BMJ Open

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

Journal:	*BMJ Open*
Manuscript ID:	bmjopen-2014-007331
Article Type:	Research
Date Submitted by the Author:	29-Nov-2014
Complete List of Authors:	HOLTEDAHL, ROBIN; Fram Rehabilition Centre, Tjomsland, Ole; South-Eastern Norway Regional Health Authority, Division of Quality and Specialist Areas Brox, Jens; Oslo University Hospital, Orthopaedic
Primary Subject Heading:	Evidence based practice
Secondary Subject Heading:	Evidence based practice, Ethics, Research methods, Surgery
Keywords:	Quality in health care < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Health economics < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Clinical trials < THERAPEUTICS, MEDICAL ETHICS, PAIN MANAGEMENT

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

Robin Holtedahl, Jens Ivar Brox, Ole Tjomsland

Fram Rehabilitation Centre, Rykkinveien 100, 1349 Rykkin, Norway Robin Holtedahl Consultant Department of Physical Medicine and Rehabilitation, Oslo University Hospital, Box 4956 Nydalen, 0424 Oslo, Norway Jens Ivar Brox Professor South-Eastern Norway Regional Health Authority, PB 404, 2303 Hamar, Norway Ole Tjomsland Director of Quality and Specialist Areas

Correspondence to: Dr Robin Holtedahl; robi-hol@online, telephone +4790248973

Key words:
Placebo effects
Invasive procedures
Biomedical ethics
Evidence based health care

Word count: 3783

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 endpoints were calculated. Linear regression was used to analyse the association between endpoints 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 endpoints, there was a large clinical effect (ES ≥0.8) after active treatment in 12 trials and after sham procedures in 11 trials. For secondary endpoints, seven and five trials showed a large clinical effect, respectively. Three trials showed a moderate difference in ES between active treatment and sham on primary endpoints (ES ≥0.5) but no trials reported a large difference. No trials showed large or moderate differences in ES on pooled secondary endpoints. Regression analysis of endpoints in active treatment and sham arms estimated an R^2 of 0.78 for primary and 0.84 for secondary endpoints. Adverse events after sham were in most cases minor and of short duration.

Conclusion The generally small differences in effect size 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.

3

SUMMARY

Key messages

- The magnitude of change in the active treatment- and placebo arms varied greatly, but about 80% of the variance in effect size of active treatment could be predicted by placebo effects, regression to the mean or spontaneous improvement.
- Adverse events related to sham procedures were mainly minor and short-lived, and frequently outweighed by positive placebo effects.

Strengths and limitations

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

1

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 percent 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. 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: (a) to assess the magnitude of change in outcome from baseline to trial endpoint in both the active treatment and placebo (sham) arms, (b) to explore the contribution of non-specific factors, including placebo, to the outcome of active treatment, and (c) to assess the level of reported adverse effects in both trial arms.

METHODS

Search strategy and selection criteria

We first conducted electronic searches for randomised placebo-controlled trials of minimally invasive interventions for cardiac, neurological and selected musculoskeletal conditions, using MEDLINE and Cochrane library to identify systematic reviews published between January 2009 and January 2014. 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 excluded open surgical and laparoscopic interventions. Where applicable, we used the “core clinical journals” filter in PubMed, which is an index of journals particularly relevant to practicing physicians. From the reviews, we selected randomised placebo-controlled trials published from January 2000 to January 2014 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 both 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 mentioned above 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, standard deviation, standard error or confidence intervals in both active and sham-groups, as well as trials with only graphic representation of data. Details of the search strategy are shown in web appendix table 1 and web appendix figure 1. We give a short description of each procedure’s introduction, therapeutic rationale and history in web appendix table 2. This review is reported in accordance with the PRISMA statement.[12]

Data extraction

We registered all continuous primary endpoints. In trials without continuous primary endpoints, with multiple endpoints or no defined primary endpoint, we selected an outcome related to pain or condition-specific endpoint. 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 endpoint. We also registered secondary endpoints in order to avoid potential bias from selective reporting in the included trials. The included and excluded secondary endpoints are shown in web appendix table 3. Endpoints describing medication, radiographic or physiological variables, social or psychological function, were not included. For the Parkinson-trials, only endpoints 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 in both trial arms were registered (tables 1 and 2).

Data synthesis

To assess clinically important change, we calculated effect size (ES, Cohen's d), based on the means and standard deviations (SD). We calculated ES both for the active and sham intervention 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 standard deviations before and after treatment calculated ES. An ES of 0.8 or more is assumed large, while an ES of 0.5 - 0.8 is considered moderate.[13] In trials with multiple secondary endpoints we calculated the pooled mean ES, without weighting. Because of 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 outcome in the active and sham groups both for primary and pooled secondary endpoints. For this analysis, we used Medcalc Statistical Software version 12.7.4.0[15]

RESULTS

Selection of interventions

The searches provided sham-controlled trials of the following interventions: percutaneous laser revascularisation of myocardium for angina pectoris, closure of foramen ovale for migraine, arthroscopic meniscectomy for meniscal tears, debridement and injection of hyaluronic acid for symptomatic osteoarthritis of the knee and injection or transplantation of biologically active material for Parkinson's disease (human retinal pigmental cells, fetal nigral cells and Neurturin). Because of the large number of described interventions for neck- and back pain syndromes, we chose to restrict the analysis to sham-controlled trials of the following interventions: epidural injections of corticosteroids for sciatica (caudal, interlaminar and transforaminal routes), percutaneous heating of the intervertebral disc for chronic low back pain (percutaneous intradiscal radiofrequency thermocoagulation and intradiscal electrothermal therapy) and vertebroplasty for vertebral body fractures. The searches provided no sham-controlled trials of arthroscopic procedures other than knee conditions.

Study selection

The study selection process is summarised in web appendix figure 1. Web appendix table 1 shows the excluded trials and the reasons for exclusion. The search provided five systematic reviews, all identified through searches in MEDLINE, none were commercially funded.[16-20] It identified a total of 71 clinical trials, twelve of them 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.[21-26] 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.

Table 1. **Included 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 2005	Food and Drug Administration / NC	Percutaneous myocardial laser revascularization / intractable angina pectoris	Laser turned on but no procedure performed	MAE in hospital (high dose): 4.1%	MAE in hospital: 0
Salem 2004	Ethics committee / NC			No procedural AE	
Sihvonen 2013	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 2002	Review Board / NC	Arthroscopic debridement / Knee osteoarthritis	Simulated arthroscopy preparation, intravenous anaesthesia, skin incisions, no instruments entered knee, knee manipulated	No procedural AE	
Pham 2004	Review Board /	Hyaluronic acid /	Intraarticular	No MAE	

	NC	Knee osteoarthritis	injection of saline solution	Any mAE: 81.7%	Any mAE: 1.2%
Altman 2004	Ethics committee / C			No MAE	
				mAE: 12.8%	mAE: 8%
Chevalier 2010	ClinicalTrials.org / C			No MAE	
				mAE: 35,8%	mAE: 33,8%
Kallmes 2009	Review Board / NC	Percutaneous vertebroplasty with PMMA cement injection / vertebral compression fracture	Conscious sedation + local anesthaesia, pressure put on spine, simulation of odor with mixing of PMMA to imitate the smell during the active procedure	No MAE	
				mAE: 14%	mAE: 16%
Buchbinder 2009	Ethics committee at each participating center / NC		Conscious sedation + local anesthaesia, needle inserted to rest on the lamina, PMMA container opened to imitate the smell during the active procedure	No procedural AE	
Cohen 2012	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 2005	Ethics committee / NC		2 mL saline into interspinous ligament	No MAE	
				mAE: 9%	mAE: 10%
Valat 2002	Ethics committee / NC		2 mL saline into epidural space, interlaminar approach	No MAE	
				mAE: 6%	mAE: 8%
Iversen 2011	Ethics committee / NC		Subcutaneous injection of 2 mL saline superficial to the sacral hiatus	Not reported	
Freeman 2005	Ethics committee / C	Intradiscal electrothermal therapy (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 2003	Review Board / NC		17-gauge needle introduced onto the outer annulus, mock electrode passage shown on monitor, generator noises produced	Not reported	
Kvarstein 2009	Ethics committee / NC	Percutaneous intradiscal radiofrequency thermocoagulation (PIRFT) / discogenic low back pain	17-gauge canula and RF-probe inserted into annulus, no RF current applied	Not reported	
Olanow 2003	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 2010	Review Board / C	Gene delivery of AAV2-Neurturin / Parkinson's disease	Scalp incisions, partial thickness burr holes, no intracranial injections	MAE: 4	MAE: 0
				Most frequent mAE: headache: 68%	Most frequent mAE: headache: 50%
Gross 2011	Review Board / C	Transplantation of human retinal pigmental cells / Parkinson's disease	Scalp incisions, partial thickness burr holes, no cell transplantation	1 death	0 deaths
				MAE: 23%	MAE: 0
LeWitt 2011	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 2008	Ethics committee / C	Patent foramen ovale closure with STARFlex Septal Repair Implant / migraine	General anesthesia, skin incision in the groin	MAE (possibly or probably related to procedure): 11%	MAE (possibly or probably related to procedure): 4%
C=commercial; NC=non-commerical; MAE=major adverse events; mAE=minor adverse events; PMMA=polymethylmethacrylate; AAV2 =adeno-associated; GAD=glutamic acid decarboxylase					

Fourteen trials from the systematic reviews fulfilled at least four of the five methodological criteria.[27 28 31-42] Seven trials provided through searches in MEDLINE fulfilled the same criteria.[29 30 43-47] All trials reported approval of study protocol prior to patient enrolment (table 1). Eight trials were commercially funded.[32 33 41 44-47] 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 endpoints after active treatment, while 11 trials showed a similar ES after the sham procedure (figure 1, table 2).

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

Author / procedure	**Limit disease duration / time to follow-up (months)**	**Trial arm / no of patients randomised**	**ES primary endpoint**	**ES pooled secondary endpoints (no of endpoints)**
Leon 2005 / Percutaneous myocardial laser revascularization	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 revascularization	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
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
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**

VAS=Visual Analogue Scale; NRS=Numerical Rating Scale; UPDRS=Unified Parkinson's Disease Rating Scale; Womac=Western Ontario and McMaster Universities Osteoarthritis Index

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

On pooled secondary endpoints, 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 endpoint during treatment, while this was the case for two of the sham groups (not reported to be related to the procedure). On secondary endpoints, deterioration occurred in two active treatment and two sham groups (table 2).

Differences in outcome between active treatment and sham
Better results on primary endpoints were reported with active treatment compared to sham in 14 of the 21 trials, but the differences were small. Three trials (one epidural study[37], one discogenic pain study[40] and one Parkinson study[46]) reported a moderate effect but none showed a large effect (figure 2, table 2). Seven trials reported a better primary endpoint outcome after sham than after active treatment.

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

On regression analyses, effect sizes in the sham groups predicted about 80 % of the variance of ES in the active treatment groups, both on primary and pooled secondary endpoints (figure 3 and 4).

Adverse events

Eighteen studies provided information about adverse events (AE) (table 1). Three of these trials reported no procedural adverse events in any of the groups.[27 29 35] Major AEs were reported after active treatment in four trials[28 44 45 47] including one death in one of the Parkinson studies.[45] In the sham groups, one trial[47] listed three major AEs possibly or probably related to the procedure, all thought to be caused by anti-platelet 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.

DISCUSSION

Principal findings

Analysis of 21 sham-controlled trials of minimally invasive procedures showed that the effect sizes in the active arms were predicted by the effect sizes in the sham arms. There was a large ES on primary endpoints 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 endpoints.

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 endpoints varied from around zero 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 three, and in the epidural trials by a factor of two. 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 endpoints 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 study

It is our opinion that the calculation of effect sizes in both active treatment and placebo arms is a strength of the present study. This made it possible to assess the magnitude of change in both arms and the contribution of non-specific factors to change in the active treatment arms. The calculation of effect sizes provides an alternative assessment to probability estimates. Another strength of the study is the supplementary analyses of pooled secondary endpoints, enabling a more comprehensive evaluation than using primary endpoints alone. Reports of tactically motivated use of primary and secondary endpoints before publication in order to improve study results strengthen the argument for registering all relevant secondary endpoints.[48] Our finding that a majority of trials reported better results on primary than on secondary endpoints 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 the Cochrane Library and MEDLINE, but 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 realize 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.[49] More recently, Bickett et al. hypothesised that epidural injection of small volumes of saline might have physiological effects.[50] 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.[51] In spite of their heterogeneous nature, the 21 selected trials share a medico-technological 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 endpoints in the active treatment and sham 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.[26 52]

Non-specific factors

The role of non-specific factors, primarily spontaneous remission or statistical regression-to-the-mean, in placebo-controlled studies is controversial.[53] 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 effect sizes in the range of 0.2 to 0.3.[54] 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.[55] Of 198 included trials fourteen had a no-treatment arm. The mean ES in the placebo groups was about 0.5, while it was only slightly above zero 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 endpoints. 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 (i.e. waiting list), we cannot rule out that the changes appearing during the trial period also reflect non-specific factors, i.e. 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 trials, there were moderate to large improvements in the sham groups even at one-year follow-up.[43-45] Other authors have also found improvements several years after sham surgery, indistinguishable from conventional surgery.[26 56] This is in agreement with recent insights into the neurobiological effects of placebo and their relation to underlying psychological mechanisms, principally expectation and conditioning.[57]

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.[58] 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 heterogeneity of the interventional techniques and variable outcome specifications, making standardization difficult to achieve. Existing ethical guidelines accept the role of placebo-controlled trials when certain conditions are met.[59] There must be genuine equipoise, i.e. 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.[60 61]

The selected trials gave a detailed description of adverse events in both active and sham-treated groups (table 1). The safety concerns frequently raised as an argument against the use of sham were generally not supported. Major adverse events related to the sham procedure were reported in only one of the trials[47] and they were short-lived and not life threatening. Minor adverse events were more frequent, but also of limited duration. Positive placebo-induced effects generally outweighed adverse events, 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 symptoms, research ethics committees should consider such factors in their risk-benefit assessments of planned sham controlled trials.[62 63]

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 facet joint injections, radiofrequency neurotomy, acromioplasty, percutaneous coronary intervention and, more recently, robotic surgery.[64-69] 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 of 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 adverse events in sham-controlled trials are overrated and could be considered acceptable in view of the potential

personal harm and societal costs associated with unproven minimally invasive interventions.

Figure legends

Figure 1. Effect sizes of active treatment and sham, primary endpoints.

Figure 2. Differences in effect size between active treatment and sham.

Figure 3. Association between effect sizes of primary endpoints in active treatment and sham arms. Linear regression, 95% confidence intervals. N=21.

Figure 4. Association between effect sizes of pooled secondary endpoints in active treatment and sham arms. Linear regression, 95% confidence intervals. N=19.

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-S15;

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**:915-9.

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. 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.

17. 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.

18. 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.

19. 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-E304.

20. 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.

21. 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.

22. 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.

23. Karppinen J, Malmivaara A, Kurunlahti M et al. Periradicular infiltration for sciatica: a randomized controlled trial. *Spine* 2001; **26**:1059-67.

24. Freed CR, Greene PE, Breeze RE et al.Transplantation of embryonic dopamine neurons for severe Parkinson's disease. *N Engl J Med* 2001;**344**:710-9.

25. 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.

26. 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.

27. 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.

28. 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**:1812-9.

29. 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.

30. 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.

31. 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**:1611-7.

32. 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.

33. 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**:113-9.

34. 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.
35. 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.
36. 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.
37. 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.
38. 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.
39. 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.

40. 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.

41. 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.

42. 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.

43. 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.

44. 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.

45. 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.

46. 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.

47. 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.

48. 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.

49. 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.

50. Bicket MC, Gupta A, Brown CH 4th 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.

51. Kaptchuk TJ, Goldman P, Stone DA et al. Do medical devices have enhanced placebo effects? *J Clin Epidemiol* 2000;**53**:786-92.

52. 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.

53. Turner JA, Deyo RA, Loeser JD et al. The importance of placebo effects in pain treatment and research. *JAMA* 1994;**271**:1609-14.

54. Hróbjartsson A, Gøtzsche PC. Placebo interventions for all clinical conditions. *Cochrane Database Syst Rev* 2010; CD003974.

55. 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.

56. Marchand S, Kupers RC, Bushnell MC et al. Analgesic and placebo effects of thalamic stimulation. *Pain* 2003;**105**:481-8.

57. Finniss DG, Kaptchuk TJ, Miller F et al. Biological, clinical, and ethical advances of placebo effects. *Lancet* 2010;**375**:686-95.

58. 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. doi: 10.1186/1745-6215-12-50.

