgical) for osteoarthritis of the hand, hip and knee.[61] Of 198 included trials, 14 had a no-treatment arm. The mean ES in the placebo groups was about 0.5, while it was only slightly above 0 in the no-treatment groups. The difference between the placebo and no-treatment groups was larger than the difference between the placebo and active treatment groups. Trials using injections, acupuncture and surgery had the largest placebo effects, and the effects were larger for subjective than objective end points. The authors concluded that there is a significant placebo effect on pain, stiffness and function in symptomatic osteoarthritis.

Because the trials in the present study did not include a no-treatment arm (ie, waiting list), we cannot rule out that the changes appearing during the trial period also reflect non-specific factors, that is, spontaneous improvement or regression to the mean. Such mechanisms would be expected to be most prominent in trials with brief illness duration before inclusion and with longer time to follow-up, while improvements in chronic, unremitting conditions such as Parkinson's disease would be more likely attributed to placebo. Interestingly, in three of the four included Parkinson's trials, there were moderate to large improvements in the sham groups even at 1-year follow-up.[49 50 51] Other authors have also found improvements several years after sham surgery, indistinguishable from conventional surgery.[32 62] This is in agreement with recent insights into the neurobiological effects of placebo and their relation to underlying psychological mechanisms, principally expectation and conditioning.[63]

Are ethical objections to sham justified?

The use of sham in controlled surgical trials is a divisive issue, with scepticism, even frank opposition, being voiced by both ethics committees, involved surgeons and anaesthetists, and potential patients.[10] Ethical arguments include the inherent risks of sham procedures combined with the lack of obvious benefits to the participants. Barriers related primarily to feasibility include problems with patient and assessor blinding, differing technical expertise, the

