

## Is the placebo powerless? Update of a systematic review with 52 new randomized trials comparing placebo with no treatment

A. HRÓBJARTSSON & P. C. GØTZSCHE

From The Nordic Cochrane Centre, Rigshospitalet, Copenhagen, Denmark

**Abstract.** Hróbjartsson A, Gøtzsche PC (The Nordic Cochrane Centre, Copenhagen Ø, Denmark). Is the placebo powerless? Update of a systematic review with 52 new randomized trials comparing placebo with no treatment (Review). *J Intern Med* 2004; **256**: 91–100.

**Background.** It is widely believed that placebo interventions induce powerful effects. We could not confirm this in a systematic review of 114 randomized trials that compared placebo-treated with untreated patients.

**Aim.** To study whether a new sample of trials would reproduce our earlier findings, and to update the review.

**Methods.** Systematic review of trials that were published since our last search (or not previously identified), and of all available trials.

**Results.** Data was available in 42 out of 52 new trials (3212 patients). The results were similar to our previous findings. The updated review summarizes data from 156 trials (11 737 patients). We found no statistically significant pooled effect in 38 trials with

binary outcomes, relative risk 0.95 (95% confidence interval 0.89–1.01). The effect on continuous outcomes decreased with increasing sample size, and there was considerable variation in effect also between large trials; the effect estimates should therefore be interpreted cautiously. If this bias is disregarded, the pooled standardized mean difference in 118 trials with continuous outcomes was  $-0.24$  ( $-0.31$  to  $-0.17$ ). For trials with patient-reported outcomes the effect was  $-0.30$  ( $-0.38$  to  $-0.21$ ), but only  $-0.10$  ( $-0.20$  to  $0.01$ ) for trials with observer-reported outcomes. Of 10 clinical conditions investigated in three trials or more, placebo had a statistically significant pooled effect only on pain or phobia on continuous scales.

**Conclusion.** We found no evidence of a generally large effect of placebo interventions. A possible small effect on patient-reported continuous outcomes, especially pain, could not be clearly distinguished from bias.

**Keywords:** meta-analysis, pain, placebo effect, placebos, systematic review.

### Background

Within a few years in the 1950s it became a common conception that effects of placebo interventions were large, and that numerous randomized trials had reliably documented these effects in a wide

range of clinical conditions. To a considerable extent this prevailing opinion was caused by a paper by Beecher 'The Powerful Placebo' [1]. However, in 1997, Kienle and Kiene showed that Beecher's influential paper was flawed [2]. Beecher, and the vast majority of placebo investigators, had not

compared patients randomized to a placebo-treated group and to an untreated group. Instead the effect had been estimated as the uncontrolled before–after difference in a placebo group in a randomized trial, which fails to distinguish the effect of placebo from spontaneous remission, and other factors [3].

Randomized trials comparing placebo and no-treatment groups were considered to be very rare [4], but despite the lack of reliable evidence many commentators continued to believe in dramatic effects of placebos [5]. This opinion was challenged in 2001 when we published a systematic review of 114 randomized clinical trials that had compared placebo-treated with untreated patients [6]. We found no evidence that placebo interventions in general had clinically important effects, and a possible effect on patient-reported continuous outcomes, for example pain, could not be clearly distinguished from bias. This result surprised us and others, caused considerable public attention, and was described as ‘a challenge to core beliefs’ [7].

In the light of the unexpected findings, it is important to explore whether the results can be reproduced in subsequently published trials. The inclusion of new trials would also provide more data for assessments of the effect of placebos on specific clinical conditions. We have therefore updated our review with trials identified since our literature search in 1999.

## Methods

Our aims were to: (i) review new trials with placebo and no-treatment groups and to compare their results with those from the first version of the review, and (ii) update the review by including the new trials and reconduct all analyses.

We wished to study whether there was any tendency of a general effect of placebo interventions (across various health conditions), to investigate the effects on specific health conditions, and to assess whether the effect differed for patient-reported outcomes (for example pain) and observer-reported outcomes (for example hypertension).