59. Snyder L. Review of the American College of Physicians Ethics Manual, Sixth Edition. *Ann Intern Med* 2012;**156**(1_Part_2):73-104.

60. Flum DR. Interpreting surgical trials with subjective outcomes: avoiding UnSPORTsmanlike conduct. *JAMA* 2006;**296**:2483-5.

61. Heckerling PS. Placebo surgery research: a blinding imperative. *J Clin Epidemiol* 2006;59:876-80.

62. 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.

63. Redberg RF. Sham controls in medical device trials. *N Engl J Med* 2014;**371**:892-3.

64. 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.

65. Ketola S, Lehtinen J, Rousi T et al. No evidence of long-term benefits of arthroscopicacromioplasty in the treatment of shoulder impingement syndrome: Five-year results of a randomised controlled trial. *Bone Joint Res* 2013;**2**:132-9.

66. 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.

67. 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**,Supplement:S142-7.

68. Vitale MA, Arons RR, Hurwitz S et al.The rising incidence of acromioplasty. *J Bone Joint Surg Am* 2010;**92**:1842-50.

69. 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.

Contributors: RH initiated and planned the project and searched databases. JIB and OT assisted in developing search strategies. Article screening and data extraction was carried out by RH. Quality of data extraction and checking was carried out by JIB and OT. Statistical analysis was undertaken by RH, who also wrote the draft. OT and JIB reviewed the draft and contributed to manuscript revisions. RH is the guarantor for this study.

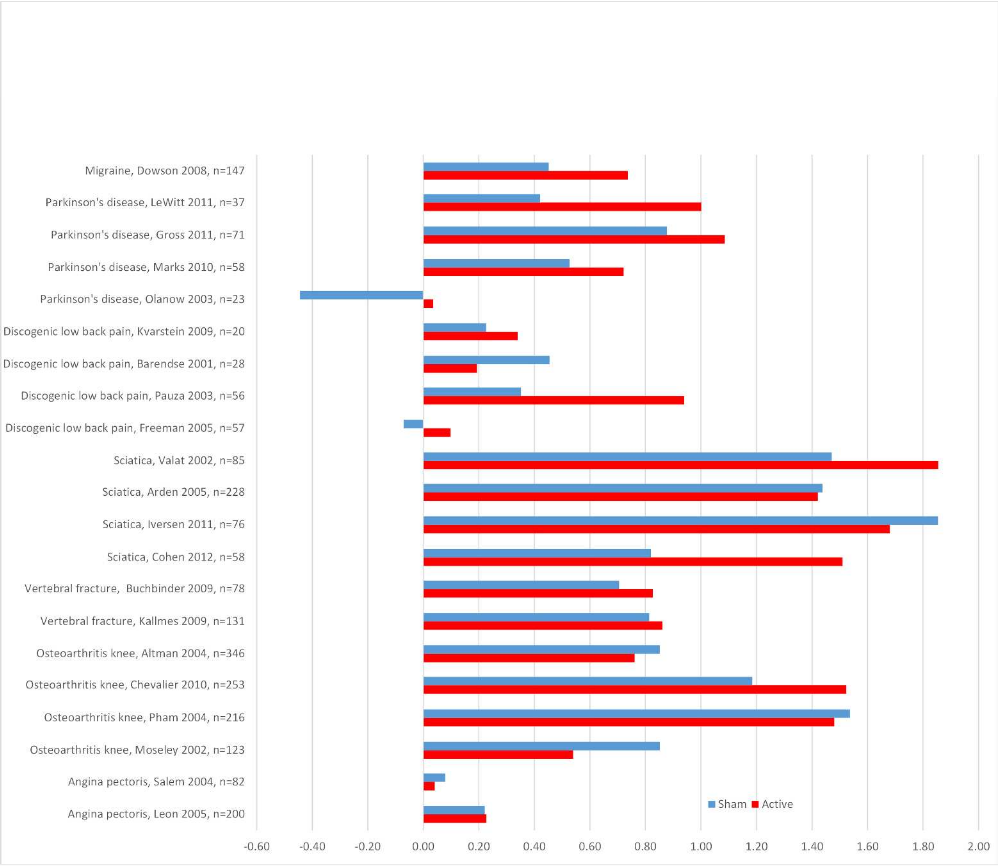
Funding sources: None.

All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

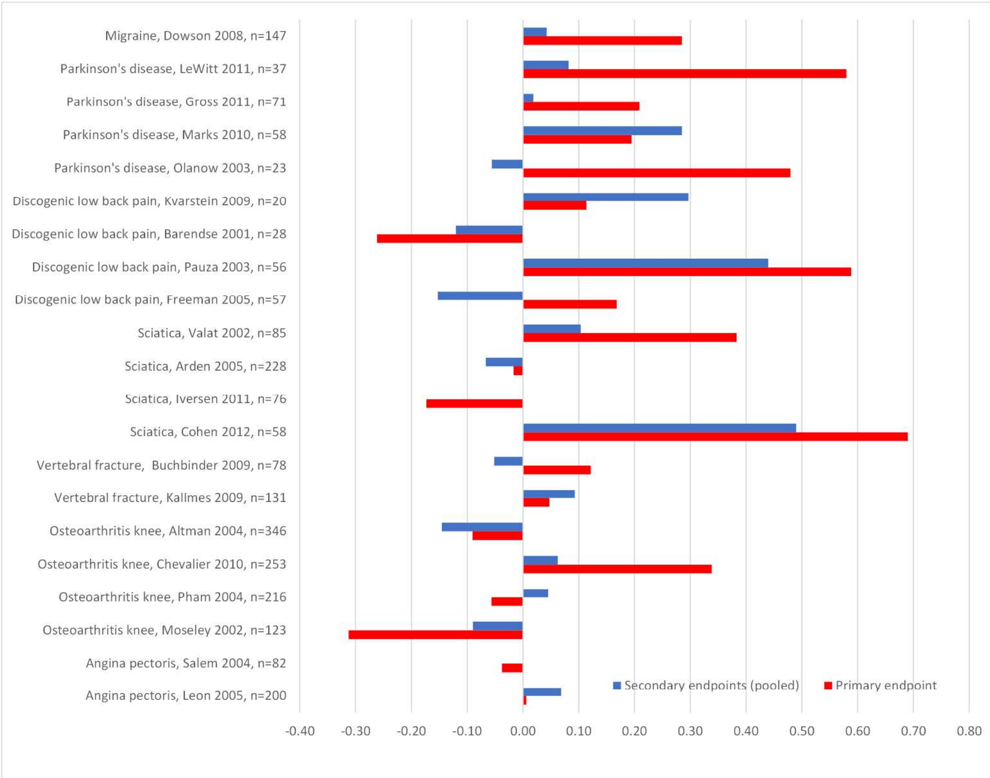
Ethical approval: Ethical approval was not required for this work.

Data sharing: Dataset can be obtained from Robin Holtedahl (robi-hol@online.no).

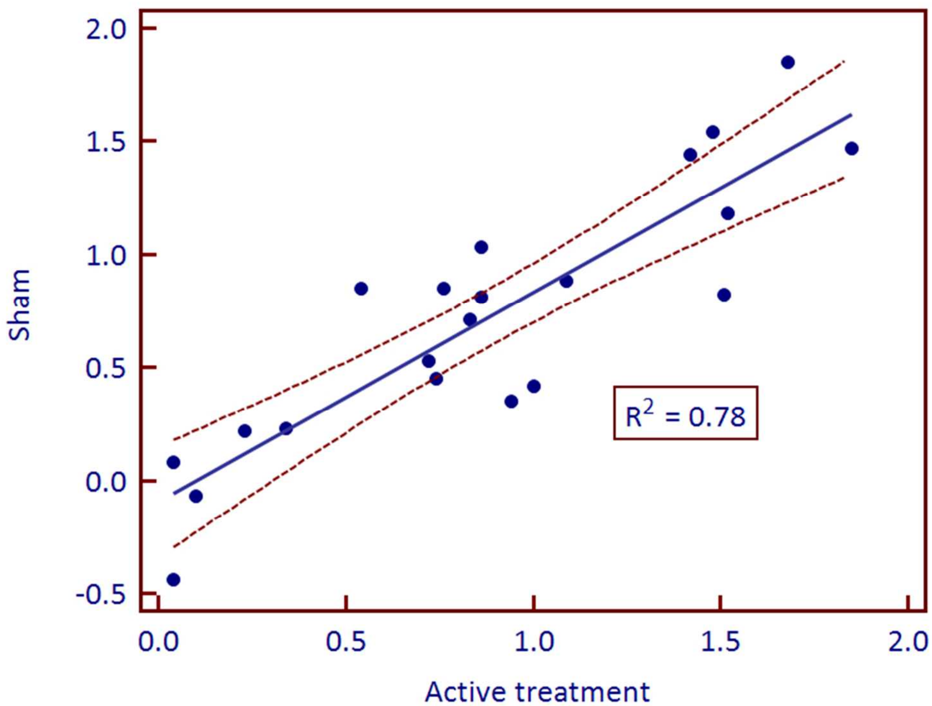
The lead author affirms that this manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned (and, if relevant, registered) have been explained.



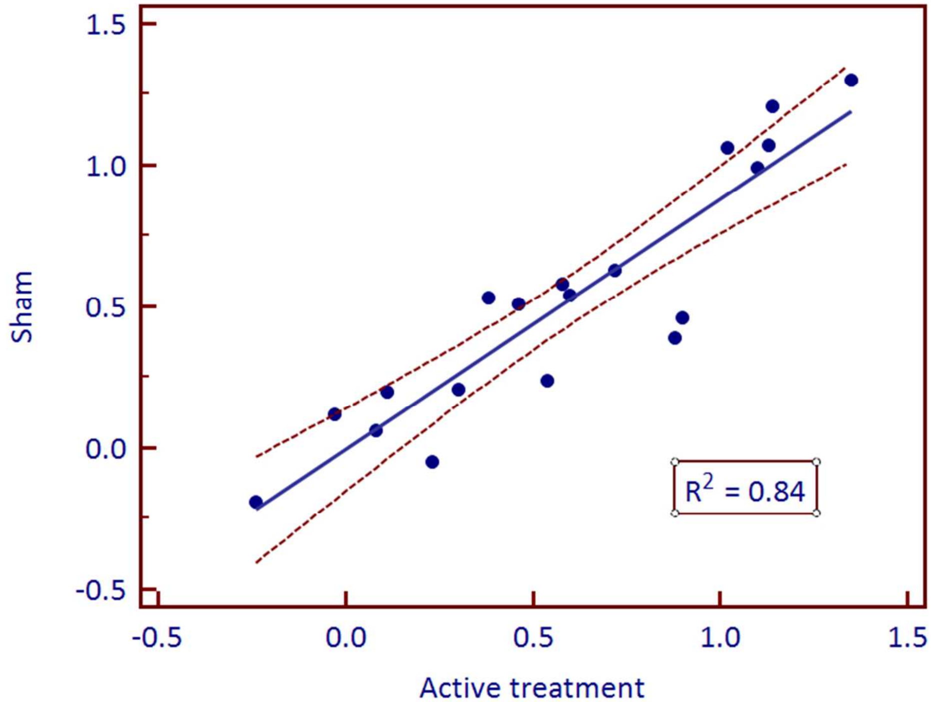
Effect sizes of active treatment and sham, primary endpoints.
250x216mm (300 x 300 DPI)



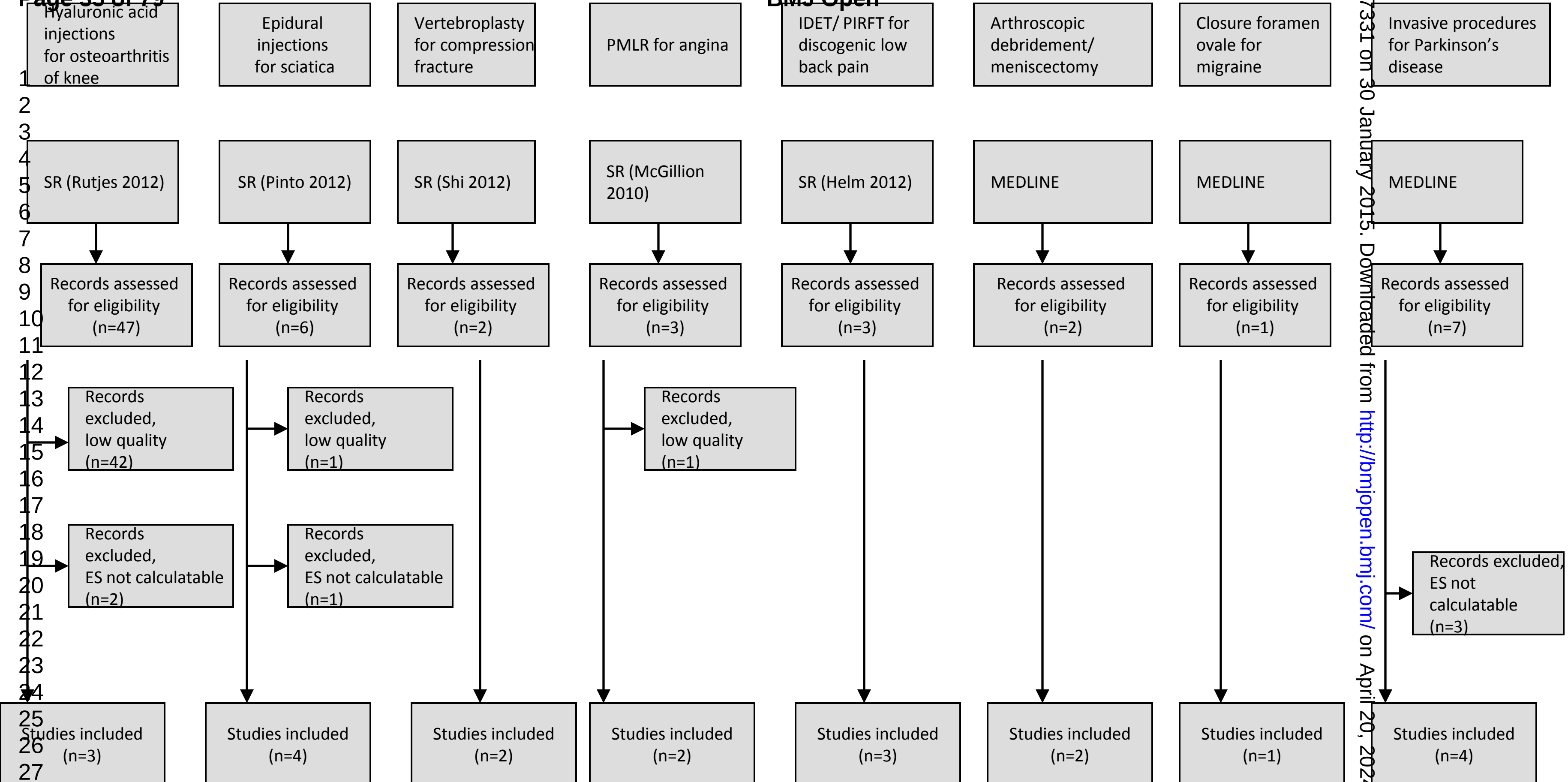
Differences in effect size between active treatment and sham.
221x173mm (300 x 300 DPI)



Association between effect sizes of primary endpoints in active treatment and sham arms. Linear regression, 95% confidence intervals. N=21.
67x50mm (300 x 300 DPI)



Association between effect sizes of pooled secondary endpoints in active treatment and sham arms. Linear regression, 95% confidence intervals. N=19.
67x50mm (300 x 300 DPI)


Hyaluronic acid injections for osteoarthritis of knee
Epidural injections for sciatica
Vertebroplasty for compression fracture
PMLR for angina
IDET/ PIRFT for discogenic low back pain
Arthroscopic debridement/ meniscectomy
Closure foramen ovale for migraine
Invasive procedures for Parkinson's disease
SR (Rutjes 2012)
SR (Pinto 2012)
SR (Shi 2012)
SR (McGillion 2010)
SR (Helm 2012)
MEDLINE
MEDLINE
MEDLINE
Records assessed for eligibility (n=47)
Records assessed for eligibility (n=6)
Records assessed for eligibility (n=2)
Records assessed for eligibility (n=3)
Records assessed for eligibility (n=3)
Records assessed for eligibility (n=2)
Records assessed for eligibility (n=1)
Records assessed for eligibility (n=7)
Records excluded, low quality (n=42)
Records excluded, low quality (n=1)
Records excluded, low quality (n=1)
Records excluded, ES not calculatable (n=2)
Records excluded, ES not calculatable (n=1)
Records excluded, ES not calculatable (n=3)
Studies included (n=3)
Studies included (n=4)
Studies included (n=2)
Studies included (n=2)
Studies included (n=3)
Studies included (n=2)
Studies included (n=1)
Studies included (n=4)