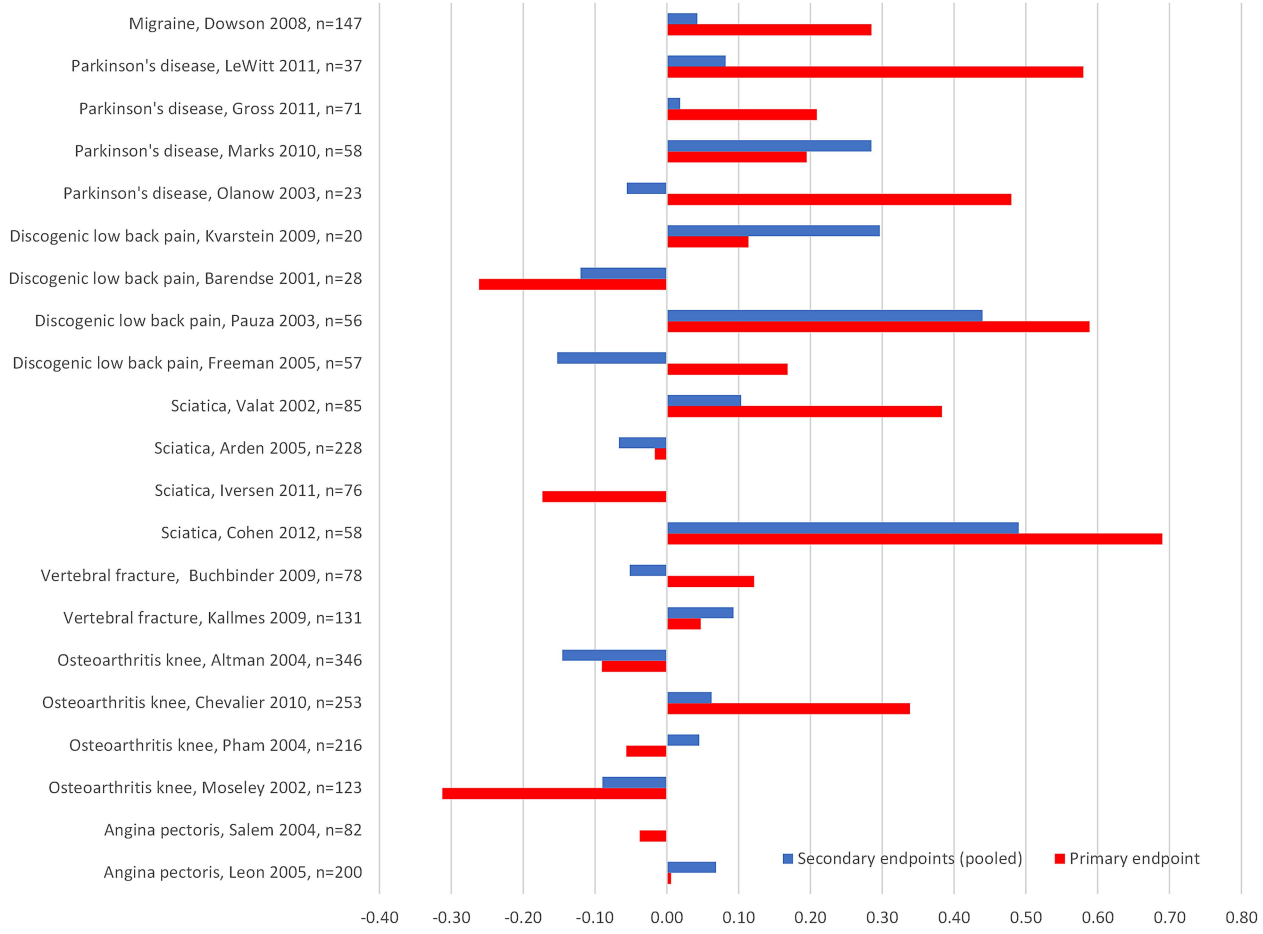

Figure 3 Differences in effect size between active treatment and sham.

heterogeneity of the interventional techniques and variable outcome specifications, making standardisation difficult to achieve. Existing ethical guidelines accept the role of placebo-controlled trials when certain conditions are met.[64] There must be genuine equipoise, that is, conflicting or weak evidence of the effectiveness of a procedure. Blinding of both participants and assessors must be assured and participants must freely consent to suspend knowledge of whether they are receiving sham or conventional treatment. The health risks and consequences of placebo or delayed treatment must be minimal, and outweighed by the societal importance of establishing the clinical utility of the intervention in question.[65 66]

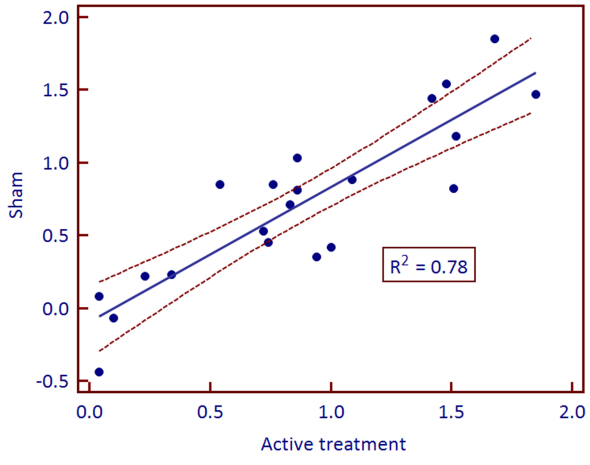

Figure 4 Association between effect sizes of primary end points in active treatment and sham arms. Linear regression, 95% CI; N=21.

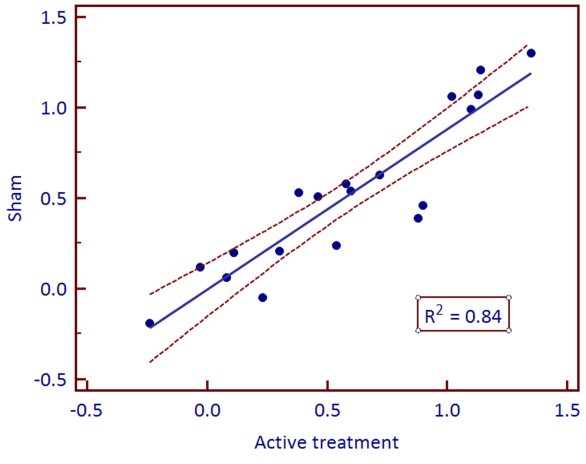

Figure 5 Association between effect sizes of pooled secondary end points in active treatment and sham arms. Linear regression, 95% CI; N=19.