We searched for trials published from 1999 to 2002, and also included trials published before 1999 that had not been identified previously. We repeated the methods of the earlier versions of the review [6, 8], except for the use of the  $I^2$ -test in the updated version.

We pragmatically defined a placebo intervention as any intervention, which was clearly labelled a placebo

in a trial report. Early in 2003 we searched The Cochrane Library, Medline, Embase, PsychInfo and Biological Abstracts for randomized clinical trials with a placebo group and a no-treatment group. Trials were excluded if it was clear that: (i) allocation of patients was conducted without concealment, for example by day of month; or (ii) the person who assessed observer-reported outcomes was aware of group assignments; or (iii) the dropout rate exceeded 50%.

One outcome per trial was extracted for the main analyses, preferably the primary outcome of the trial report. For binary outcomes we calculated the relative risk (if below 1, it indicates a positive effect of the placebo intervention). For continuous outcomes we calculated the standardized mean difference (a negative value indicates a positive effect of the placebo intervention). We calculated the pooled relative risks, and the pooled standardized mean differences, with random effect models as suggested by DerSimonian and Laird [9]. A random effects model results in a pooled effect which basically is a weighted mean of the effect found in each individual trial, besides also incorporating the variation in effect between the trials.

The different results reported in various trials can be a result of random variation or true differences in effect, so-called heterogeneity. We examined heterogeneity by calculating the DerSimonian and Laird's  $Q$  statistic [9], and the  $I^2$ -statistic [10]. Both were compared with a chi-square distribution with degrees of freedom equal to the number of trials minus one. We used the  $Q$  statistic for testing the presence of heterogeneity, and the  $I^2$ -statistic for estimating the degree of heterogeneity. The  $I^2$ -statistic can be interpreted as the proportion of the observed discrepancy in the estimation of effect, within a group of trials, which cannot be accounted for by random variation [10]. All results are reported with 95% confidence intervals and all  $P$ -values are two-tailed.

We calculated the pooled effects of placebo overall for trials with binary outcomes and for trials with continuous outcomes. We also calculated the pooled effect on separate clinical conditions when they had been studied in three trials or more, and the pooled effect of trials with patient-reported and observer-reported outcomes. For each trial we plotted the effect by the inverse of its standard error. Asymmetry in such ‘funnel plots’ reflects that the effects of individual studies decrease with increasing sample size. The degree of funnel plot asymmetry

was assessed both visually, and formally by a linear regression analysis [11].

We conducted 10 preplanned comparisons of the results obtained in two or more subgroups of trials to explore whether the effect of placebo was related to type of intervention (physical, psychological or pharmacological), outcome (various subgroups of patient-reported and observer-reported outcomes), or aspects of methodological quality (concealment of allocation, blinding of treatment provider, blinding of outcome evaluator, dropout rates 15% or lower, lack of co-intervention, clearly stated primary outcome, clearly stated aim of studying effects of placebo, and non-Gaussian distributions). We conducted two unplanned comparisons. In one we explored the effect in trials with sample sizes of 50 patients or more. In the second we explored the effect in the trials with both clearly adequate concealment, dropout rate of 15% or lower, and a sample size of 50 patients or more.

## Results

### *New trials identified in 2003*

Of 250 potentially eligible trial reports, we excluded 128 that addressed nonclinical or nonrandomized studies, 53 that did not compare a placebo group with a no-treatment group, six duplicate publications, and

a further 11 trials for other reasons, for example dropout rates over 50%. Of the remaining 52 trials, 40 had been published after 1998. We were unable to extract relevant outcome data in nine trials, and one trial investigated adverse effects. The analyses were therefore based on 42 trials with 3212 patients (counting only patients in the placebo and no-treatment groups). There were six trials with binary outcomes (489 patients) and 36 trials with continuous outcomes (2723 patients).