Appendix table 1. Search phrases, no of eligible studies and reasons for exclusion

Procedure	Search phrase MEDLINE	Source	Eligible studies	Excluded, ES not calculatable	Excluded, other methodo-logical reasons	Included studies
PMLR	Percutaneous myocardial laser revascularization	McGillion 2010 (17)	3	-	1	Salem 2004, Leon 2005
PIRFT /IDET	Intradiscal OR annular AND thermal AND "low back pain"	Helm 2012 (18)	3	-	-	Kvarstein, 2009
				-	-	Freeman 2005, Pauza 2003
Epidural injection corticosteroids	Epidural AND corticosteroid* AND sciatica	Pinto 2012 (16)	6	Karppinen 2001	1	Iversen 2011, Valat 2002, Arden 2005, Cohen 2012
Intraarticular hyaluronic acid for osteoarthritis knee	Hyaluron* OR viscosuppl* AND knee AND osteoarthritis	Rutjes 2012 (15)	47	Lundsgaard 2008, Petrella 2008	42	Chevalier 2010, Altman 2004, Pham 2004
Vertebroplasty	vertebroplast*	Shi 2012 (19)	2	-	-	Kallmes 2009, Buchbinder 2009
Invasive treatment of Parkinson's disease	transplantation OR gene OR "stem cell" AND Parkinson*	MEDLINE	7	Freed 2001, Gordon 2004, McRae 2004	-	Marks 2010, Olanow 2003, Gross 2011, LeWitt 2011
Arthroscopic debridement knee osteoarthritis	debridement AND lavage AND knee AND osteoarthr*	MEDLINE	1	-	-	Moseley 2002
Meniscectomy knee	meniscectomy AND knee	MEDLINE	1	-	-	Sihvonen 2013
Foramen ovale closure for migraine	"foramen ovale" AND migraine	MEDLINE	1	-	-	Dowson 2008
Number of trials			**71**	**6**	**44**	**21**

Appendix table 2. Indications, postulated mechanisms and history of selected interventions

Invasive procedure / indication	Postulated mechanism	History	References
Percutaneous myocardial laser revascularization / intractable angina pectoris	Increasing the delivery of oxygenated blood to poorly perfused myocardium by creating channels	Introduced in the 1980s, initially transmyocardial route, later percutaneous route, now mostly abandoned	Schofield PM, McNab D. NICE evaluation of transmyocardial laser revascularisation and percutaneous laser revascularisation for refractory angina. *Heart* 2010;96:312-3.
Patent foramen ovale closure with STARFlex Septal Repair Implant / migraine	Improvement of migraine headache, believed to block the formation of microembolies to the brain	Developed in the 1990s for the prevention of stroke, later thought to cure migraine, never in clinical use for this indication	Gornall J. A very public break-up. *BMJ* 2010;340:c110
Arthroscopic debridement / Knee osteoarthritis	Unclear, no documented effect on arthritic process, but about 50% report relief of pain (Mosely)	Annually about 650.000 procedures in the USA in the mid-ninetees, but 39% decrease between 2000 and 2008.	Holmes R, Moschetti W, Martin B, Tomek I, Finlayson S. Effect of evidence and changes in reimbursement on the rate of arthroscopy for osteoarthritis. *Am J Sports Med* 2013;41:1039-43.
Arthroscopic meniscectomy / degenerative meniscal lesions	Unclear, relief of symptoms attributed to trimming damaged meniscus down to viable meniscus and removing fragments.	The most common orthopedic procedure in the United States, 700.000 per year, up 50% last 15 years	Kim S, Bosque J, Meehan JP, Jamali A, Marder R. Increase in outpatient knee arthroscopy in the United States: a comparison of National Surveys of Ambulatory Surgery, 1996 and 2006. *J Bone Joint Surg Am* 2011;93:994-1000.
Viscosupplementation with hyaluronic acid / Knee osteoarthritis	Improve joint lubrication by increasing HA levels in joint, in spite of short half-lives (Marshall 2000)	Many positive reports since late 1980s, including sham-controlled trials. Still widely in use	Rutjes 2012 (15)
Percutaneous vertebroplasty with PMMA cement injection / vertebral compression fracture	Increase the strength of the damaged bone and alleviate pain by preventing microfractures	Numerous observational studies and single-blind trials reported substantial clinical benefits. Slight reduction of procedure since 2009	Manchikanti L, Pampati V, Hirsch JA. Analysis of utilization patterns of vertebroplasty and kyphoplasty in the Medicare population. *J Neurointerv Surg* 2013;5:467-72.
Epidural injection of corticosteroids / Sciatica	Dampen inflammatory reaction in nerve root sheaths caused by mechanical compression	Routinely used for sciatica since the 1950s (Pinto 2012). Since 2000 the number of injections increased by about 130% in the United States and 50% in the United Kingdom	Manchikanti L, Falco FJ, Singh V, Pampati V, Parr AT, Benyamin RM, Fellows B, Hirsch JA. Utilization of interventional techniques in managing chronic pain in the Medicare population: analysis of growth patterns from 2000 to 2011. *Pain Physician* 2012;15:E969-82

Percutaneous intradiscal radiofrequency and thermocoagulation (PIRFT and IDET) / discogenic low back pain	Placement of a electrode or RF-probe into the annulus and applying heat or current to destruct nociceptors/ annulus	Introduced in 1996 (IDET), later mostly abandoned	Helm 2012 (18)
Fetal nigral transplantation / Parkinson's disease	Restoration of dopamin levels in basal ganglia through injection of growth factors, GAD gene or nigral dopamine neurons	Based on animal models and a few small observational trials from about 2000. None in routine clinical use due to insufficient evidence	-
Gene delivery of AAV2-Neurturin / Parkinson's disease			
Transplantation of human retinal pigmental cells / Parkinson's disease			
Insertion of AAV-GAD gene into subthalamic nucleus / Parkinson's disease			

Appendix table 3. Included and excluded secondary endpoints.		
Author	**Included secondary endpoints**	**Excluded secondary endpoints (means not reported, or irrelevant)**
Leon 2005		
	Time to onset angina	Improvement in angina class
	Time to onset ST depression	Radioisotope imaging
	Overall health	
	Frequency angina	
	Stability angina	
	Physical functioning	
	Disease perception	
	Treatment satisfaction	
	PCS	
	MCS	
Salem 2004		
		Proportion improved CCS angina class
		Medication usage
		Seattle Angina Questionnaire
		Left EF
		Angina stability
		Angina frequency
		Physical limitation
		Treatment satisfactioin
		Disease perception
Sihvonen 2013	WOMET score	-
	Knee pain at rest	
	Knee pain after exercise	
	15D score	
Moseley 2002		
	Arthritis Impact Scale	-
	Physical functioning Scale	
	Walking-bending	
	SF-36 Pain	
	SF-36 Physical functioning	
Pham 2004		
	Lequesne's algofunctional index	-
	Global assessment	
	% painful days	
Chevalier 2010		
	Womac C function	-
Altman 2004		
	Womac stiffness	-
	Womac physical	
Kallmes 2009		
	Pain Intensity	Opioid use

	SF-36 PCS	
	SF-36 MCS	
	Pain Frequency Index	
	Pain Bothersomeness Index	
	EQ-SD Index	
	SOF-ADL	
Buchbinder 2009		
	Roland-Morris Disability Questionnaire	-
	Life Questionnaire of the European Foundation	
	European Quality of Life–5 Dimensions	
Cohen 2012		
	Oswestry Disability Index	-
	Back pain	
Arden 2005		
	Leg pain	Analgesic use
	Back pain	
Valat 2002		
	Roland-Morris Disability Questionnaire	Dallas Pain Questionnaire
	Straight leg raising	
	Schober's test	
Iversen 2011		
		VAS back and leg pain, European Quality of Life scale
Freeman 2005		
	Modifiede Somatic Perception Questionnaire	SF-36 Mental, Role Physical/ Mental, Social Function
	Low Back Pain Outcome Score	
	SF-36 Physical Function	
	SF-36 Pain	
	SF-36 General Health	
	SF-36 Vitality	
Pauza 2003		
	VAS Pain	-
	SF-36 Physical Function	
	SF-36 Pain	
Kvarstein 2009		
	SF-36 Bodily pain	SF-36 Mental, Role Physical/ Mental, Social Function
	SF-36 Physical function	
	Oswestry Disability Index	
	SF-36 General health	
	SF-36 Vitality	
Olanow 2003		
	UPDRS motor on	Mean L-dopa dose equivalents
	UPDRS ADL off	
	UPDRS ADL on	

	% Off time day	
	% On time without dyskinesia	
Marks 2010		
	UPDRS OFF 1	Mean L-dopa dose equivalents
	UPDRS OFF 2	
	UPDRS ON 1	
	UPDRS ON 2	
	UPDRS ON 3	
	On without dyskinesia	
	On with dyskinesia	
Gross 2011		
	UPDRS ON	Mean L-dopa dose equivalents
	UPDRS ADL	
LeWitt 2011		
	UPDRS 1	Timed walking
	UPDRS2	BPRS other than taps
	UPDRS4	Dyskinesia rating scale
	Schwab and England ADL scale	Patient's diary
	BPRS taps 60 s	Clinical global impression
	Hoehan and Yahr stage	
	PDQ-39 total	
Dowson 2008		
	Headache Impact Test	-

PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	2
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	4
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	4
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	5
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	5
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	5, Appendix table 1
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	5, Appendix 1
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	5-6
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	5-6
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	5
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	6
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	6

PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	5
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	6, 7, Appendix Flow chart
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	7-13
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	7,9 Appendix table 1
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	10-13
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	10-13
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	7,9
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	14, Fig. 3,4
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	14,15-17
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	13-17
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	18
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	29

PRISMA 2009 Checklist

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(6): e1000097. doi:10.1371/journal.pmed1000097

For more information, visit: **www.prisma-statement.org**.

Page 2 of 2

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

Robin Holtedahl, Jens Ivar Brox, Ole Tjomsland

Fram Rehabilitation Centre, Rykkinveien 100, 1349 Rykkin, Norway Robin Holtedahl Consultant Department of Physical Medicine and Rehabilitation, Oslo University Hospital, Box 4956 Nydalen, 0424 Oslo, Norway Jens Ivar Brox Professor South-Eastern Norway Regional Health Authority, PB 404, 2303 Hamar, Norway Ole Tjomsland Director of Quality and Specialist Areas

Correspondence to: Dr Robin Holtedahl; robi-hol@online, telephone +4790248973

Key words:
Placebo effects
Invasive procedures
Biomedical ethics
Evidence based health care

Word count: 3783

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 ~~Exploratory~~A s~~S~~ystematic review and meta-analysis.

Data Sources and Study Selection We searched MEDLINE and Cochrane library ~~for~~ to identify systematic reviews of musculoskeletal, neurological and cardiac~~ological~~ conditions published between January 2009 and January 2014 ~~including~~ ~~and randomised clinical trials~~ ~~musculoskeletal, neurological and cardiological conditions~~ comparing selected minimally invasive ~~procedures~~ with placebo (sham) procedures. ~~We selected the most recent systematic review with low risk of bias published in core medical journals.~~ We searched ~~For procedures that were not evaluated in systematic reviews we searched~~ MEDLINE for additional randomised controlled trials published between January 2000 and January 2014. ~~trials with low risk of bias.~~

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

Results We included 2~~2~~1 trials involving 2519~~472~~ adult participants. For primary endpoints, there was a large clinical effect (ES ≥0.8) after active treatment in 12 trials and after sham procedures in 11 trials. For secondary endpoints, seven and five trials showed a large clinical effect, respectively. Three trials showed a moderate difference in ES between active treatment and sham on primary endpoints (ES ≥0.5) but no trials reported a large difference. No trials showed large or moderate differences in ES on pooled secondary endpoints. Regression analysis of endpoints in active treatment and sham arms estimated an R^2 of 0.7~~9~~8 for primary and 0.84 for secondary endpoints. Adverse events after sham were in most cases minor and of short duration.

Conclusion The generally small differences in effect size between active treatment and sham suggest that non-specific mechanisms, ~~principally~~ 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~~sham-controlled trials were generally not substantiated.

SUMMARY

Article focus

- Many minimally invasive procedures have gained increased popularity during the last two decades in spite of limited evidence of their clinical effectiveness.
- Systematic review and meta-analysis of published randomised double-blind placebo-controlled studies of minimally invasive procedures, with special emphasis on the magnitude of change in the placebo (sham) arms.
- Assessment of adverse events in the trials' active treatment and placebo arms.

Key messages

- The magnitude of change in the active treatment- and placebo arms varied greatly, but about 80% of the variancetion in effect size of active treatment could be explained predicted by placebo effects, regression to the mean or spontaneous improvement.
- Adverse events related to sham procedures were mainly minor and short-lived, and frequently outweighed by positive placebo effects.

Strengths and limitations

- Strict selection criteria of trials, with low risk of biasmainly based on high-quality systematic reviews with low risk of bias.
- Calculation of effect sizes on primary and pooled secondary endpoints both in active treatment and shamin both active treatment and sham arms.
- Heterogenous interventions, outcome measures and timing of assessment.
- Searches limited to MEDLINE and Cochrane library

4

INTRODUCTION

It is normally assumed that medical practices are based on firm clinical evidence, and that new practices or techniques are ~~not~~ introduced ~~before~~ when superiority, or at least non-inferiority, has~~, has~~ been demonstrated compared to established treatments. However, medical history reveals numerous examples contradicting this assumption. Forty-two percent 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. 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~~study, we analysed placebo-controlled trials of ~~selected~~ minimally invasive ~~procedures~~interventions in musculoskeletal, neurological and cardiac~~ological~~ conditions. The aims were threefold: (a) to assess the magnitude of change in outcome from baseline to trial endpoint in both the active treatment and placebo (sham) arms,~~;~~ (b) to explore the contribution of non-specific factors, including placebo, to the outcome of active treatment, and (c) to assess the level of reported adverse effects in both trial arms.

METHODS

Search strategy and selection criteria

The main focus was evaluation of minimally invasive procedures that were claimed to have substantial clinical effects in cardiological, neurological and musculoskeletal conditions. We excluded oOpen surgical interventions were excluded.

We first conducted electronic searches for randomised placebo-controlled trials of minimally invasive interventions for cardiacological, neurological and selected musculoskeletal conditions, using MEDLINE and Cochrane library to identify systematic reviews published from between January 2009 to and January 2014-. 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 excluded open surgical and laparoscopic interventions. Where applicable, we used the "core clinical journals" filter in PubMed, which is an index of journals particularly relevant to practicing physicians. From the selected reviews, Wwe selected randomised placebo-controlled randomised placebo-controlled trials published from January 2000 to January 2014 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 both 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[2911]. Two of the authors (RH and JIB) independently assessed analysed the five methodological criteria in the RCTs included selected from systematic reviews. to ascertain that they complied with the five criteria were fulfilled.

For interventions that were not evaluated in systematic reviews, Wwe next searched MEDLINE and Cochrane library 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 mentioned aboverisk of bias in the additional RCTs that were identified from this search. not selected from systematic reviews, based on the same five criteria that were used for the selection of trials from systematic reviews.

6

The searches provided trials of the following interventions: percutaneous laser revascularisation of myocardium for angina pectoris, closure of foramen ovale for migraine, arthroscopic meniscectomy for meniscal tears, debridement and injection of hyaluronic acid for symptomatic osteoarthritis of the knee, epidural injections of corticosteroids for sciatica, percutaneous heating of the intervertebral disc for chronic low back pain, vertebroplasty for vertebral body fractures, and injection or transplantation of biological tissue for Parkinson's disease.

The rationale for the introduction of most of these interventions is that a physiological derangement can be brought back to an original, healthy state by invasive techniques. Promising results in initial pragmatic uncontrolled trials in some cases led to widespread clinical implementation, even though some subsequent larger and methodologically more rigorous trials failed to replicate the initial findings. Another common feature of the included interventions is that their rationale is primarily based on improvements in subjective outcome, including pain and health related quality of life.

The searches provided no sham-controlled trials of percutaneous heating of the cervical intervertebral disc, lumbar facet joint injections, chemonucleolysis, transmyocardial laser revascularization for angina, deep brain stimulation for Parkinson's disease or arthroscopic procedures (other than knee conditions). No studies of radiofrequency denervation or intradiscal steroid injection for low back pain were found that provided SD, which is a requirement for calculation of effect size.