The selected trials gave a detailed description of AEs in active as well as sham-treated groups (table 1). The safety concerns frequently raised as an argument against the use of sham were generally not supported. Major AEs related to the sham procedure were reported in only 1 of the trials[53] and they were short-lived and not life-threatening. Minor AEs were more frequent, but of limited duration. Positive placebo-induced effects generally outweighed AEs, thus weakening ethical arguments against the use of sham interventions. In our opinion, the consequences of the continued use of unproven invasive procedures are of a different magnitude. In the light of studies supporting the beneficial effects of sham procedures, at least for pain and Parkinson's symptoms, research ethics committees should consider such factors in their risk–benefit assessments of planned sham-controlled trials.[67 68]

Clinical implications

The present results are pertinent to the ongoing discussion about wasteful and unproven medical practices, and underscore the necessity for a continual assessment of existing or novel unproven procedures. Minimally invasive techniques have lowered the threshold for interventions and led to their application to a wider clinical spectrum (indication creep) without an ongoing evaluation of effectiveness or safety.[4] The last two decades have seen dramatic increases in the use of several of the described procedures, as well as interventions we have not investigated, such as acromioplasty, percutaneous coronary intervention and recently, robotic surgery.[69–74] In light of the results in the present study, placebo effects might well explain a large part of the purported effects of such procedures. When clinicians and regulators are faced with claims of large treatment effects for insufficiently tested procedures, their default mode should be watchful scepticism. The standards of the evaluation process before approval and reimbursement for devices and procedures need to be strengthened, and economic or regulatory incentives that perpetuate the use of undocumented or harmful procedures should be abrogated.

CONCLUSION

Sham-controlled trials are unique in their ability to discriminate between true treatment effects and non-specific effects. The results of the present study suggest that placebo and other non-specific effects explain a large part of their purported benefits. Further, results indicate that the risks of AEs in sham-controlled trials are over-rated and could be considered acceptable in view of the potential personal harm and societal costs associated with unproven minimally invasive interventions.

Author affiliations
[1]Fram Rehabilitation Centre, Bærum, Norway
[2]Department of Physical Medicine and Rehabilitation, Oslo University Hospital, Oslo, Norway
[3]South-Eastern Norway Regional Health Authority, Hamar, Norway

Contributors RH initiated and planned the project and searched databases; also made article screening, data extraction and statistical analysis, who also wrote the draft. He is the guarantor. JIB and OT assisted in developing search strategies and also quality of data extraction and checking and they reviewed the draft and contributed to manuscript revisions.

Funding This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests None.

Provenance and peer review Not commissioned; internally peer reviewed.

Data sharing statement Extra data can be accessed via the Dryad data repository at http://datadryad.org/ with the doi:10.5061/dryad.105cb.