A description of the individual trials and their results can be found in the Appendix or in the forthcoming Cochrane version of this review [12]. The trials investigated 14 clinical conditions: depression, insomnia, pain, nausea, phobia, smoking, vitiligo, hypertension, obesity, jet lag, secondary erectile dysfunction, dry eye, patient involvement in adolescent diabetic care and difficulty of colonoscopy.

*Trials with binary outcomes.* As we identified only six trials with binary outcomes, a funnel plot would not be meaningful. Heterogeneity was not detected ( $P = 0.79$ ,  $I^2 = 0\%$ ). There was no statistically significant pooled effect of placebo overall, relative risk 0.95 (95% CI 0.83–1.08), or for the trials with patient-reported or observer-reported outcomes (Table 1). Only pain had been investigated by at least three trials; an effect was not detected (Table 2).

**Table 1** Effect of placebo treatment

	<i>n</i>	<i>k</i>	Pooled values	$I^2$ (%)
<i>Trials identified in 2003</i>				
Binary outcomes <sup>a</sup>				
Overall	489	6	0.95 (0.83–1.08)	0
Patient-reported	449	4	0.96 (0.83–1.11)	0
Observer-reported	40	2	0.88 (0.61–1.25)	nc
Continuous outcomes <sup>b</sup>				
Overall	2723	36	−0.17 (−0.26 to −0.07)	38
Patient-reported	2118	22	−0.19 (−0.32 to −0.07)	38
Observer-reported	605	14	−0.09 (−0.25 to 0.07)	0
<i>All trials</i>				
Binary outcomes <sup>a</sup>				
Overall	4284	38	0.95 (0.89–1.01)	38
Patient-reported	2377	27	0.95 (0.88–1.03)	35
Observer-reported	1907	11	0.91 (0.81–1.03)	45
Continuous outcomes <sup>b</sup>				
Overall	7453	118	−0.24 (−0.31 to −0.17)	45
Patient-reported	5199	75	−0.30 (−0.38 to −0.21)	45
Observer-reported	2254	43	−0.10 (−0.20 to 0.01)	23

*n*, number of trial participants; *k*, number of trials; nc, not calculable. <sup>a</sup>Pooled relative risk (a figure below 1 indicates a beneficial effect, 95% confidence intervals in brackets). <sup>b</sup>Pooled standardized mean difference (a negative figure indicates a beneficial effect).  $I^2$  the percentage of the variation of the estimated effect between the trials that cannot be accounted for by random variation.

**Table 2** Effect of placebo treatment on clinical problems investigated in 3 trials or more

	<i>n</i>	<i>k</i>	Pooled values	<i>I</i> <sup>2</sup> (%)
<i>Trials identified in 2003</i>				
Binary outcomes <sup>a</sup>				
Pain	154	3	0.99 (0.81–1.21)	9
Continuous outcomes <sup>b</sup>				
Pain	1231	17	-0.23 (-0.38 to -0.09)	29
Hypertension	179	3	0.02 (-0.29 to 0.32)	0
Nausea	161	3	-0.22 (-0.53 to 0.09)	0
Obesity	60	3	0.09 (-0.42 to 0.60)	0
<i>All trials</i>				
Binary outcomes <sup>a</sup>				
Smoking	887	6	0.88 (0.71–1.09)	79
Pain	525	5	0.98 (0.88–1.10)	0
Nausea	497	5	0.92 (0.80–1.06)	0
Depression	152	3	1.03 (0.78–1.34)	0
Continuous outcomes <sup>b</sup>				
Pain	2,833	44	-0.25 (-0.35 to -0.16)	26
Smoking	703	3	-0.53 (-1.29 to 0.23)	75
Hypertension	308	10	-0.17 (-0.46 to 0.12)	29
Nausea	288	5	-0.31 (-0.63 to 0.01)	41
Anxiety	257	6	-0.06 (-0.31 to 0.18)	0
Obesity	188	8	-0.20 (-0.57 to 0.17)	32
Insomnia	164	6	-0.19 (-0.50 to 0.12)	0
Depression	106	4	-0.27 (-0.69 to 0.15)	11
Asthma	81	3	-0.34 (-0.83 to 0.14)	53
Phobia	57	3	-0.63 (-1.17 to -0.08)	0