From the most recently published systematic review of each procedure, we selected randomised placebo-controlled trials that according to the review fulfilled at least four of the following criteria: random allocation, allocation concealment, blinding of participant, blinding of assessor and intention-to-treat analysis. For procedures that were not evaluated in systematic reviews, we searched MEDLINE and Cochrane library for randomised placebo-controlled trials. Two of the reviewers (OT and JIB) independently assessed the risk of bias in the RCTs that were not selected from systematic reviews, based on the same five criteria that were used for the selection of trials from systematic reviews.

Only English language journals were included. We excluded crossover trials, trials that did not report results as means, standard deviation, standard error or confidence intervals in both active and sham-groups, as well as trials with only graphic representation of data. We excluded reviews with declared commercial conflicts of interest in order to avoid the risk of financially motivated selection of trials in these reviews.

7

~~Commercially funded RCTs were not excluded, because all the included trials had been screened for bias either by the high-quality systematic reviews or by two of the authors, using strict methodological criteria.~~ Details of the search strategy are shown in web appendix table 1 and web appendix figure 1. We give aA short description of each procedure's introduction, therapeutic rationale and history, is given in web appendix table 2. This review is reported in accordance with the PRISMA statement.[12]

Data extraction

We registered all continuous primary endpoints. In trials without continuous primary endpoints, with multiple endpoints or no defined primary endpoint, we selected an outcome related to pain or condition-specific endpoint. 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 endpoint. We also registered secondary endpoints in order to avoid potential bias from selective reporting in the included trials. The included and excluded secondary endpoints are shown in web appendix table 3. Endpoints describing medication, radiographic or physiological variables, social or psychological function, were not included. For the ~~Parkinson-trials~~Parkinson-trials, only endpoints in the off-medication state were registered. Results fromon 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 in both trial arms were registered (Ttables 1 and 2).

Data synthesis

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

1

8

with a recommended procedure.[14] Unadjusted linear regression analyses were used to explore the association between outcome in the active and sham groups both for primary and pooled secondary endpoints. For this analysis, we used Medcalc Statistical Software version 12.7.4.0[15]

RESULTS

Selection of interventions

The searches provided sham-controlled trials of the following interventions: percutaneous laser revascularisation of myocardium for angina pectoris, closure of foramen ovale for migraine, arthroscopic meniscectomy for meniscal tears, debridement and injection of hyaluronic acid for symptomatic osteoarthritis of the knee~~, epidural injections of corticosteroids for sciatica, percutaneous heating of the intervertebral disc for chronic low back pain (two techniques), vertebroplasty for vertebral body fractures,~~ and injection or transplantation of biologically active material for Parkinson's disease (human retinal pigmental cells, fetal nigral cells and Neurturin ~~3 techniques~~). Because of the large number of described interventions for neck- and back pain syndromes, we chose to restrict the analysis to sham-controlled trials of the following interventions:

~~No studies of radiofrequency denervation or intradiscal steroid injection for low back pain were found that provided SD, which is a requirement for calculation of effect size. The~~ epidural injections of corticosteroids for sciatica (~~sacral~~caudal, interlaminar ~~or~~ and transforaminal routes), percutaneous heating of the intervertebral disc for chronic low back pain (percutaneous intradiscal radiofrequency thermocoagulation ~~or~~ and intradiscal electrothermal therapy) and vertebroplasty for vertebral body fractures. The searches provided no sham-controlled trials of arthroscopic procedures other than knee conditions.

Study selection

The study selection process is summarised in web appendix figure 1. Web appendix table 1 shows the excluded trials and the reasons for exclusion. The search provided five systematic reviews, all identified through searches in MEDLINE, none were commercially funded.[16-20] It identified a total of 7~~1~~1~~0~~ clinical trials, t~~en~~welve of them were not identified from the systematic reviews. Forty-~~three~~ four trials were excluded for methodological reasons, principally ~~due to~~ risk of bias. Six additional trials were excluded because ES could not be calculated.[21-26] Finally, 2~~2~~1 clinical trials with a total of 25~~72~~19 participants were included in the present review (table 1). Trial interventions in active treatment and sham arms are also shown.

Table 1. **Included 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 2005	Food and Drug Administration / NC	Percutaneous myocardial laser revascularization / intractable angina pectoris	Laser turned on but no procedure performed	MAE in hospital (high dose): 4.1%	MAE in hospital: 0
Salem 2004	Ethics committee / NC			No procedural AE	
Sihvonen 2013	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 2002	Review Board / NC	Arthroscopic debridement / Knee osteoarthritis	Simulated arthroscopy preparation, intravenous anaesthesia, skin incisions, no instruments entered knee, knee manipulated	No procedural AE	
Pham 2004	Review Board / NC	Hyaluronic acid / Knee osteoarthritis	Intraarticular injection of saline solution	No MAE	
				Any mAE: 81.7%	Any mAE: 1.2%
Altman 2004	Ethics committee / C			No MAE	
				mAE: 12.8%	mAE: 8%
Chevalier 2010	ClinicalTrials.org / C			No MAE	
				mAE: 35,8%	mAE: 33,8%
Kallmes 2009	Review Board / NC	Percutaneous vertebroplasty with PMMA cement injection / vertebral compression fracture	Conscious sedation + local anesthaesia, pressure put on spine, simulation of odor with mixing of PMMA to imitate the smell during the active procedure	No MAE	
				mAE: 14%	mAE: 16%
Buchbinder 2009	Ethics committee at each participating center~~Australian Clinical Trial Register~~ / NC		Conscious sedation + local anesthaesia, needle inserted to rest on the lamina, PMMA container opened to imitate the smell during the active procedure	No procedural AE	

Cohen 2012	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 2005	Ethics committee / NC		2 mL saline into interspinous ligament	No MAE mAE: 9%	mAE: 10%
Valat 2002	Ethics committee / NC		2 mL saline into epidural space, interlaminar approach	No MAE mAE: 6%	mAE: 8%
Iversen 2011	Ethics committee / NC		Subcutaneous injection of 2 mL saline superficial to the sacral hiatus	Not reported	
Freeman 2005	Ethics committee / C	Intradiscal electrothermal therapy (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 2003	Review Board / NC		17-gauge needle introduced onto the outer annulus, mock electrode passage shown on monitor, generator noises produced	Not reported	
Kvarstein 2009	Ethics committee / NC	Percutaneous intradiscal radiofrequency thermocoagulation (PIRFT) / discogenic low back pain	17-gauge canula and RF-probe inserted into annulus, no RF current applied	Not reported	
Olanow 2003	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 2010	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 2011	Review Board /	Transplantation of	Scalp incisions,	1 death	0 deaths

	C	human retinal pigmental cells / Parkinson's disease	partial thickness burr holes, no cell transplantation	MAE: 23%	MAE: 0
LeWitt 2011	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 2008	Ethics committee / C	Patent foramen ovale closure with STARFlex Septal Repair Implant / migraine	General anesthesia, skin incision in the groin	MAE (possibly or probably related to procedure): 11%	MAE (possibly or probably related to procedure): 4%

C=commercial; NC=non-commerical; MAE=major adverse events; mAE=minor adverse events; PMMA=polymethylmethacrylate; AAV2 =adeno-associated; GAD=glutamic acid decarboxylase

F~~if~~our~~if~~teen trials ~~were selected~~ from the systematic reviews fulfilled at least four of the five methodological criteria.~~.~~[27 28 31-4~~3~~2] ~~, one trial did not fulfil the methodological criteria and was excluded.~~ Seven trials ~~were selected~~ provided through searches in MEDLINE fulfilled the same criteria.[29 ~~28~~ 30 4~~4~~3-4~~8~~7] ~~and all these trial fulfilled at least four of the five methodological criteria. The two authors who independently screened the individual trials, with special emphasis on concealment of treatment allocation and blinding, found the risk of bias to be generally low.~~

All trials reported approval of study protocol prior to patient enrolment (table 1). Eight trials were commercially funded.[32 33 41 44~~4~~-47~~8~~] Most of the trials had few participants, ranging from 20 to 346 (median 80).

Clinical outcomes after active treatment and sham

Twelve of the 2~~2~~1 trials showed a large ES on primary endpoints after active treatment, while 11 trials showed a similar ES after the sham procedure (figure 1, table 2).

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

Author / procedure	**Limit disease duration / time to follow-up (months)**	**Trial arm / no of patients randomised**	**ES primary endpoint**	**ES pooled secondary endpoints (no of endpoints)**

Leon 2005 / Percutaneous myocardial laser revascularization	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 revascularization	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**
~~**Petrella 2002 /** Hyaluronic acid~~			~~Womac pain~~	~~(8)~~
	~~None / 1~~	~~Active / 25~~	~~0.35~~	~~0.54~~
		~~Sham / 28~~	~~0.15~~	~~0.40~~

13

~~ES active treatment vs sham~~			~~**0.22**~~	~~**0.14**~~
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
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		0.99

1

			1.47	
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
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		0.30

			1.00	
		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**
VAS=Visual Analogue Scale; NRS=Numerical Rating Scale; UPDRS=Unified Parkinson's Disease Rating Scale; Womac=Western Ontario and McMaster Universities Osteoarthritis Index				

ES on primary endpoints was moderate in ~~four~~ three of the active treatment groups and in two of the sham groups.

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

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

Differences in outcome between active treatment and sham
Better results on primary endpoints were reported with active treatment compared to sham in 1~~5~~4 of the 2~~2~~1 trials, but the differences were small. Three trials (one epidural study[3~~8~~7], one discogenic pain study[40~~1~~] and one Parkinson study[4~~5~~6]) reported a moderate effect but none showed a large effect (figure 2, table 2). Seven trials reported a better primary endpoint outcome after sham than after active treatment.

Nineteen trials reported ~~On~~ secondary endpoints, 1~~2~~1 of these~~/~~19~~20 trials~~ reported better outcome after active treatment than after sham, but in no case did the differences reach a moderate ES (figure 2, table 2). In twelve ~~132/1920 trials~~trials, ~~with both primary and secondary endpoints,~~ the outcome was better for primary than for pooled secondary endpoints. This bore no relation to funding source.

On regression ~~analyses~~analyses, effect sizes in the sham groups ~~explained~~ predicted about 80 % of the variance ~~tion~~ of

16

ES in the active treatment groups, both on primary and pooled secondary endpoints (figure 3 and 4).

Adverse events

EighNineteen of the 221 studies provided information about adverse events (AE) (table 1). Three of these trials reported no procedural adverse events in any of the groups.[27 29 356] Major AEs were reported after active treatment in four trials[28 445 456 478] including one death in one of the Parkinson studies.[465] In the sham groups, one trial[478] listed three major AEs possibly or probably related to the procedure, all thought to be caused by anti-platelet 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.

DISCUSSION

Principal findings

AOur analysis of these 21 selected sham-controlled trials of minimally invasive procedures showed that the general lack of clinicaleffect sizes effect in the active in the selected trials arms werewas predicted by the effect sizes mainly due to large effects in the sham arms and not to small effects in the active treatment arms. There was In these 221 selected sham-controlled trials of invasive procedures, there was a large clinical ESeffect on primary endpoints in about half of both similar number of the active and sham interventions, but none of the trials showed a large . The difference in ES ESeffect between active treatment and sham groups eitherm on primary or secondary endpoints.

T was moderate in three trials, while none demonstrated a large effect. On pooled secondary endpoints, none of the trials showed even a moderate clinical effect.

Our analysis of effect sizes showed that the general lack of clinical effect in the selected trials was mainly due to large effects in the sham arms and not to small effects in the active treatment arms. However, 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 endpoints varied from around zero 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 three, and in the epidural trials by a factor of two. 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 endpoints 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 study

It is our opinion that the calculation of effect sizes in both active treatment and placebo arms is a strength of the present study. This made it possible to- assess the magnitude of change in both arms as well as and the contribution of non-specific factors to change in the active treatment arms. The calculation of effect sizes provides an alternative assessment to probability estimates. Another strength of the study is the supplementary analyses of pooled secondary endpoints, enabling a more comprehensive evaluation than using primary endpoints alone. Reports of tactically motivated use of primary and secondary endpoints before publication in order to improve study results strengthen the argument for registering all relevant secondary endpoints.[48] Our finding that a majority of trials reported better results on primary than on secondary endpoints 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 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.

We applied principles from guidelines for conducting systematic reviews and meta-analyses and included an . The independent assessment of methodological trial quality by two of the authors. authors gives added confidence in the trial selection. WHowever, we cannot rule out that we have missed relevant trials because we limited our search to the Cochrane Library and MEDLINE, but 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 realize that this selection process and the fact that we relied on previous methodological evaluations may have contributed to unrecognised selection bias.

The present It must be emphasised that this review is exploratory, being limited to some selected minimally invasive procedures in cardiology, neurology, and musculoskeletal

18

conditions. to certain conditions and interventions, and also excluding interventions for life threatening conditions. We applied principles from guidelines for conducting systematic reviews and meta-analyses. By 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 realize that this selection process and the fact that we relied on previous methodological evaluations may contribute to unrecognised selection bias. The strengths of the present systematic review include the use of strictly defined selection criteria to minimise bias.

For five six of the eighttwelvenine procedures we identified selected trials that, according to from the most recent systematic reviews published in core clinical journals, fulfilled at least four of the following criteria: random allocation, allocation concealment, blinding of participant, blinding of assessor and intention-to-treat analysis.

For the remaining , all of them by authors without declared commercial interests. From these reviews, we selected trials that complied with a set of predefined methodological criteria. tThreefour proceduressix four procedures, additional trials that were identified by directly through MEDLINE searches and the same criteria were used to assess bias. We cannot rule out that we have missed relevant trials because we limited our search to the Cochrane Library and MEDLINE, but most relevant trials are likely to have been identified by our searches. It must be emphasised that our limitation to certain conditions, as well as the heterogeneous nature of selected interventions, imply that our findings cannot be generalised to minimally invasive procedures in all medical disciplines. We believe, however, that the same methodology could be applied to more systematic analyses of the role of placebo effects in other conditions and procedures.

We applied principles from guidelines for conducting systematic reviews and meta-analyses. By selecting core journals and trials that had previously been methodologically evaluated in systematic reviewsreviews, it was our intention to reduce the risk of bias by excluding studies of low quality. We realize that this selection process and the fact that we relied on previous methodological evaluations may contribute to unrecognized selection bias. We also emphasise that our limitation to certain conditions and highly heterogeneous interventions implies that our findings cannot be generalised to minimally invasive procedures in all medical disciplines.

The calculation of effect sizes in both active treatment and placebo arms enabled us to assess the magnitude of change in both groups. This in turn made it possible, through regression analysis, to show that non-specific effects, including placebo, can largely explain the variation in outcomes after the active

19

interventions. The calculation of effect sizes provides a better assessment of clinically important effects than using probability estimates, and supplementary analyses of pooled secondary endpoints contribute to a more comprehensive evaluation than using primary endpoints alone. Reports of tactically motivated manipulation of primary and secondary endpoints before publication in order to improve study results are also arguments in favour of registering all relevant secondary endpoints.[49] Our finding that a majority of trials reported better results on primary than on secondary endpoints might lend support to such a hypothesis. However, according to the authors, all trials had sought and gained approval of the protocol from ethics committee and/ or review board (table 1).

The described indications and procedures are heterogeneous, encompassing both neurological, orthopaedic and cardiological specialties. While some procedures are, or have been, in wide clinical use, some are still in the clinical trial phase. are still considered experimental. Other sources of heterogeneity are variable duration of disease before inclusion, the selection of outcome measures and time to follow-up. RThough esults our findings cannot be generalised to minimally invasive procedures in all medical disciplines, but we hypothesise that a similar methodology could be applied to more systematic analyses of the role of non-specific effects in other minimally invasive procedures.
We emphasise that our limitation to certain conditions and interventions implies that our findings cannot be generalised to minimally invasive procedures in all medical disciplines. Other sources of heterogeneity wereare variable duration of disease before inclusion, time to follow-up and variable and outcome measures. The contribution of spontaneous improvement relative to placebo effect might be expected to be greater with longer time to follow-up.

screened for bias using the same methodology.

Moreover, we Our calculation of ed the effect sizes in both active treatment and placebo arms, enabled ing us to assess the magnitude of change in both groups. This in turn made it possible, through regression analysis, to show that non-specific effects, including placebo, can largely explain the variation in outcomes after the active interventions. The calculation of effect sizes provides a better assessment of clinically important effects than using probability estimates, and supplementary analyses of pooled secondary endpoints contribute to a more comprehensive evaluation than using primary endpoints alone. Reports of tactically motivated manipulation of primary and secondary endpoints before publication in order to improve study results are also arguments in favour of registering all relevant secondary endpoints.[49] Our finding that a majority of

1

20

trials reported better results on primary than on secondary endpoints might lend support to such a hypothesis. However, according to the authors, all trials had sought and gained approval of the protocol from ethics committee and/ or review board (table 1).