REFERENCES

1. Prasad V, Vandross A, Toomey C, *et al*. A decade of reversal: an analysis of 146 contradicted medical practices. *Mayo Clin Proc* 2013;88:790–8.
2. Tiago V, Pereira TV, Horwitz RI, *et al*. Empirical evaluation of very large treatment effects of medical interventions. *JAMA* 2012;308:1676–84.
3. Roberts AL, Kewman DG, Mercier L, *et al*. The power of nonspecific effects in healing: implications for psychosocial and biological treatments. *Clin Psychol Rev* 1993;13:375–91.
4. Garner S, Littlejohns P. Disinvestment from low value clinical interventions: NICEly done? *BMJ* 2011;343:d4519.
5. Scott IA, Elshaug AG. Foregoing low-value care: how much evidence is needed to change beliefs? *Intern Med J* 2013;43:107–9.
6. Obremskey WT, Pappas N, Attallah-Wasif E, *et al*. Level of evidence in orthopaedic journals. *J Bone Joint Surg Am* 2005;87:2632–8.
7. Reeves B. Health-technology assessment in surgery. *Lancet* 1999;353(Suppl 1):S13–15.
8. Wenner DM, Brody BA, Jarman AF, *et al*. Do surgical trials meet the scientific standards for clinical trials? *J Am Coll Surg* 2012;215:722–30.
9. Ezekiel JE, Miller FG. The ethics of placebo-controlled trials—a middle ground. *N Engl J Med* 2001;345:915–19.
10. Campbell MK, Entwistle VA, Cuthbertson BH, *et al*. Developing a placebo-controlled trial in surgery: issues of design, acceptability and feasibility. *Trials* 2011;12:50.
11. Cochrane Handbook for Systematic Reviews of Interventions. Version 5.1.0. http://handbook.cochrane.org (11.11.2014).
12. Liberati A, Altman DG, Tetzlaff J, *et al*. The PRISMA statement for reporting systematic reviews and meta-analyses of studies that evaluate healthcare interventions: explanation and elaboration. *BMJ* 2009;339:b2700.
13. Cohen J. *Statistical power analysis for the behavioral sciences*. Hillsdale, NJ: Lawrence Erlbaum Associates, 1988.
14. Durlak JA. How to select, calculate, and interpret effect sizes. *J Pediatr Psychol* 2009;34:917–28.
15. MedCalc Software bvba, Ostend, Belgium. http://www.medcalc.org (08.03.2014).
16. Nath S, Nath CA, Pettersson K, *et al*. 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.
17. 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.
18. Cohen SP, Hurley RW, Buckenmaier CC, *et al*. Randomized placebo-controlled study evaluating lateral branch radiofrequency denervation for sacroiliac joint pain. *Anesthesiology* 2008;109:279–88.
19. Tekin I, Mirzai H, Ok G, *et al*. A comparison of conventional and pulsed radiofrequency denervation in the treatment of chronic facet joint pain. *Clin J Pain* 2007;23:524–9.
20. 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.