*n*, number of trial participants; *k*, number of trials. <sup>a</sup>Pooled relative risk (a figure below 1 indicates a beneficial effect, 95% CI in brackets). <sup>b</sup>Pooled standardized mean difference (a negative figure indicates a beneficial effect). *I*<sup>2</sup>: the percentage of the variation of the estimated effect between the trials that cannot be accounted for by random variation.

*Trials with continuous outcomes.* On inspection the funnel plot of the 36 trials was asymmetrical without a clear single peak (not shown). The effect of placebo decreased with increasing sample size, but this tendency was not statistically significant ( $P = 0.16$ ). There was no statistically significant heterogeneity ( $I^2 = 20\%$ ,  $P = 0.14$ ). There was a small effect of placebo overall, standardized mean difference  $-0.17$  ( $-0.26$  to  $-0.07$ ) (Table 1). A similar effect was found for trials with patient-reported outcomes, but no statistically significant pooled effect was seen for trials with observer-reported outcomes, standardized mean difference  $-0.09$  ( $-0.25$  to  $0.07$ ) (Table 1). Four clinical conditions had been investigated in three trials or more. We found a statistically significant pooled effect of placebo on pain, but not on hypertension, nausea or obesity (Table 2).

### All trials

As there were no statistically significant differences between the pooled results of the previously analysed 114 trials and the newly included 42 trials (data not shown), we present the combined results in the following.

We included 182 trials; data could be extracted from 156 trials (11 737 patients, which is 38% more than in our previous review). There were 38 trials (4284 patients) with binary outcomes and 118 trials (7453 patients) with continuous outcomes.

A description of the individual trials can be found in the forthcoming updated Cochrane version of this review [12]. The trials investigated 46 clinical conditions: depression, insomnia, pain, nausea, phobia, smoking, vitiligo, hypertension, obesity, jet lag, secondary erectile dysfunction, dry eye, patient involvement in adolescent diabetic care, difficulty of colonoscopy, alcohol abuse, Alzheimer's disease, anaemia, anxiety, asthma, attention-deficit hyperactivity disorder, bacterial infections, benign prostatic hyperplasia, carpal tunnel syndrome, common cold, compulsive nail biting, enuresis, epilepsy, faecal soiling, herpes simplex infection, hypercholesterolaemia, hyperglycaemia, ileus, infertility, insufficient cervical dilatation, labour, marital discord, menopause, mental handicap, orgasmic difficulties, Parkinson's disease, poor oral hygiene, Raynaud's disease, schizophrenia, sea sickness, stress related to dental treatment and undiagnosed ailments.

*Trials with binary outcomes.* The funnel plot was symmetrical (Fig. 1) and the effect of placebo did not change with increasing sample size ( $P = 0.41$ ). There was small to moderate heterogeneity ( $P = 0.01$ ,  $I^2 = 38\%$ ). There was no statistically significant pooled effect of placebo overall, relative risk 0.95 (0.88–1.01), or on patient-reported or observer-reported outcomes (Table 1). Placebo had no statistically significant effect either on conditions that had been investigated in at least three independent trials (nausea, pain, relapse in prevention of smoking and depression), but confidence intervals were wide (Table 2). There was substantial heterogeneity between the trials investigating smoking ( $P < 0.001$ ,  $I^2 = 79\%$ ) (Table 2).