The present study has several potential limitations. The described indications and procedures are heterogeneous, encompassing both neurological, orthopaedic and cardiological specialties. While some procedures are, or have been, in wide clinical use, some are still considered experimental. Duration of disease before inclusion, time to follow-up and outcome measures varied considerably, adding to the heterogeneity. The contribution of spontaneous improvement relative to placebo effect might be expected to be greater with longer time to follow-up. We cannot exclude that we have missed may have missed relevant trials because we limited our search to the Cochrane Library and MEDLINE, but most relevant trials are likely to be identified by our searches. or because of publication bias in the MEDLINE searches, though this is less likely for trials selected from the included systematic reviews. 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.

We applied principles from guidelines for conducting systematic reviews and meta-analyses and included an . The iindependent assessment of methodological trial quality by two of the authors. authors gives added confidence in the trial selection. WHowever, we cannot rule out that we have missed relevant trials because we limited our search to the Cochrane Library and MEDLINE, but 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 realize 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.[5049] More recently, Bickett et al. hypothesised that epidural injection of small volumes of saline might have physiological effects.[501] 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.[512] In spite of their heterogeneous nature, the 2~~2~~1 selected trials share a medico-technological context in which an a priori enhanced placebo response could be expected. If an ES >0.8 is considered as mega-placebo, ~~nearly~~ 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 endpoints in the active treatment and sham 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.[26 523]

Non-specific factors

The role of non-specific factors, primarily spontaneous remission or statistical regression-to-the-mean, in placebo-controlled studies is controversial.[534] 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 effect sizes in the range of 0.2 to 0.3. [545] 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.[565] Of 198 included trials fourteen had a no-treatment arm. The mean ES in the placebo groups was about 0.5, while it was only slightly above zero 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 endpoints. 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 (i.e. waiting list), we cannot rule out that the changes appearing during the trial period also reflect non-specific factors, i.e. 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 trials, there were moderate to large improvements in the sham groups even at one-year follow-up.[443-465] Other authors have also found improvements several years after sham surgery, indistinguishable from conventional surgery.[26 567] This is in agreement with recent insights into the neurobiological effects of placebo and their relation to underlying psychological mechanisms, principally expectation and conditioning.[578]

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.[589] 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 heterogeneity of the interventional techniques and variable outcome specifications, making standardization difficult to achieve. Existing ethical guidelines accept the role of placebo-controlled trials when certain conditions are met.[60-59] There must be genuine equipoise, i.e. 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.[60 61 62]

The selected trials gave a detailed description of adverse events in both active and sham-treated groups (table 1). The safety concerns frequently raised as an argument against the use of sham were generally not supported. Major adverse events related to the sham procedure were reported in only one of the trials[47~~8~~] and they were short-lived and not life threatening. Minor adverse events were more frequent, but also of limited duration. Positive placebo-induced effects generally outweighed adverse events, thus weakening ethical arguments against the use of sham interventions. In our opinion, the ~~ethical~~ 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 symptoms, research ethics committees should consider such factors in their risk-benefit assessments of planned sham controlled trials.[62 63 ~~6~~4]

~~Behov for mer homogene pas mat~~

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 -facet joint injections, radiofrequency neurotomy, acromioplasty, percutaneous coronary intervention and, more recently, robotic surgery.[6~~5~~4-~~70~~69] 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 of 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 ~~associated with minimally invasive~~~~the selected interventions explain~~ explain a large part of their purported benefits ~~of the selected procedures~~. Further, results indicate that the risks of adverse events in sham-controlled trials are overrated, and ~~. The risks are~~ could be considered ~~, and in many cases could might be viewed as~~ acceptable ~~, not least~~ in view of the potential ~~for large~~ personal harm and societal ~~costs harm~~ costs ~~and~~ associated ~~dubious ethics of~~ with ~~the continued use of~~ unproven minimally invasive interventions.

Figure legends

Figure 1. Effect sizes of active treatment and sham, primary endpoints.

Figure 2. Differences in effect size between active treatment and sham.

Figure 3. Association between effect sizes of primary endpoints in active treatment and sham arms. Linear regression, 95% confidence intervals. N=2~~2~~1.

Figure 4. Association between effect sizes of pooled secondary endpoints in active treatment and sham arms. Linear regression, 95% confidence intervals. N=~~20~~19.

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.

25

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-S15;

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**:915-9.

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. 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.

17. 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.

18. 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.

19. 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-E304.

20. 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.

21. 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.

22. 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.

23. Karppinen J, Malmivaara A, Kurunlahti M et al. Periradicular infiltration for sciatica: a randomized controlled trial. *Spine* 2001; **26**:1059-67.

24. Freed CR, Greene PE, Breeze RE et al.Transplantation of embryonic dopamine neurons for severe Parkinson's disease. *N Engl J Med* 2001;**344**:710-9.

25. 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.

26. 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.

27. 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.

28. 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**:1812-9.

29. 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.

30. 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.

31. 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**:1611-7.

32. 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.

33. 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

29

randomised, multicentre, double-blind, placebo controlled trial. *Ann Rheum Dis* 2010;**69**:113-9.

~~34.Petrella RJ, DiSilvestro MD, Hildebrand C. Effects of hyaluronate sodium on pain and physical functioning in osteoarthritis of the knee: a randomized, double-blind, placebo-controlled clinical trial. *Arch Intern Med* 2002;**162**:292-8.~~

~~35.~~34. 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.

~~36.~~35. 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.

~~37.~~36. 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.

~~38.~~37. 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.

~~39.~~38. Arden NK, Price C, Reading I et al. A multicentre randomized controlled trial

1

of epidural corticosteroid injections for sciatica: the WEST study. *Rheumatology* 2005;**44**:1399-406.

~~40.~~39. 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.

~~41.~~40. 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.

~~42.~~41. 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.

~~43.~~42. 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.

~~44.~~43. 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.

~~45.~~44. 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.

~~46.~~45. 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.

~~47.~~46. 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.

~~48.~~47. 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.

~~49.~~48. 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.

~~50.~~49. 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.

~~51.~~50. Bicket MC, Gupta A, Brown CH 4th 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.

~~52.~~51. Kaptchuk TJ, Goldman P, Stone DA et al. Do medical devices have enhanced placebo effects? *J Clin Epidemiol* 2000;**53**:786-92.

~~53.~~52. 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.

~~54.~~53. Turner JA, Deyo RA, Loeser JD et al. The importance of placebo effects in pain treatment and research. *JAMA* 1994;**271**:1609-14.

~~55.~~54. Hróbjartsson A, Gøtzsche PC. Placebo interventions for all clinical conditions. *Cochrane Database Syst Rev* 2010; CD003974.

~~56.~~55. 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.

33

~~57.~~56. Marchand S, Kupers RC, Bushnell MC et al. Analgesic and placebo effects of thalamic stimulation. *Pain* 2003;**105**:481-8.

~~58.~~57. Finniss DG, Kaptchuk TJ, Miller F et al. Biological, clinical, and ethical advances of placebo effects. *Lancet* 2010;**375**:686-95.

~~59.~~58. 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. doi: 10.1186/1745-6215-12-50.

~~60.~~59. Snyder L. Review of the American College of Physicians Ethics Manual, Sixth Edition. *Ann Intern Med* 2012;**156**(1_Part_2):73-104.

~~61.~~60. Flum DR. Interpreting surgical trials with subjective outcomes: avoiding UnSPORTsmanlike conduct. *JAMA* 2006;**296**:2483-5.

~~62.~~61. Heckerling PS. Placebo surgery research: a blinding imperative. *J Clin Epidemiol* 2006;59:876-80.

~~63.~~62. 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.

~~64.~~63. Redberg RF. Sham controls in medical device trials. *N Engl J Med* 2014;**371**:892-3.

~~65.~~64. Haahr JP, Østergaard S, Dalsgaard J et al. Exercises versus arthroscopic decompression in

34

patients with subacromial impingement: a randomised, controlled study in 90 cases with a one year follow up. *Ann Rheum Dis* 2005;**64**:760-4.

~~66.~~65. Ketola S, Lehtinen J, Rousi T et al. No evidence of long-term benefits of arthroscopicacromioplasty in the treatment of shoulder impingement syndrome: Five-year results of a randomised controlled trial. *Bone Joint Res* 2013;**2**:132-9.

~~67.~~66. 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.

~~68.~~67. 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**,Supplement:S142-7.

~~69.~~68. Vitale MA, Arons RR, Hurwitz S et al.The rising incidence of acromioplasty. *J Bone Joint Surg Am* 2010;**92**:1842-50.

~~70.~~69. 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.

35

Contributors: RH initiated and planned the project and searched databases. JIB and OT assisted in developing search strategies. Article screening and data extraction was carried out by RH. Quality of data extraction and checking was carried out by JIB and OT. Statistical analysis was undertaken by RH, who also wrote the draft. OT and JIB reviewed the draft and contributed to manuscript revisions. RH is the guarantor for this study.

Funding sources: None.

All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

Ethical approval: Ethical approval was not required for this work.

Data sharing: Dataset can be obtained from Robin Holtedahl (robi-hol@online.no).

The lead author affirms that this manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned (and, if relevant, registered) have been explained.

1

BMJ Open

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

Journal:	*BMJ Open*
Manuscript ID:	bmjopen-2014-007331.R1
Article Type:	Research
Date Submitted by the Author:	18-Dec-2014
Complete List of Authors:	HOLTEDAHL, ROBIN; Fram Rehabilition Centre, Tjomsland, Ole; South-Eastern Norway Regional Health Authority, Division of Quality and Specialist Areas Brox, Jens; Oslo University Hospital, Orthopaedic
Primary Subject Heading:	Evidence based practice
Secondary Subject Heading:	Evidence based practice, Ethics, Research methods, Health services research
Keywords:	Quality in health care < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Health economics < HEALTH SERVICES ADMINISTRATION & MANAGEMENT, Clinical trials < THERAPEUTICS, MEDICAL ETHICS, PAIN MANAGEMENT

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

Robin Holtedahl, Jens Ivar Brox, Ole Tjomsland

Fram Rehabilitation Centre, Rykkinveien 100, 1349 Rykkin, Norway Robin Holtedahl Consultant Department of Physical Medicine and Rehabilitation, Oslo University Hospital, Box 4956 Nydalen, 0424 Oslo, Norway Jens Ivar Brox Professor South-Eastern Norway Regional Health Authority, PB 404, 2303 Hamar, Norway Ole Tjomsland Director of Quality and Specialist Areas

Correspondence to: Dr Robin Holtedahl; robi-hol@online, telephone +4790248973

Key words:
Placebo effects
Invasive procedures
Biomedical ethics
Evidence based health care

Word count: 5801

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 endpoints were calculated. Linear regression was used to analyse the association between endpoints 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 endpoints, there was a large clinical effect (ES ≥0.8) after active treatment in 12 trials and after sham procedures in 11 trials. For secondary endpoints, seven and five trials showed a large clinical effect, respectively. Three trials showed a moderate difference in ES between active treatment and sham on primary endpoints (ES ≥0.5) but no trials reported a large difference. No trials showed large or moderate differences in ES on pooled secondary endpoints. Regression analysis of endpoints in active treatment and sham arms estimated an R^2 of 0.78 for primary and 0.84 for secondary endpoints. Adverse events after sham were in most cases minor and of short duration.

Conclusion The generally small differences in effect size 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.

SUMMARY

Key messages

- The magnitude of change in the active treatment and placebo arms varied greatly, but about 80% of the variance in effect size of active treatment could be predicted by placebo effects, regression to the mean or spontaneous improvement.
- Adverse events related to sham procedures were mainly minor and short-lived, and frequently outweighed by positive placebo effects.

Strengths and limitations

- Selection of trials with low risk of bias
- Calculation of effect sizes on primary and pooled secondary endpoints in both active treatment and 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 percent 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. 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: (a) to assess the magnitude of change in outcome from baseline to trial endpoint in both the active treatment and placebo (sham) arms, (b) to explore the contribution of non-specific factors, including placebo, to the outcome of active treatment, and (c) 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 both 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 mentioned above 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, standard deviation, standard error or confidence intervals in both 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 endpoints. In trials without continuous primary endpoints, with multiple endpoints or no defined primary endpoint, we selected an outcome related to pain or condition-specific endpoint. 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 endpoint. We also registered secondary endpoints in order to avoid potential bias from selective reporting in the included trials. Endpoints describing medication, radiographic or physiological variables, social or psychological function, were not included. For the Parkinson-trials, only endpoints 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 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 standard deviations (SD). We calculated ES both for the active and sham intervention 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 standard deviations before and after treatment calculated ES. An ES of 0.8 or more is assumed large, while an ES of 0.5 - 0.8 is considered moderate.[13] In trials with multiple secondary endpoints we calculated the pooled mean ES, without weighting. Because of 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 outcome in the active and sham groups both for primary and pooled secondary endpoints. For this analysis, we used Medcalc Statistical Software version 12.7.4.0[15]

RESULTS

Selection of interventions and trials

The searches provided sham-controlled trials of 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: SD not provided (n=1),[16] compound primary endpoint (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, twelve of them 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.

Table 1. **Included 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 2005	Food and Drug Administration / NC	Percutaneous myocardial laser revascularization / intractable angina pectoris	Laser turned on but no procedure performed	MAE in hospital (high dose): 4.1%	MAE in hospital: 0
Salem 2004	Ethics committee / NC			No procedural AE	
Sihvonen 2013	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 2002	Review Board / NC	Arthroscopic debridement / Knee osteoarthritis	Simulated arthroscopy preparation, intravenous anaesthesia, skin incisions, no instruments entered knee, knee manipulated	No procedural AE	
Pham 2004	Review Board / NC	Hyaluronic acid / Knee osteoarthritis	Intraarticular injection of saline solution	No MAE	
				Any mAE: 81.7%	Any mAE: 1.2%
Altman 2004	Ethics committee / C			No MAE	
				mAE: 12.8%	mAE: 8%
Chevalier 2010	ClinicalTrials.org / C			No MAE	
				mAE: 35,8%	mAE: 33,8%
Kallmes 2009	Review Board / NC	Percutaneous vertebroplasty with PMMA cement injection / vertebral compression fracture	Conscious sedation + local anesthaesia, pressure put on spine, simulation of odor with mixing of PMMA to imitate the smell during the active procedure	No MAE	
				mAE: 14%	mAE: 16%
Buchbinder 2009	Ethics committee at each participating center / NC		Conscious sedation + local anesthaesia, needle inserted to rest on the lamina, PMMA container opened to imitate the smell during the active procedure	No procedural AE	
Cohen 2012	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 2005	Ethics committee / NC		2 mL saline into interspinous ligament	No MAE	
				mAE: 9%	mAE: 10%

Valat 2002	Ethics committee / NC		2 mL saline into epidural space, interlaminar approach	No MAE	
				mAE: 6%	mAE: 8%
Iversen 2011	Ethics committee / NC		Subcutaneous injection of 2 mL saline superficial to the sacral hiatus	Not reported	
Freeman 2005	Ethics committee / C	Intradiscal electrothermal therapy (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 2003	Review Board / NC		17-gauge needle introduced onto the outer annulus, mock electrode passage shown on monitor, generator noises produced	Not reported	
Kvarstein 2009	Ethics committee / NC	Percutaneous intradiscal radiofrequency thermocoagulation (PIRFT) / discogenic low back pain	17-gauge canula and RF-probe inserted into annulus, no RF current applied	Not reported	
Olanow 2003	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 2010	Review Board / C	Gene delivery of AAV2-Neurturin / Parkinson's disease	Scalp incisions, partial thickness burr holes, no intracranial injections	MAE: 4	MAE: 0
				Most frequent mAE: headache: 68%	Most frequent mAE: headache: 50%
Gross 2011	Review Board / C	Transplantation of human retinal pigmental cells / Parkinson's disease	Scalp incisions, partial thickness burr holes, no cell transplantation	1 death	0 deaths
				MAE: 23%	MAE: 0
LeWitt 2011	Review Board /	Insertion of AAV-	Insertion of catheter	No MAE	

	C	GAD gene into subthalamic nucleus / Parkinson's disease	caudal to nucleus, infusion of saline	mAE (probably related to procedure): 56%	mAE (probably related to procedure): 14%
Dowson 2008	Ethics committee / C	Patent foramen ovale closure with STARFlex Septal Repair Implant / migraine	General anesthesia, skin incision in the groin	MAE (possibly or probably related to procedure): 11%	MAE (possibly or probably related to procedure): 4%
C=commercial; NC=non-commerical; MAE=major adverse events; mAE=minor adverse events; PMMA=polymethylmethacrylate; AAV2 =adeno-associated; GAD=glutamic acid decarboxylase					

Fourteen trials from the systematic reviews fulfilled at least four of the five methodological criteria.[33 34 37-48] Seven trials provided through searches in MEDLINE fulfilled the same criteria.[35 36 49-53] The included and excluded secondary endpoints 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.[38 39 47 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 endpoints after active treatment, while 11 trials showed a similar ES after the sham procedure (figure 2, table 2).