21. van Wijk RM, Geurts JW, Wynne HJ, *et al*. Radiofrequency denervation of lumbar facet joints in the treatment of chronic low back pain: a randomized, double-blind, sham lesion-controlled trial. *Clin J Pain* 2005;21:335–44.
22. Rutjes AWS, Jünl P, da Costa BR, *et al*. Viscosupplementation for osteoarthritis of the knee. A systematic review and meta-analysis. *Ann Intern Med* 2012;157:180–91.
23. Pinto RZ, Maher CG, Ferreira ML, *et al*. Epidural corticosteroid injections in the management of sciatica: a systematic review and meta-analysis. *Ann Intern Med* 2012;157:865–77.
24. McGillion M, Cook A, Victor JC, *et al*. Effectiveness of percutaneous laser revascularization therapy for refractory angina. *Vasc Health Risk Manag* 2010;6:735–47.
25. Helm S II, Deer TR, Manchikanti L, *et al*. Effectiveness of thermal annular procedures in treating discogenic low back pain. *Pain Physician* 2012;15:E279–304.
26. Shi MM, Cai XZ, Lin T, *et al*. Is there really no benefit of vertebroplasty for osteoporotic vertebral fractures? A meta-analysis. *Clin Orthop Relat Res* 2012;470:2785–99.
27. Petrella RJ, Cogliano A, Decaria J. Combining two hyaluronic acids in osteoarthritis of the knee: a randomized, double-blind, placebo-controlled trial. *Clin Rheumatol* 2008;27:975–81.
28. Lundsgaard C, Dufour N, Fallentin E, *et al*. Intra-articular sodium hyaluronate 2 mL versus physiological saline 20 mL versus physiological saline 2 mL for painful knee osteoarthritis: a randomized clinical trial. *Scand J Rheumatol* 2008;37:142–50.
29. Karppinen J, Malmivaara A, Kurunlahti M, *et al*. Periradicular infiltration for sciatica: a randomized controlled trial. *Spine* 2001;26:1059–67.
30. 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.
31. 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.
32. 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.
33. 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.
34. 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.
35. Pham T, Le Henanff A, Ravaud P, *et al*. Evaluation of the symptomatic and structural efficacy of a new hyaluronic acid compound, NRD101, in comparison with diacerein and placebo in a 1 year randomised controlled study in symptomatic knee osteoarthritis. *Ann Rheum Dis* 2004;63:1611–17.
36. Altman RD, Akermark C, Beaulieu AD, *et al*. Efficacy and safety of a single intra-articular injection of non-animal stabilized hyaluronic acid (NASHA) in patients with osteoarthritis of the knee. *Osteoarthritis Cartilage* 2004;12:642–9.
37. Chevalier X, Jerosch J, Goupille P, *et al*. Single, intra-articular treatment with 6 ml hylan G-F 20 in patients with symptomatic primary osteoarthritis of the knee: a randomised, multicentre, double-blind, placebo controlled trial. *Ann Rheum Dis* 2010;69:113–19.
38. 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.
39. 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.
40. Iversen T, Solberg TK, Romner B, *et al*. Effect of caudal epidural steroid or saline injection in chronic lumbar radiculopathy: multicentre, blinded, randomised controlled trial. *BMJ* 2011;343:d5278.
41. Cohen SP, White RL, Kurihara C, *et al*. Epidural steroids, etanercept, or saline in subacute sciatica: a multicenter, randomized trial. *Ann Intern Med* 2012;156:551–9.
42. Arden NK, Price C, Reading I, *et al*. A multicentre randomized controlled trial of epidural corticosteroid injections for sciatica: the WEST study. *Rheumatology* 2005;44:1399–406.
43. Valat JP, Giraudeau B, Rozenberg S, *et al*. Epidural corticosteroid injections for sciatica: a randomised, double blind, controlled clinical trial. *Ann Rheum Dis* 2003;62:639–43.
44. Pauza KJ, Howell S, Dreyfuss P, *et al*. A randomized, placebo-controlled trial of intradiscal electrothermal therapy for the treatment of discogenic low back pain. Spine *J* 2004;4:27–35.
45. 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* 2005;30:2369–77.
46. Kvarstein G, Måwe L, Indahl A, *et al*. A randomized double-blind controlled trial of intra-annular radiofrequency thermal disc therapy—a 12-month follow-up. *Pain* 2009;145:279–86.
47. Moseley JB, O'Malley K, Petersen NJ, *et al*. A controlled trial of arthroscopic surgery for osteoarthritis of the knee. *N Engl J Med* 2002;347:81–8.