*Trials with continuous outcomes.* The funnel plot was asymmetrical, and a single peak could not be

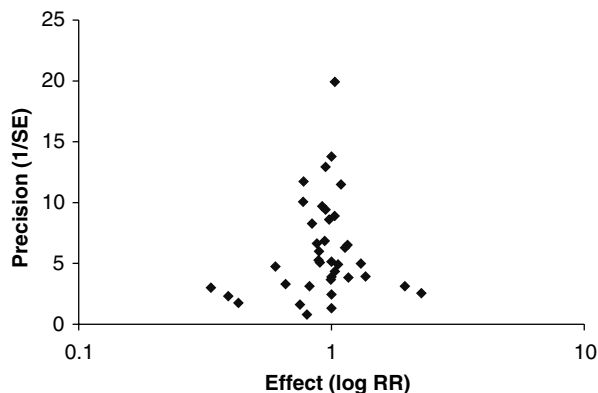


Fig. 1 Funnel plot of 38 trials with binary outcomes. RR, relative risk; SE, standard error.

identified as the effects of large trials varied considerably (Fig. 2). For example, for the 10 largest trials the standardized mean difference spanned from 0.15 to  $-0.66$ . There was no statistically significant association between the effect of placebo and sample size ( $P = 0.24$ ). This is in contrast to the first version of our review ( $P = 0.05$ ). The degree of small trial bias was almost the same in the two versions (close to identical intercepts in Egger's regression analysis), and the lack of statistical significance is therefore probably caused by the large variability amongst big trials in the updated version. There was moderate heterogeneity ( $P < 0.001$ ,  $I^2 = 45\%$ ).

Because of these problems, it is a questionable procedure to pool all the trials, and we show the results mainly for completeness. There was an overall positive effect of placebo for continuous outcomes, standardized mean difference  $-0.24$  ( $-0.31$  to  $-0.17$ ). The effect for patient-reported outcomes was  $-0.30$  ( $-0.38$  to  $-0.21$ ), whereas no statistically significant pooled effect was found for observer-reported outcomes, standardized mean difference

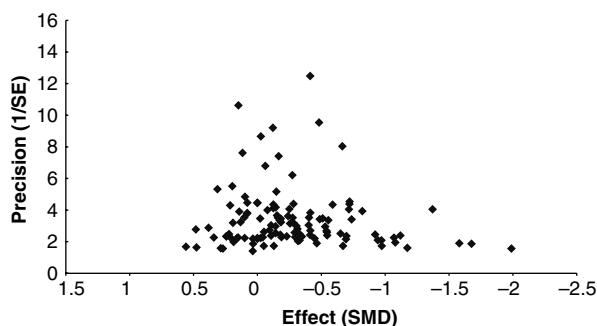


Fig. 2 Funnel plot of 118 trials with continuous outcomes. SMD, standardized mean difference; SE, standard error.

$-0.10$  ( $-0.20$  to  $0.01$ ). This considerable difference between patient- and observer-reported outcomes is statistically significant ( $P = 0.002$ ) (Table 1).

Ten clinical problems had been investigated in at least three trials with continuous outcomes: pain, obesity, asthma, hypertension, insomnia, nausea, depression, anxiety, phobia and smoking. Confidence intervals were wide for most conditions and placebo had a statistically significant effect only on pain and phobia (Table 2). There was substantial heterogeneity for trials investigating the effect on smoking ( $P = 0.02$ ,  $I^2 = 75\%$ ) (Table 2).

**Subgroup analyses.** There were no statistically significant differences in the effect of placebo between the three types of placebo interventions (Table 3), or other types of subgroups, except for a small negative effect of placebo interventions in five trials with continuous laboratory data (Table 4). The effect of placebo in the trials with clearly concealed allocation, dropout rate of 15% or less, and sample size of 50 or more, was relative risk 0.98 (0.87–1.09) for three trials with binary outcomes, and standardized mean difference  $-0.14$  ( $-0.32$  to  $-0.03$ ) for three trials with continuous outcomes. The degree of heterogeneity was lower for trials with clearly concealed allocation:  $I^2$  was 3% when outcomes were binary and 1% when continuous. For trials with unclear concealment of allocation  $I^2$  was 42% when outcomes were binary and 47% when continuous.