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

Author / procedure	**Limit disease duration / time to follow-up (months)**	**Trial arm / no of patients randomised**	**ES primary endpoint**	**ES pooled secondary endpoints (no of endpoints)**
Leon 2005 / Percutaneous myocardial laser revascularization	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 revascularization	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		0.46	

			0.83	
		Sham / 40	0.71	0.51
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	>6 / 12		Brief Pain Inventory	(5)

thermocoagulation				
		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
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**
VAS=Visual Analogue Scale; NRS=Numerical Rating Scale; UPDRS=Unified Parkinson's Disease Rating Scale; Womac=Western Ontario and McMaster Universities Osteoarthritis Index				

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

On pooled secondary endpoints, 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 endpoint during treatment, while this was the case for two of the sham groups (not reported to be related to the procedure). On secondary endpoints, deterioration occurred in two active treatment and two sham groups (table 2).

Differences in outcome between active treatment and sham

Better results on primary endpoints were reported with active treatment compared to sham in 14 of the 21 trials, but the differences were small. Three trials (one epidural study[43], one discogenic pain study[46] and one Parkinson study[52]) reported a moderate effect but none showed a large effect (figure 3, table 2). Seven trials reported a better primary endpoint outcome after sham than after active treatment.

Nineteen trials reported secondary endpoints, 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, table 2). In twelve trials, the outcome was better for primary than for pooled secondary endpoints. This bore no relation to funding source.

On regression analyses, effect sizes in the sham groups predicted about 80 % of the variance of ES in the active treatment groups, both on primary and pooled secondary endpoints (figure 4 and 5).

Adverse events

Eighteen studies provided information about adverse events (AE) (table 1). Three of these trials reported no procedural adverse events in any of the groups.[33 35 41] Major AEs were reported after active treatment in four trials[34 50 51 53] including one death in one of the Parkinson studies.[51] In the sham groups, one trial[53] listed three major AEs possibly or probably related to the procedure, all thought to be caused by anti-platelet 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.

DISCUSSION

Principal findings

Analysis of 21 sham-controlled trials of minimally invasive procedures showed that the effect sizes in the active treatment arms were predicted by the effect sizes in the sham arms. There was a large ES on primary endpoints 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 endpoints.

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 endpoints varied from around zero 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 three, and in the epidural trials by a factor of two. 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 endpoints 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 study

It is our opinion that the calculation of effect sizes in both active treatment and placebo arms is a strength of the present study. This made it possible to assess the magnitude of change in both arms and the contribution of non-specific factors to change in the active treatment arms. The calculation of effect sizes provides an alternative assessment to probability estimates. Another strength of the study is the supplementary analyses of pooled secondary endpoints, enabling a more comprehensive evaluation than using primary endpoints alone. Reports of tactically motivated use of primary and secondary endpoints before publication in order to improve study results strengthen the argument for registering all relevant secondary endpoints.[54] Our finding that a majority of trials reported better results on primary than on secondary endpoints 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 the Cochrane Library and MEDLINE, but 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 realize 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] More 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 medico-technological 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 endpoints in the active treatment and sham 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 effect sizes in the range of 0.2 to 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 fourteen had a no-treatment arm. The mean ES in the placebo groups was about 0.5, while it was only slightly above zero 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 endpoints. 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 (i.e. waiting list), we cannot rule out that the changes appearing during the trial period also reflect non-specific factors, i.e. 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 trials, there were moderate to large improvements in the sham groups even at one-year follow-up.[49-51] 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.[64] 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 heterogeneity of the interventional techniques and variable outcome specifications, making standardization difficult to achieve. Existing ethical guidelines accept the role of placebo-controlled trials when certain conditions are met.[65] There must be genuine equipoise, i.e. 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.[66 67]

The selected trials gave a detailed description of adverse events in both active and sham-treated groups (table 1). The safety concerns frequently raised as an argument against the use of sham were generally not supported. Major adverse events related to the sham procedure were reported in only one of the trials[53] and they were short-lived and not life threatening. Minor adverse events were more frequent, but also of limited duration. Positive placebo-induced effects generally outweighed adverse events, 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 symptoms, research ethics committees should consider such factors in their risk-benefit assessments of planned sham controlled trials.[68 69]

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, more recently, robotic surgery.[70-75]

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 of 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 adverse events in sham-controlled trials are overrated and could be considered acceptable in view of the potential personal harm and societal costs associated with unproven minimally invasive interventions.

Figure legends

Figure 1. Flow chart of study selection in the present meta-analysis.

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

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

Figure 4. Association between effect sizes of primary endpoints in active treatment and sham arms. Linear regression, 95% confidence intervals. N=21.

Figure 5. Association between effect sizes of pooled secondary endpoints in active treatment and sham arms. Linear regression, 95% confidence intervals. N=19.

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-S15;

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**:915-9.

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;13:383-398.

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**:1411-6.

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-E304.

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**:710-9.

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**:1812-9.

35. 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.

36. 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.

37. 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**:1611-7.

38. 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.

39. 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**:113-9.

40. 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.

41. 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.

42. 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.

43. 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.

44. 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.

45. 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.

46. 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.

47. 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.

48. 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.

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 4th 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; 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. 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. doi: 10.1186/1745-6215-12-50.

65. Snyder L. Review of the American College of Physicians Ethics Manual, Sixth Edition. *Ann Intern Med* 2012;**156**(1_Part_2):73-104.

66. Flum DR. Interpreting surgical trials with subjective outcomes: avoiding UnSPORTsmanlike conduct. *JAMA* 2006;**296**:2483-5.

67. Heckerling PS. Placebo surgery research: a blinding imperative. *J Clin Epidemiol* 2006;59:876-80.

68. 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.

69. Redberg RF. Sham controls in medical device trials. *N Engl J Med* 2014;**371**:892-3.

70. 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.

71. 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.

72. 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.

73. 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**,Supplement:S142-7.

74. Vitale MA, Arons RR, Hurwitz S et al.The rising incidence of acromioplasty. *J Bone Joint Surg Am* 2010;**92**:1842-50.

75. 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.

Contributors: RH initiated and planned the project and searched databases. JIB and OT assisted in developing search strategies. Article screening and data extraction was carried out by RH. Quality of data extraction and checking was carried out by JIB and OT. Statistical analysis was undertaken by RH, who also wrote the draft. OT and JIB reviewed the draft and contributed to manuscript revisions. RH is the guarantor for this study.

Funding sources: None.

All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

Ethical approval: Ethical approval was not required for this work.

Data sharing: Dataset can be obtained from Robin Holtedahl (robi-hol@online.no).

The lead author affirms that this manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned (and, if relevant, registered) have been explained.

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

Robin Holtedahl, Jens Ivar Brox, Ole Tjomsland

Fram Rehabilitation Centre, Rykkinveien 100, 1349 Rykkin, Norway Robin Holtedahl Consultant Department of Physical Medicine and Rehabilitation, Oslo University Hospital, Box 4956 Nydalen, 0424 Oslo, Norway Jens Ivar Brox Professor South-Eastern Norway Regional Health Authority, PB 404, 2303 Hamar, Norway Ole Tjomsland Director of Quality and Specialist Areas

Correspondence to: Dr Robin Holtedahl; robi-hol@online, telephone +4790248973

Key words:
Placebo effects
Invasive procedures
Biomedical ethics
Evidence based health care

Word count: 5801

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 endpoints were calculated. Linear regression was used to analyse the association between endpoints 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 endpoints, there was a large clinical effect (ES ≥0.8) after active treatment in 12 trials and after sham procedures in 11 trials. For secondary endpoints, seven and five trials showed a large clinical effect, respectively. Three trials showed a moderate difference in ES between active treatment and sham on primary endpoints (ES ≥0.5) but no trials reported a large difference. No trials showed large or moderate differences in ES on pooled secondary endpoints. Regression analysis of endpoints in active treatment and sham arms estimated an R^2 of 0.78 for primary and 0.84 for secondary endpoints. Adverse events after sham were in most cases minor and of short duration.

Conclusion The generally small differences in effect size 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.

3

SUMMARY

Key messages

- The magnitude of change in the active treatment and placebo arms varied greatly, but about 80% of the variance in effect size of active treatment could be predicted by placebo effects, regression to the mean or spontaneous improvement.
- Adverse events related to sham procedures were mainly minor and short-lived, and frequently outweighed by positive placebo effects.

Strengths and limitations

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

1

4

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 percent 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. 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: (a) to assess the magnitude of change in outcome from baseline to trial endpoint in both the active treatment and placebo (sham) arms, (b) to explore the contribution of non-specific factors, including placebo, to the outcome of active treatment, and (c) to assess the level of reported adverse effects in both trial arms.

METHODS

Search strategy and selection criteria

We ~~first~~ conducted electronic searches for randomised placebo-controlled trials of minimally invasive interventions for cardiac, neurological and ~~selected~~ musculoskeletal conditions. ~~We primarily searched for interventions addressing subjective endpoints, including pain states, but included trials for Parkinson's disease. Open surgical and laparoscopic interventions and interventions targeting hard endpoints (i.e. hypertension) were excluded.~~ 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~~ing~~ 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". ~~Sett inn søkestrategi, søkeord osv.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 excluded open surgical and laparoscopic interventions. Where applicable, W~~We ~~w~~e 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~~From the reviews~~, we selected randomised placebo-controlled trials published ~~from~~ later than January 2000 ~~to January 2014~~. ~~Dette er ikke helt persist, fordi du har gjort søk på sham RCT på studier publisert etter siste inklusjonsdato I SR.~~ We excluded ~~earlier~~ 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

6

these criteria both 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 mentioned above 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, standard deviation, standard error or confidence intervals in both 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 endpoints. In trials without continuous primary endpoints, with multiple endpoints or no defined primary endpoint, we selected an outcome related to pain or condition-specific endpoint. 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 endpoint. We also registered secondary endpoints in order to avoid potential bias from selective reporting in the included trials. Endpoints describing medication, radiographic or physiological variables, social or psychological function, were not included. For the Parkinson-trials, only endpoints 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 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 standard deviations (SD). We calculated ES both for the active and sham intervention 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 standard deviations before and after treatment calculated ES. An ES of 0.8 or more is assumed large, while an ES of 0.5 - 0.8 is considered moderate.[13] In trials with multiple secondary endpoints we

7

calculated the pooled mean ES, without weighting. Because of 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 outcome in the active and sham groups both for primary and pooled secondary endpoints. For this analysis, we used Medcalc Statistical Software version 12.7.4.0[15]

RESULTS

Selection of interventions and trials

The searches provided sham-controlled trials of 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)), . Because of the large number of described interventions for neck- and back pain syndromes, we chose to restrict the analysis to sham-controlled trials of the following interventions: 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 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: SD not provided (n=1),[16](ref) compound primary endpoint (n=1),[17] (ref) risk of false positive response because of only one diagnostic block (n=4).[18-21] (

Study selection

The study selection process is summarised in web appendix figure 1. Web appendix table 2 shows the excluded trials and the reasons for exclusion. 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, twelve of them were not identified from the systematic

1

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.

Table 1. **Included 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 2005	Food and Drug Administration / NC	Percutaneous myocardial laser revascularization / intractable angina pectoris	Laser turned on but no procedure performed	MAE in hospital (high dose): 4.1%	MAE in hospital: 0
Salem 2004	Ethics committee / NC			No procedural AE	
Sihvonen 2013	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 2002	Review Board / NC	Arthroscopic debridement / Knee osteoarthritis	Simulated arthroscopy preparation, intravenous anaesthesia, skin incisions, no instruments entered knee, knee manipulated	No procedural AE	
Pham 2004	Review Board / NC	Hyaluronic acid / Knee osteoarthritis	Intraarticular injection of saline solution	No MAE Any mAE: 81.7%	Any mAE: 1.2%
Altman 2004	Ethics committee / C			No MAE mAE: 12.8%	mAE: 8%
Chevalier 2010	ClinicalTrials.org / C			No MAE mAE: 35,8%	mAE: 33,8%
Kallmes 2009	Review Board / NC	Percutaneous vertebroplasty with	Conscious sedation + local anesthaesia,	No MAE	

9

		PMMA cement injection / vertebral compression fracture	pressure put on spine, simulation of odor with mixing of PMMA to imitate the smell during the active procedure	mAE: 14%	mAE: 16%
Buchbinder 2009	Ethics committee at each participating center / NC		Conscious sedation + local anesthaesia, needle inserted to rest on the lamina, PMMA container opened to imitate the smell during the active procedure	No procedural AE	
Cohen 2012	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 2005	Ethics committee / NC		2 mL saline into interspinous ligament	No MAE	
				mAE: 9%	mAE: 10%
Valat 2002	Ethics committee / NC		2 mL saline into epidural space, interlaminar approach	No MAE	
				mAE: 6%	mAE: 8%
Iversen 2011	Ethics committee / NC		Subcutaneous injection of 2 mL saline superficial to the sacral hiatus	Not reported	
Freeman 2005	Ethics committee / C	Intradiscal electrothermal therapy (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 2003	Review Board / NC		17-gauge needle introduced onto the outer annulus, mock electrode passage shown on monitor, generator noises produced	Not reported	
Kvarstein 2009	Ethics committee / NC	Percutaneous intradiscal radiofrequency thermocoagulation (PIRFT) / discogenic	17-gauge canula and RF-probe inserted into annulus, no RF current applied	Not reported	

1

		low back pain			
Olanow 2003	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 2010	Review Board / C	Gene delivery of AAV2-Neurturin / Parkinson's disease	Scalp incisions, partial thickness burr holes, no intracranial injections	MAE: 4	MAE: 0
				Most frequent mAE: headache: 68%	Most frequent mAE: headache: 50%
Gross 2011	Review Board / C	Transplantation of human retinal pigmental cells / Parkinson's disease	Scalp incisions, partial thickness burr holes, no cell transplantation	1 death	0 deaths
				MAE: 23%	MAE: 0
LeWitt 2011	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 2008	Ethics committee / C	Patent foramen ovale closure with STARFlex Septal Repair Implant / migraine	General anesthesia, skin incision in the groin	MAE (possibly or probably related to procedure): 11%	MAE (possibly or probably related to procedure): 4%
C=commercial; NC=non-commerical; MAE=major adverse events; mAE=minor adverse events; PMMA=polymethylmethacrylate; AAV2 =adeno-associated; GAD=glutamic acid decarboxylase					

Fourteen trials from the systematic reviews fulfilled at least four of the five methodological criteria.[33 34 37-48] Seven trials provided through searches in MEDLINE fulfilled the same criteria.[35 36 49-53] The included and excluded secondary endpoints 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.[38 39 47 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 endpoints after active treatment, while 11 trials showed a similar ES after the sham procedure (figure ~~1~~2, table 2).

11

Table 2. **Effect size (ES) on primary and pooled secondary endpoints, showing differences between active treatment and sham arms.**				
Author / procedure	**Limit disease duration / time to follow-up (months)**	**Trial arm / no of patients randomised**	**ES primary endpoint**	**ES pooled secondary endpoints (no of endpoints)**
Leon 2005 / Percutaneous myocardial laser revascularization	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 revascularization	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

1

12

		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
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		1.10

13

			1.85	
		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
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	≥60 / 6		UPDRS 3 off	(7)

subthalamic nucleus				
		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**
VAS=Visual Analogue Scale; NRS=Numerical Rating Scale; UPDRS=Unified Parkinson's Disease Rating Scale; Womac=Western Ontario and McMaster Universities Osteoarthritis Index				

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

On pooled secondary endpoints, 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 endpoint during treatment, while this was the case for two of the sham groups (not reported to be related to the procedure). On secondary endpoints, deterioration occurred in two active treatment and two sham groups (table 2).

Differences in outcome between active treatment and sham

Better results on primary endpoints were reported with active treatment compared to sham in 14 of the 21 trials, but the differences were small. Three trials (one epidural study[43], one discogenic pain study[46] and one Parkinson study[52]) reported a moderate effect but none showed a large effect (figure ~~2~~3, table 2). Seven trials reported a better primary endpoint outcome after sham than after active treatment.

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

On regression analyses, effect sizes in the sham groups predicted about 80 % of the variance of ES in the active treatment groups, both on primary and pooled secondary endpoints (figure ~~3~~4 and ~~4~~5).