48. 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.
49. 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.
50. Marks WJ Jr, Bartus RT, Siffert J, *et al*. Gene delivery of AAV2-neurturin for Parkinson's disease: a double-blind, randomised, controlled trial. *Lancet Neurol* 2010;9:1164–72.
51. Gross RE, Watts RL, Hauser RA, *et al*. Intrastriatal transplantation of microcarrier-bound human retinal pigment epithelial cells versus sham surgery in patients with advanced Parkinson's disease: a double-blind, randomised, controlled trial. *Lancet Neurol* 2011;10:509–19.
52. LeWitt PA, Rezai AR, Leehey MA, *et al*. AAV2-GAD gene therapy for advanced Parkinson's disease: a double-blind, sham-surgery controlled, randomised trial. *Lancet Neurol* 2011;10:309–19.
53. 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.
54. Hannink G, Gooszen HG, Rovers MM. Comparison of registered and published primary outcomes in randomized clinical trials of surgical interventions. *Ann Surg* 2013;257:818–23.
55. Birch S. A review and analysis of placebo treatments, placebo effects, and placebo controls in trials of medical procedures when sham is not inert. *J Altern Complement Med* 2006;12:303–10.
56. Bicket MC, Gupta A, Brown CH IV, *et al*. Epidural injections for spinal pain: a systematic review and meta-analysis evaluating the "control" injections in randomized controlled trials. *Anesthesiology* 2013;119:907–31.
57. Kaptchuk TJ, Goldman P, Stone DA, *et al*. Do medical devices have enhanced placebo effects? *J Clin Epidemiol* 2000;53:786–92.
58. Einvik G, Tjomsland O, Kvernebo K, *et al*. Preoperative expectations and clinical outcome of transmyocardial laser treatment in patients with angina pectoris. *Tidsskr Nor Laegeforen* 2002;122:2102–4.
59. Turner JA, Deyo RA, Loeser JD, *et al*. The importance of placebo effects in pain treatment and research. *JAMA* 1994;271:1609–14.
60. Hróbjartsson A, Gøtzsche PC. Placebo interventions for all clinical conditions. *Cochrane Database Syst Rev* 2010;(1):CD003974.
61. Zhang W, Robertson J, Jones AC, *et al*. The placebo effect and its determinants in osteoarthritis: meta-analysis of randomised controlled trials. *Ann Rheum Dis* 2008;67:1716–23.
62. Marchand S, Kupers RC, Bushnell MC, *et al*. Analgesic and placebo effects of thalamic stimulation. *Pain* 2003;105:481–8.
63. Finniss DG, Kaptchuk TJ, Miller F, *et al*. Biological, clinical, and ethical advances of placebo effects. *Lancet* 2010;375:686–95.
64. Snyder L. Review of the American College of Physicians Ethics Manual, Sixth Edition. *Ann Intern Med* 2012;156(1_Part_2):73–104.
65. Flum DR. Interpreting surgical trials with subjective outcomes: avoiding UnSPORTsmanlike conduct. *JAMA* 2006;296:2483–5.
66. Heckerling PS. Placebo surgery research: a blinding imperative. *J Clin Epidemiol* 2006;59:876–80.
67. Brim RL, Miller FG. The potential benefit of the placebo effect in sham-controlled trials: implications for risk-benefit assessments and informed consent. *J Med Ethics* 2013;39:703–7.
68. Redberg RF. Sham controls in medical device trials. *N Engl J Med* 2014;371:892–3.
69. Haahr JP, Østergaard S, Dalsgaard J, *et al*. Exercises versus arthroscopic decompression in patients with subacromial impingement: a randomised, controlled study in 90 cases with a one year follow up. *Ann Rheum Dis* 2005;64:760–4.
70. Ketola S, Lehtinen J, Rousi T, *et al*. No evidence of long-term benefits of arthroscopic acromioplasty in the treatment of shoulder impingement syndrome: five-year results of a randomised controlled trial. *Bone Joint Res* 2013;2:132–9.

71. Herrlin S, Hållander M, Wange P, *et al*. Arthroscopic or conservative treatment of degenerative medial meniscal tears: a prospective randomised trial. *Knee Surg Sports Traumatol Arthrosc* 2007;15:393–401.
72. Yu E, Cil A, Harmsen WS, *et al*. Arthroscopy and the dramatic increase in frequency of anterior acromioplasty from 1980 to 2005: an epidemiologic study. *Arthroscopy* 2010;26(Suppl):S142–7.
73. Vitale MA, Arons RR, Hurwitz S, *et al*. The rising incidence of acromioplasty. *J Bone Joint Surg Am* 2010;92:1842–50.
74. Brox JI, Staff PH, Ljunggren AE, *et al*. Arthroscopic surgery compared with supervised exercises in patients with rotator cuff disease (stage II impingement syndrome). *BMJ* 1993;307:899–903.