Table 3 Effect of three types of placebo treatment in all trials

	<i>n</i>	<i>k</i>	Pooled values	$I^2$ (%)
Type of placebo, binary data <sup>a</sup>				
Pharmacological	3119	22	0.97 (0.88–1.07)	45
Physical	925	8	0.95 (0.86–1.05)	0
Psychological	240	8	0.89 (0.74–1.06)	69
Type of placebo, continuous data <sup>b</sup>				
Pharmacological	4156	38	$-0.14$ ( $-0.26$ to $-0.03$ )	61
Physical	2092	37	$-0.30$ ( $-0.42$ to $-0.17$ )	41
Psychological	1205	43	$-0.31$ ( $-0.44$ to $-0.18$ )	17

*n*, number of trial participants; *k*, number of trials. <sup>a</sup>Pooled relative risk (a figure below 1 indicates a beneficial effect, 95% CI in brackets). <sup>b</sup>Pooled standardized mean difference (a negative figure indicates a beneficial effect).  $I^2$ : the percentage of the variation of the estimated effect between the trials that cannot be accounted for by random variation. The typical pharmacological placebo intervention was a tablet without active content. The typical physical placebo implied some kind of manual procedure, for example sham acupuncture. The typical psychological placebo was a nondirectional, neutral discussion between patient and treatment provider, a so-called 'attention placebo'.

**Table 4** Effect of placebo treatment according to types of outcomes in all trials

	<i>n</i>	<i>k</i>	Pooled values	<i>I</i> <sup>2</sup> (%)
<i>Binary outcomes</i>				
Observer-reported <sup>a</sup>				
Laboratory procedures	1423	4	0.92 (0.73–1.17)	72
Patient cooperation not essential	340	3	0.88 (0.68–1.14)	54
Patient cooperation essential	144	4	0.92 (0.77–1.09)	0
Patient-reported <sup>b</sup>				
Potentially observable	901	17	0.92 (0.79–1.06)	54
Nonobservable	1476	10	0.98 (0.90–1.07)	0
<i>Continuous outcomes</i>				
Observer-reported <sup>a</sup>				
Laboratory procedures	729	5	0.16 (0.01–0.30)	0
Patient cooperation not essential	880	21	–0.14 (–0.32 to 0.04)	24
Patient cooperation essential	645	17	–0.23 (–0.38 to –0.07)	0
Patient-reported <sup>b</sup>				
Potentially observable	1849	20	–0.34 (–0.48 to –0.20)	31
Nonobservable	3350	55	–0.30 (–0.40 to –0.20)	46

*n*, number of trial participants; *k*, number of trials. <sup>a</sup>Pooled relative risk (a figure below 1 indicates a beneficial effect, 95% CI in brackets).

<sup>b</sup>Pooled standardized mean difference (a negative figure indicates a beneficial effect). *I*<sup>2</sup>: the percentage of the variation of the estimated effect between the trials that cannot be accounted for by random variation.

The degree of heterogeneity was also lower for trials with observer-reported continuous outcomes (*I*<sup>2</sup> = 23%) than for patient-reported outcomes (*I*<sup>2</sup> = 45%).

*Trials without extracted outcome data.* In 26 trials, outcome data had not been reported in a way that was suited for meta-analysis. There was no clear tendency for the findings in these trials to be different from the findings in the 156 trials we meta-analysed.