Adverse events

Eighteen studies provided information about adverse events (AE) (table 1). Three of these trials reported no procedural adverse events in any of the groups.[33,35,41] Major AEs were reported after active treatment in four trials[34,50,51,53] including one death in one of the Parkinson studies.[51] In the sham groups, one trial[53] listed three major AEs possibly or probably related to the procedure, all thought to be caused by anti-platelet 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.

DISCUSSION

Principal findings

Analysis of 21 sham-controlled trials of minimally invasive procedures showed that the effect sizes in the active treatment arms were predicted by the effect sizes in the sham arms. There was a large ES on primary endpoints 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 endpoints.

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 endpoints varied from around zero 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 three, and in the epidural trials by a factor of two. 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 endpoints 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 study

It is our opinion that the calculation of effect sizes in both active treatment and placebo arms is a strength of the present study. This made it possible to assess the magnitude of change in both arms and the contribution of non-specific factors to change in the active treatment arms. The calculation of effect sizes provides an alternative assessment to probability estimates. Another strength of the study is the supplementary analyses of pooled secondary endpoints, enabling a more comprehensive evaluation than using primary endpoints alone. Reports of tactically motivated use of primary and secondary endpoints before publication in order to improve study results strengthen the argument for registering all relevant secondary endpoints.[54] Our finding that a majority of trials reported better results on primary than on secondary endpoints 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 the Cochrane Library and MEDLINE, but 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 realize 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] More 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 medico-technological 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 endpoints in the active treatment and sham 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 effect sizes in the range of 0.2 to 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 fourteen had a no-treatment arm. The mean ES in the placebo groups was about 0.5, while it was only slightly above zero 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 endpoints. 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 (i.e. waiting list), we cannot rule out that the changes appearing during the trial period also reflect non-specific factors, i.e. 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 trials, there were moderate to large improvements in the sham groups even at one-year follow-up.[49-51] 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.[64] 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 heterogeneity of the interventional techniques and variable outcome specifications, making standardization difficult to achieve. Existing ethical guidelines accept the role of placebo-controlled trials when certain conditions are met.[65] There must be genuine equipoise, i.e. 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.[66 67]

The selected trials gave a detailed description of adverse events in both active and sham-treated groups (table 1). The safety concerns frequently raised as an argument against the use of sham were generally not supported. Major adverse events related to the sham procedure were reported in only one of the trials[53] and they were short-lived and not life threatening. Minor adverse events were more frequent, but

19

also of limited duration. Positive placebo-induced effects generally outweighed adverse events, 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 symptoms, research ethics committees should consider such factors in their risk-benefit assessments of planned sham controlled trials.[68 69]

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, more recently, robotic surgery.[70-75] 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 of 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 adverse events in sham-controlled trials are overrated and could be considered acceptable in view of the potential personal harm and societal costs associated with unproven minimally invasive interventions.

Figure legends
Figure 1. Flow chart of study selection in the present meta-analysis.

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

1

Figure ~~2~~3. Differences in effect size between active treatment and sham.

Figure ~~3~~4. Association between effect sizes of primary endpoints in active treatment and sham arms. Linear regression, 95% confidence intervals. N=21.

Figure ~~4~~5. Association between effect sizes of pooled secondary endpoints in active treatment and sham arms. Linear regression, 95% confidence intervals. N=19.

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.

21

7. Reeves B. Health-technology assessment in surgery. *Lancet* 1999;**353**:Suppl 1:S13-S15;

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**:915-9.

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.

1

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;13:383-398.

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**:1411-6.

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.

~~16.~~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.

~~17.~~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.

~~18.~~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.

~~19.~~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-E304.

~~20.~~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.

~~21.~~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.

24

~~22.~~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.

~~23.~~29. Karppinen J, Malmivaara A, Kurunlahti M et al. Periradicular infiltration for sciatica: a randomized controlled trial. *Spine* 2001; **26**:1059-67.

~~24.~~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**:710-9.

~~25.~~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.

~~26.~~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.

~~27.~~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.

~~28.~~34. Leon MB, Kornowski R, Downey WE et al. A blinded, randomized, placebo-controlled trial of

25

percutaneous laser myocardial revascularization to improve angina symptoms in patients with severe coronary disease. *J Am Coll Cardiol* 2005;**46**:1812-9.

~~29.~~35. 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.

~~30.~~36. 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.

~~31.~~37. 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**:1611-7.

~~32.~~38. 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.

~~33.~~39. 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**:113-9.

1

26

~~34.~~40. 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.

~~35.~~41. 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.

~~36.~~42. 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.

~~37.~~43. 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.

~~38.~~44. 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.

~~39.~~45. 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.

~~40.~~46. Pauza KJ, Howell S, Dreyfuss P et al. A randomized, placebo-controlled trial of intradiscal

27

electrothermal therapy for the treatment of discogenic low back pain. *Spine* J 2004;**4**:27-35.

~~41.~~47. 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.

~~42.~~48. 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.

~~43.~~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.

~~44.~~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.

~~45.~~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.

1

28

~~46.~~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.

~~47.~~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.

~~48.~~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.

~~49.~~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.

~~50.~~56. Bicket MC, Gupta A, Brown CH 4th 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.

29

~~51.~~57. Kaptchuk TJ, Goldman P, Stone DA et al. Do medical devices have enhanced placebo effects? *J Clin Epidemiol* 2000;**53**:786-92.

~~52.~~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.

~~53.~~59. Turner JA, Deyo RA, Loeser JD et al. The importance of placebo effects in pain treatment and research. *JAMA* 1994;**271**:1609-14.

~~54.~~60. Hróbjartsson A, Gøtzsche PC. Placebo interventions for all clinical conditions. *Cochrane Database Syst Rev* 2010; CD003974.

~~55.~~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.

~~56.~~62. Marchand S, Kupers RC, Bushnell MC et al. Analgesic and placebo effects of thalamic stimulation. *Pain* 2003;**105**:481-8.

~~57.~~63. Finniss DG, Kaptchuk TJ, Miller F et al. Biological, clinical, and ethical advances of placebo effects. *Lancet* 2010;**375**:686-95.

~~58.~~64. Campbell MK, Entwistle VA, Cuthbertson BH et al. Developing a placebo-controlled trial in

1

surgery: issues of design, acceptability and feasibility. *Trials* 2011;**12**:50. doi: 10.1186/1745-6215-12-50.

~~59.~~65. Snyder L. Review of the American College of Physicians Ethics Manual, Sixth Edition. *Ann Intern Med* 2012;**156**(1_Part_2):73-104.

~~60.~~66. Flum DR. Interpreting surgical trials with subjective outcomes: avoiding UnSPORTsmanlike conduct. *JAMA* 2006;**296**:2483-5.

~~61.~~67. Heckerling PS. Placebo surgery research: a blinding imperative. *J Clin Epidemiol* 2006;59:876-80.

~~62.~~68. 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.

~~63.~~69. Redberg RF. Sham controls in medical device trials. *N Engl J Med* 2014;**371**:892-3.

~~64.~~70. 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.

~~65.~~71. 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

31

randomised controlled trial. *Bone Joint Res* 2013;**2**:132-9.

~~66.~~72. 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.

~~67.~~73. 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**,Supplement:S142-7.

~~68.~~74. Vitale MA, Arons RR, Hurwitz S et al.The rising incidence of acromioplasty. *J Bone Joint Surg Am* 2010;**92**:1842-50.

~~69.~~75. 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.

Contributors: RH initiated and planned the project and searched databases. JIB and OT assisted in developing search strategies. Article screening and data extraction was carried out by RH. Quality of data extraction and checking was carried out by JIB and OT. Statistical analysis was undertaken by RH, who also wrote the draft. OT and JIB reviewed the draft and contributed to manuscript revisions. RH is the guarantor for this study.

Funding sources: None.

1

All authors have completed the ICMJE uniform disclosure form at www.icmje.org/coi_disclosure.pdf and declare: no support from any organisation for the submitted work; no financial relationships with any organisations that might have an interest in the submitted work in the previous three years; no other relationships or activities that could appear to have influenced the submitted work.

Ethical approval: Ethical approval was not required for this work.

Data sharing: Dataset can be obtained from Robin Holtedahl (robi-hol@online.no).

The lead author affirms that this manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned (and, if relevant, registered) have been explained.

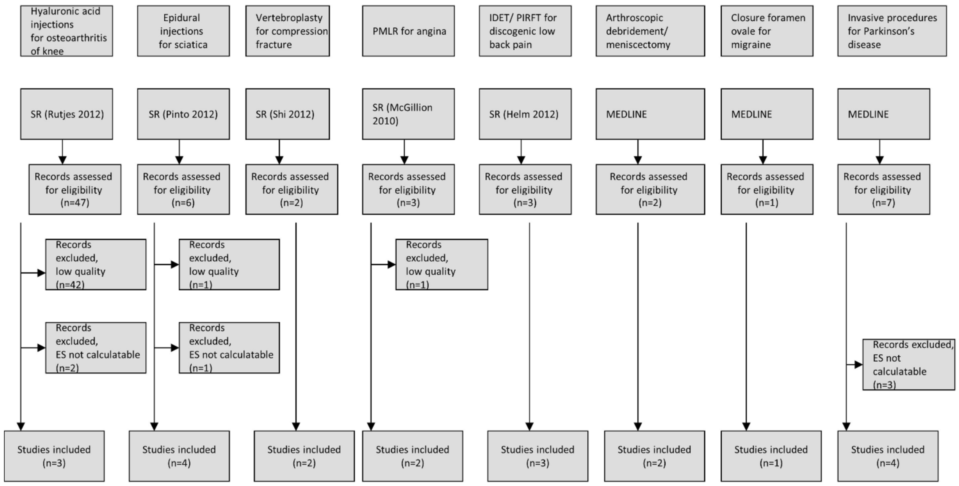


Figure 1 Flow chart of study selection in the present meta-analysis. SR = systematic review
254x190mm (300 x 300 DPI)

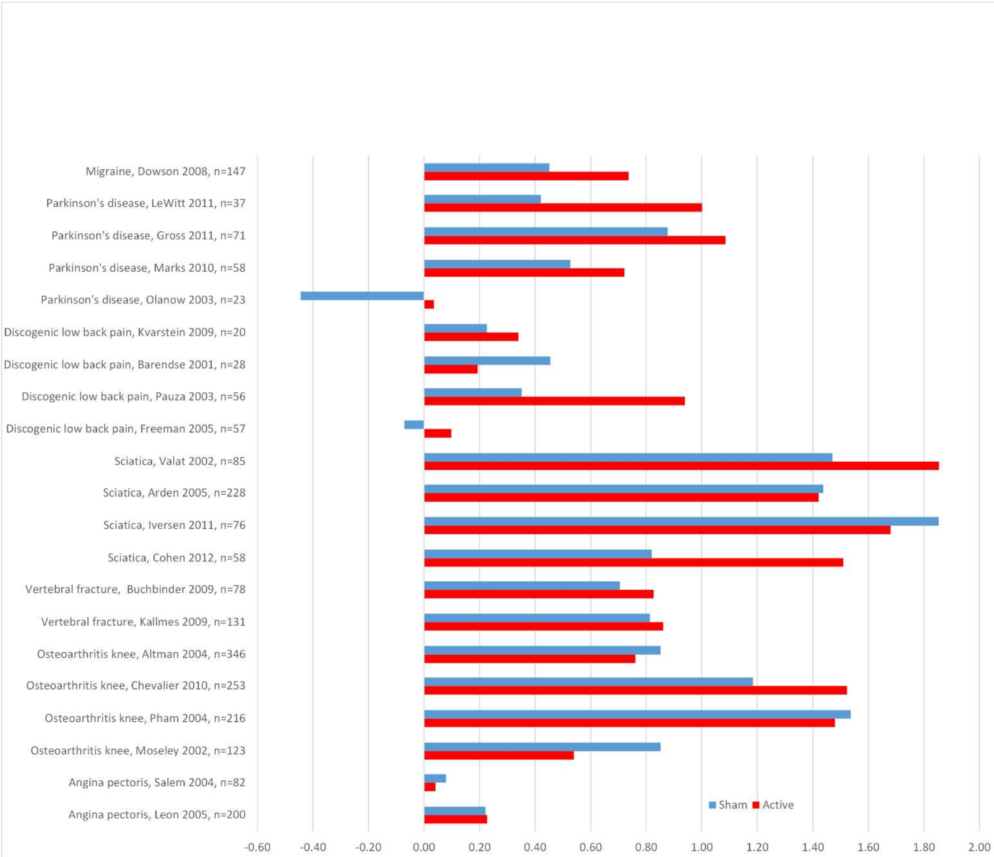


Figure 2 Effect sizes of active treatment and sham, primary endpoints.
250x216mm (300 x 300 DPI)

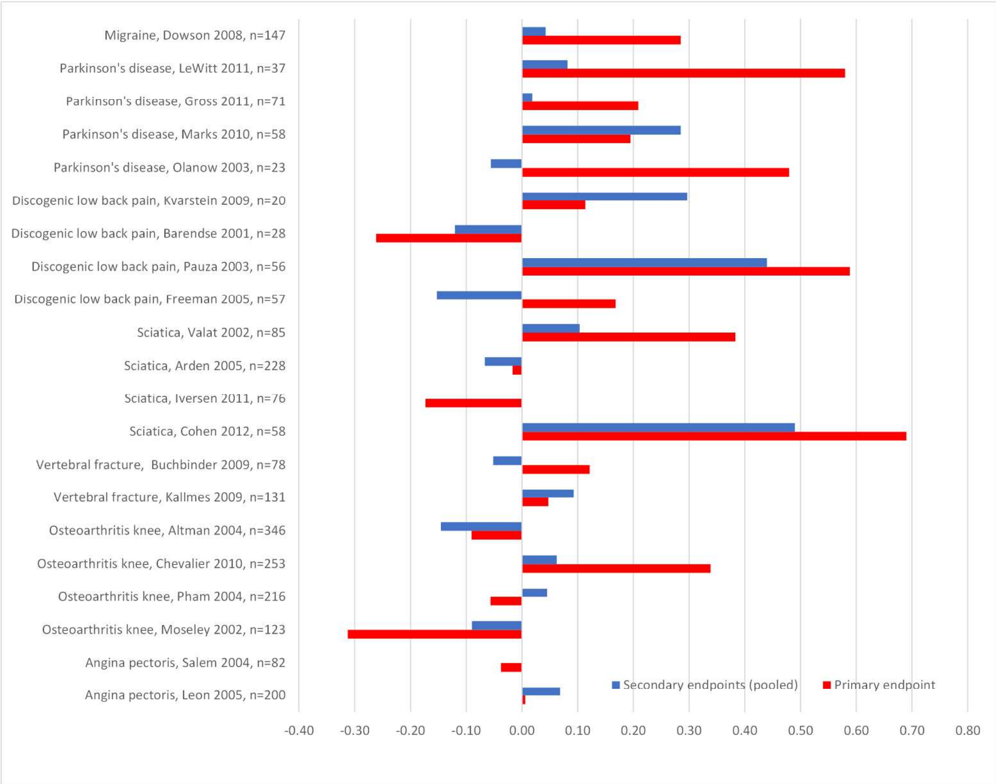


Figure 3 Differences in effect size between active treatment and sham.
221x173mm (300 x 300 DPI)

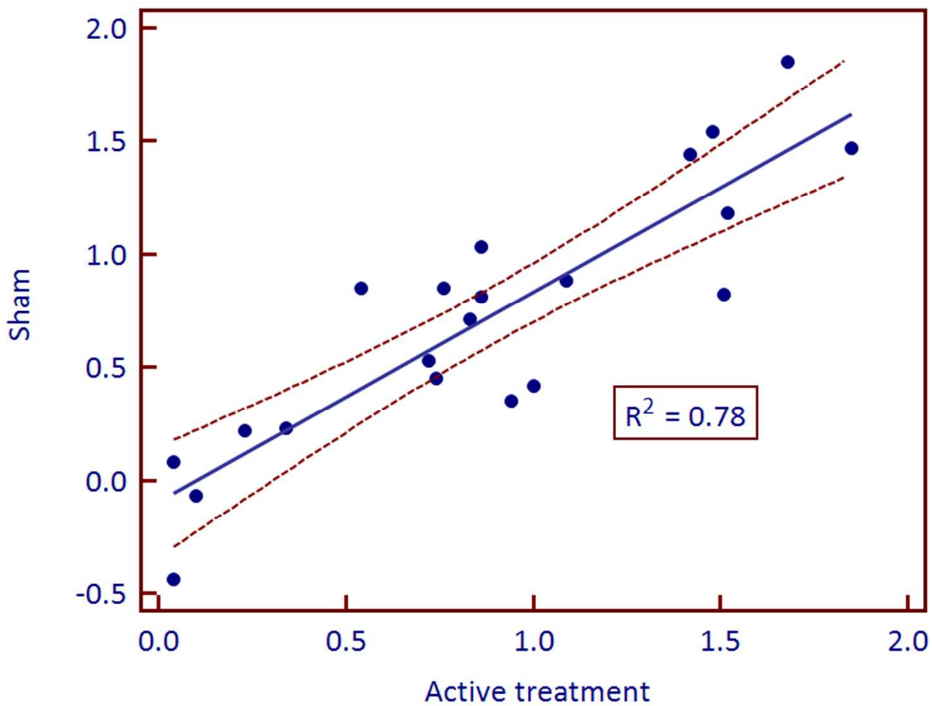


Figure 4 Association between effect sizes of primary endpoints in active treatment and sham arms. Linear regression, 95% confidence intervals. N=21.
67x50mm (300 x 300 DPI)