#### *Clinical conditions with a statistically significant effect of placebo*

*Pain.* Forty-four trials (2833 patients) evaluated the effect on pain as a continuous outcome, for example measured on a 100-mm visual analogue scale. The heterogeneity amongst trials was close to being statistically significant (*P* = 0.06), but the degree of heterogeneity was low (*I*<sup>2</sup> = 26%). The pooled standardized mean difference was –0.25 (–0.35 to –0.16). The effect was also seen in six trials with clearly concealed allocation, standardized mean difference –0.22 (–0.44 to 0.00), and in 15 trials with dropout rates below 15%, standardized mean difference –0.25 (–0.42 to –0.09). As the mean standard deviation of pain measurements on 100 mm visual analogue scales was 24 mm, the effect of placebo corresponded to a reduction in pain intensity of 6 mm (3.8–8.4). The funnel plot of the pain trials was asymmetri-

cal (not shown), indicating smaller effects in larger trials (*P* = 0.05). The pooled standardized mean difference in 10 trials with 75 patients or more was –0.17 (–0.32 to –0.02). The apparent positive result was not reproduced in the five trials with binary outcomes (565 patients), pooled relative risk 0.98 (0.88–1.10).

*Phobia.* Three trials (57 patients) evaluated the effect of placebo on phobia as a continuous outcome, for example assessment of fear of snakes. The pooled standardized mean difference was –0.63 (–1.17 to –0.08). No heterogeneity was detected (*P* = 0.52, *I*<sup>2</sup> = 0%). The trials were very small with sample sizes of 14, 18 and 25 patients, and the concealment of allocation was unclear. Phobia was not investigated in trials with binary outcomes.

## Discussion

The majority of the 42 new trials we analysed were published after 1998 and therefore reflect contemporary clinical practice and research methodology. The results were very similar, however, to those we have reported on previously, based on 114 older trials [6, 8].

In the combined analyses of all 156 trials, we did not find a statistically significant effect of placebo interventions in trials with binary outcomes, or when continuous outcomes were reported by observers,

whereas a statistically significant effect was observed for trials with patient-reported continuous outcomes, especially for pain. Our results are incompatible with placebo interventions in general causing large effects on most clinical conditions.

It is important to realize that the difference between placebo groups and no-treatment groups does not equal the effect of placebo as such a comparison is unblinded. Thus, even if there were no true effect of placebo, one would expect to measure differences due to reporting bias, attrition bias and other forms of bias related to lack of blinding [3]. Reporting bias is particularly problematic. Most patients are polite and prone to please the investigators by reporting improvement, even when no improvement was felt. It is difficult to separate such reporting bias from true effects of placebo, but we suspect reporting bias occurred. For example, the estimated effects of placebo were three times higher for patient-reported continuous outcomes than for observer-reported outcomes (standardized mean difference  $-0.30$  vs.  $-0.10$ ,  $P = 0.002$ ).

Another reason why the pooled effect from the trials with continuous outcomes should be interpreted with considerable caution is that the funnel plot of the trials was asymmetrical without a clear peak, indicating not only that small trials tended to have larger effects than big trials, but also that the heterogeneity amongst large trials was considerable (see Results). This pattern can be caused by lack of identification of small trials with neutral or negative effects, by lower quality of the small trials or by differential true effects amongst different types of trials [11].

Patients in untreated control groups may seek alternative treatments outside the setting of a trial more often than patients in placebo groups. We expected that this 'co-intervention' bias could be more pronounced in trials where placebo was the only treatment offered. However, we found no statistically significant difference in pooled effects between such trials, and trials where placebo was added to a standard treatment (also given to the untreated group).

To find no reliable evidence of an effect is not the same as evidence of no effect. Our sample of trials was very large, but it was also heterogeneous. We conducted several subgroup analyses without finding large effects of placebo (except for the effect on phobia which we regard as unreliable due to few, small, low

quality trials). However, we cannot exclude the possibility that in the process of pooling heterogeneous trials the existence of such a group was obscured. Our conclusions are also limited to the clinical conditions and outcomes studied; for example, there were few trials reporting broad outcomes such as patient's quality of life or well-being.

Even if we disregard the likely bias involved in the assessment of pain, the estimated analgesic effect of placebo corresponded to only 6 mm on a 100-mm visual analogue scale. It is doubtful whether this represents a clinically relevant reduction in pain. A systematic review of pain trials studying mechanisms of the effect of placebo (and not primarily the clinical effect) reported higher effects [13]. It is not clear whether this reflects a true difference in effect between the two settings, or a different susceptibility to reporting bias and other biases.