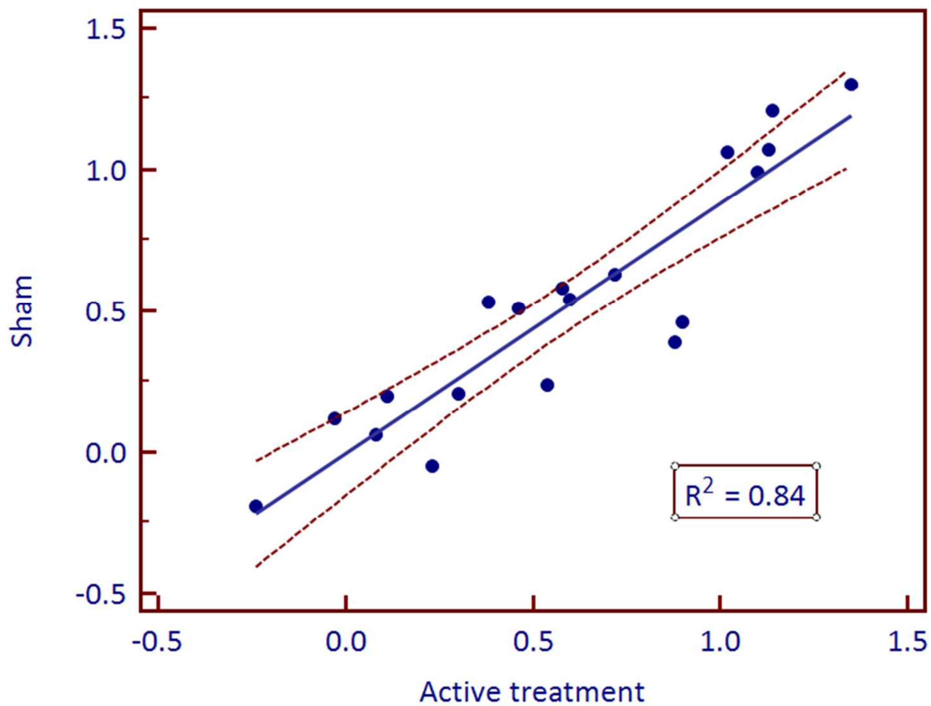


Figure 5. Association between effect sizes of pooled secondary endpoints in active treatment and sham arms. Linear regression, 95% confidence intervals. N=19.
67x50mm (300 x 300 DPI)

Appendix table 1. Indications, postulated mechanisms and history of selected interventions

Invasive procedure / indication	**Postulated mechanism**	**History**	**References**
Percutaneous myocardial laser revascularization / intractable angina pectoris	Increasing the delivery of oxygenated blood to poorly perfused myocardium by creating channels	Introduced in the 1980s, initially transmyocardial route, later percutaneous route, now mostly abandoned	Schofield PM, McNab D. NICE evaluation of transmyocardial laser revascularisation and percutaneous laser revascularisation for refractory angina. *Heart* 2010;96:312-3.
Patent foramen ovale closure with STARFlex Septal Repair Implant / migraine	Improvement of migraine headache, believed to block the formation of microemboles to the brain	Developed in the 1990s for the prevention of stroke, later thought to cure migraine, never in clinical use for this indication	Gornall J. A very public break-up. *BMJ* 2010;340:c110
Arthroscopic debridement / Knee osteoarthritis	Unclear, no documented effect on arthritic process, but about 50% report relief of pain (Mosely)	Annually about 650.000 procedures in the USA in the mid-ninetees, but 39% decrease between 2000 and 2008.	Holmes R, Moschetti W, Martin B, Tomek I, Finlayson S. Effect of evidence and changes in reimbursement on the rate of arthroscopy for osteoarthritis. *Am J Sports Med* 2013;41:1039-43.
Arthroscopic meniscectomy / degenerative meniscal lesions	Unclear, relief of symptoms attributed to trimming damaged meniscus down to viable meniscus and removing fragments.	The most common orthopedic procedure in the United States, 700.000 per year, up 50% last 15 years	Kim S, Bosque J, Meehan JP, Jamali A, Marder R. Increase in outpatient knee arthroscopy in the United States: a comparison of National Surveys of Ambulatory Surgery, 1996 and 2006. *J Bone Joint Surg Am* 2011;93:994-1000.
Viscosupplementation with hyaluronic acid / Knee osteoarthritis	Improve joint lubrication by increasing HA levels in joint, in spite of short half-lives (Marshall 2000)	Many positive reports since late 1980s, including sham-controlled trials. Still widely in use	Rutjes 2012 (15)
Percutaneous vertebroplasty with PMMA cement injection / vertebral compression fracture	Increase the strength of the damaged bone and alleviate pain by preventing microfractures	Numerous observational studies and single-blind trials reported substantial clinical benefits. Slight reduction of procedure since 2009	Manchikanti L, Pampati V, Hirsch JA. Analysis of utilization patterns of vertebroplasty and kyphoplasty in the Medicare population. *J Neurointerv Surg* 2013;5:467-72.
Epidural injection of corticosteroids / Sciatica	Dampen inflammatory reaction in nerve root sheaths caused by mechanical compression	Routinely used for sciatica since the 1950s (Pinto 2012). Since 2000 the number of injections increased by about 130% in the United States and 50% in the United Kingdom	Manchikanti L, Falco FJ, Singh V, Pampati V, Parr AT, Benyamin RM, Fellows B, Hirsch JA. Utilization of interventional techniques in managing chronic pain in the Medicare population: analysis of growth patterns from 2000 to 2011. *Pain Physician* 2012;15:E969-82

Percutaneous intradiscal radiofrequency and thermocoagulation (PIRFT and IDET) / discogenic low back pain	Placement of a electrode or RF-probe into the annulus and applying heat or current to destruct nociceptors/ annulus	Introduced in 1996 (IDET), later mostly abandoned	Helm 2012 (18)
Fetal nigral transplantation / Parkinson's disease	Restoration of dopamin levels in basal ganglia through injection of growth factors, GAD gene or nigral dopamine neurons	Based on animal models and a few small observational trials from about 2000. None in routine clinical use due to insufficient evidence	-
Gene delivery of AAV2-Neurturin / Parkinson's disease			
Transplantation of human retinal pigmental cells / Parkinson's disease			
Insertion of AAV-GAD gene into subthalamic nucleus / Parkinson's disease			

Appendix table 2. Search phrases, no of eligible studies and reasons for exclusion						
Procedure	**Search phrase MEDLINE**	**Source**	**Eligible studies**	**Excluded, ES not calculatable**	**Excluded, risk of bias**	**Included studies**
PMLR	Percutaneous myocardial laser revascularization	McGillion 2010 (17)	3	-	1	Salem 2004, Leon 2005
PIRFT /IDET	Intradiscal OR annular AND thermal AND "low back pain"	Helm 2012 (18)	3	-	-	Kvarstein, 2009
				-	-	Freeman 2005, Pauza 2003
Epidural injection corticosteroids	Epidural AND corticosteroid* AND sciatica	Pinto 2012 (16)	6	Karppinen 2001	1	Iversen 2011, Valat 2002, Arden 2005, Cohen 2012
Intraarticular hyaluronic acid for osteoarthritis knee	Hyaluron* OR viscosuppl* AND knee AND osteoarthritis	Rutjes 2012 (15)	48	Lundsgaard 2008, Petrella 2008	41	Petrella 2006, Chevalier 2010, Altman 2004, Pham 2004
Vertebroplasty	vertebroplast*	Shi 2012 (19)	2	-	-	Kallmes 2009, Buchbinder 2009
Invasive treatment of Parkinson's disease	transplantation OR gene OR "stem cell" AND Parkinson*	MEDLINE	6	Freed 2001, Gordon 2004, McRae 2004	-	Marks 2010, Olanow 2003, Gross 2011, LeWitt 2011
Arthroscopic debridement knee osteoarthritis	debridement AND lavage AND knee AND osteoarthr*	MEDLINE	1	-	-	Moseley 2002
Meniscectomy knee	meniscectomy AND knee	MEDLINE	1	-	-	Sihvonen 2013
Foramen ovale closure for migraine	"foramen ovale" AND migraine	MEDLINE	1	-	-	Dowson 2008
Number of trials			**71**	**6**	**43**	**22**

Appendix table 3. Included and excluded secondary endpoints.		
Author	**Included secondary endpoints**	**Excluded secondary endpoints (means not reported, or irrelevant)**
Leon 2005		
	Time to onset angina	Improvement in angina class
	Time to onset ST depression	Radioisotope imaging
	Overall health	
	Frequency angina	
	Stability angina	
	Physical functioning	
	Disease perception	
	Treatment satisfaction	
	PCS	
	MCS	
Salem 2004		
		Proportion improved CCS angina class
		Medication usage
		Seattle Angina Questionnaire
		Left EF
		Angina stability
		Angina frequency
		Physical limitation
		Treatment satisfactioin
		Disease perception
Sihvonen 2013	WOMET score	-
	Knee pain at rest	
	Knee pain after exercise	
	15D score	
Moseley 2002		
	Arthritis Impact Scale	-
	Physical functioning Scale	
	Walking-bending	
	SF-36 Pain	
	SF-36 Physical functioning	
Pham 2004		
	Lequesne's algofunctional index	-
	Global assessment	
	% painful days	
Chevalier 2010		
	Womac C function	-
Altman 2004		
	Womac stiffness	-
	Womac physical	
Kallmes 2009		
	Pain Intensity	Opioid use

	SF-36 PCS	
	SF-36 MCS	
	Pain Frequency Index	
	Pain Bothersomeness Index	
	EQ-SD Index	
	SOF-ADL	
Buchbinder 2009		
	Roland-Morris Disability Questionnaire	-
	Life Questionnaire of the European Foundation	
	European Quality of Life–5 Dimensions	
Cohen 2012		
	Oswestry Disability Index	-
	Back pain	
Arden 2005		
	Leg pain	Analgesic use
	Back pain	
Valat 2002		
	Roland-Morris Disability Questionnaire	Dallas Pain Questionnaire
	Straight leg raising	
	Schober's test	
Iversen 2011		
		VAS back and leg pain, European Quality of Life scale
Freeman 2005		
	Modifiede Somatic Perception Questionnaire	SF-36 Mental, Role Physical/ Mental, Social Function
	Low Back Pain Outcome Score	
	SF-36 Physical Function	
	SF-36 Pain	
	SF-36 General Health	
	SF-36 Vitality	
Pauza 2003		
	VAS Pain	-
	SF-36 Physical Function	
	SF-36 Pain	
Kvarstein 2009		
	SF-36 Bodily pain	SF-36 Mental, Role Physical/ Mental, Social Function
	SF-36 Physical function	
	Oswestry Disability Index	
	SF-36 General health	
	SF-36 Vitality	
Olanow 2003		
	UPDRS motor on	Mean L-dopa dose equivalents
	UPDRS ADL off	
	UPDRS ADL on	

	% Off time day	
	% On time without dyskinesia	
Marks 2010		
	UPDRS OFF 1	Mean L-dopa dose equivalents
	UPDRS OFF 2	
	UPDRS ON 1	
	UPDRS ON 2	
	UPDRS ON 3	
	On without dyskinesia	
	On with dyskinesia	
Gross 2011		
	UPDRS ON	Mean L-dopa dose equivalents
	UPDRS ADL	
LeWitt 2011		
	UPDRS 1	Timed walking
	UPDRS2	BPRS other than taps
	UPDRS4	Dyskinesia rating scale
	Schwab and England ADL scale	Patient's diary
	BPRS taps 60 s	Clinical global impression
	Hoehan and Yahr stage	
	PDQ-39 total	
Dowson 2008		
	Headache Impact Test	-

PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
TITLE			
Title	1	Identify the report as a systematic review, meta-analysis, or both.	1
ABSTRACT			
Structured summary	2	Provide a structured summary including, as applicable: background; objectives; data sources; study eligibility criteria, participants, and interventions; study appraisal and synthesis methods; results; limitations; conclusions and implications of key findings; systematic review registration number.	2
INTRODUCTION			
Rationale	3	Describe the rationale for the review in the context of what is already known.	4
Objectives	4	Provide an explicit statement of questions being addressed with reference to participants, interventions, comparisons, outcomes, and study design (PICOS).	4
METHODS			
Protocol and registration	5	Indicate if a review protocol exists, if and where it can be accessed (e.g., Web address), and, if available, provide registration information including registration number.	
Eligibility criteria	6	Specify study characteristics (e.g., PICOS, length of follow-up) and report characteristics (e.g., years considered, language, publication status) used as criteria for eligibility, giving rationale.	5
Information sources	7	Describe all information sources (e.g., databases with dates of coverage, contact with study authors to identify additional studies) in the search and date last searched.	5
Search	8	Present full electronic search strategy for at least one database, including any limits used, such that it could be repeated.	5, Appendix table 1
Study selection	9	State the process for selecting studies (i.e., screening, eligibility, included in systematic review, and, if applicable, included in the meta-analysis).	5-6, appendix table 2
Data collection process	10	Describe method of data extraction from reports (e.g., piloted forms, independently, in duplicate) and any processes for obtaining and confirming data from investigators.	6
Data items	11	List and define all variables for which data were sought (e.g., PICOS, funding sources) and any assumptions and simplifications made.	6
Risk of bias in individual studies	12	Describe methods used for assessing risk of bias of individual studies (including specification of whether this was done at the study or outcome level), and how this information is to be used in any data synthesis.	5-6
Summary measures	13	State the principal summary measures (e.g., risk ratio, difference in means).	6
Synthesis of results	14	Describe the methods of handling data and combining results of studies, if done, including measures of consistency (e.g., I^2) for each meta-analysis.	6

PRISMA 2009 Checklist

Section/topic	#	Checklist item	Reported on page #
Risk of bias across studies	15	Specify any assessment of risk of bias that may affect the cumulative evidence (e.g., publication bias, selective reporting within studies).	5-6
Additional analyses	16	Describe methods of additional analyses (e.g., sensitivity or subgroup analyses, meta-regression), if done, indicating which were pre-specified.	
RESULTS			
Study selection	17	Give numbers of studies screened, assessed for eligibility, and included in the review, with reasons for exclusions at each stage, ideally with a flow diagram.	7, fig. 1
Study characteristics	18	For each study, present characteristics for which data were extracted (e.g., study size, PICOS, follow-up period) and provide the citations.	7-13
Risk of bias within studies	19	Present data on risk of bias of each study and, if available, any outcome level assessment (see item 12).	7,9
Results of individual studies	20	For all outcomes considered (benefits or harms), present, for each study: (a) simple summary data for each intervention group (b) effect estimates and confidence intervals, ideally with a forest plot.	10-13, fig 2,3
Synthesis of results	21	Present results of each meta-analysis done, including confidence intervals and measures of consistency.	10-13
Risk of bias across studies	22	Present results of any assessment of risk of bias across studies (see Item 15).	7,10
Additional analysis	23	Give results of additional analyses, if done (e.g., sensitivity or subgroup analyses, meta-regression [see Item 16]).	14, Fig. 4, 5
DISCUSSION			
Summary of evidence	24	Summarize the main findings including the strength of evidence for each main outcome; consider their relevance to key groups (e.g., healthcare providers, users, and policy makers).	14,15
Limitations	25	Discuss limitations at study and outcome level (e.g., risk of bias), and at review-level (e.g., incomplete retrieval of identified research, reporting bias).	15-17
Conclusions	26	Provide a general interpretation of the results in the context of other evidence, and implications for future research.	19
FUNDING			
Funding	27	Describe sources of funding for the systematic review and other support (e.g., supply of data); role of funders for the systematic review.	31

From: Moher D, Liberati A, Tetzlaff J, Altman DG, The PRISMA Group (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. PLoS Med 6(6): e1000097. doi:10.1371/journal.pmed1000097

For more information, visit: www.prisma-statement.org.