We have no good explanation for the difference between effects of placebo when measured on a binary and on a continuous scale, but continuous scales could be more sensitive to small effects or biases.

It is a question of definition whether the effect of a placebo intervention equals the 'placebo effect', as this term is sometimes also used for other aspects of the patient-provider interaction, for example psychologically mediated effects in general, the effect of suggestion, the effect of expectancies, the effect of patients' experience of meaning, etc. [3]. Patients in a no-treatment group also interact with treatment providers, and the patients are therefore only truly untreated with respect to receiving a placebo intervention. Hence, our results do not exclude the possibility that other aspects of the patient-provider interaction, or interactions between the treatment ritual and different ways of informing patients, could have clinically useful effects.

Despite ethical concerns of the deception inherent in most placebo prescriptions [14], the clinical use of placebo interventions has been advocated in editorials and articles in leading journals [5, 15] and by influential commentators [16]. A survey of 772 randomly sampled Danish clinicians found that 48% (41–55%) of general practitioners used what they regarded a placebo intervention 10 times or more per year. Many of the interventions were vitamins for unspecified fatigue and antibiotics for viral infections. The main reasons for using a placebo intervention were to avoid confrontation with the patient and to utilize a perceived effect of placebo [17].

Placebo interventions are important in clinical research as a tool for blinding. We excluded trials in which it was clear that an unblinded assessor evaluated observer-reported outcomes, and in most cases we disregarded follow-up data. Our results therefore probably underestimate the bias caused by use of no-treatment control groups rather than placebo groups in randomized trials. Even so, our results illustrate the risk of bias involved in unblinded trials. For example, in double-blind placebo-controlled trials the effect of nonsteroidal, anti-inflammatory drugs (NSAIDs) on arthritis pain was standardized mean difference  $-0.84$  [18]. If this assessment had been based on trials with no-treatment groups, and assuming the results of our review were correct (standardized mean difference  $-0.24$  for placebo versus no treatment), the estimated effect of NSAIDs would have been  $-0.84 - 0.24 = -1.08$ , which is an overestimate of 29%.

In conclusion, we reproduced the findings of our previous review and found no evidence that placebo interventions in general have large clinical effects, and no reliable evidence that they have clinically useful effects. A possible effect on patient-reported continuous outcomes, especially on pain, could not be clearly distinguished from bias.

## Conflicts of interest

No conflict of interest was declared.

## Acknowledgements

We thank Roberto Oliveri who translated newly identified trial reports from Italian, and the researchers of the new trials who provided access to additional data.

## References

- 1 Beecher HK. The powerful placebo. *JAMA* 1955; **159**: 1602–6.
- 2 Kienle GS, Kiene H. The powerful placebo effect: fact or fiction? *J Clin Epidemiol* 1997; **12**: 1311–8.
- 3 Hróbjartsson A. What are the main methodological problems in the estimation of placebo effects? *J Clin Epidemiol* 2002; **55**: 430–5.
- 4 Ernst E, Resch KL. Concept of true and perceived placebo effects. *BMJ* 1995; **311**: 551–3.
- 5 Brown WA. The placebo effect. *Sci Am* 1998; **278**: 68–73.
- 6 Hróbjartsson A, Gøtzsche PC. Is the placebo powerless? An analysis of clinical trials comparing placebo treatment with no treatment. *N Engl J Med* 2001; **344**: 1594–602.
- 7 Brody H, Weismantel D. A challenge to core beliefs. *Adv Mind Body Med* 2001; **17**: 296–8; discussion 312–8.
- 8 Hróbjartsson A, Gøtzsche PC. Placebo treatment versus no treatment (Cochrane Review). In: *The Cochrane Library*, Issue 1. Chichester, UK: John Wiley & Sons Ltd, 2003.
- 9 DerSimonian R, Laird N. Meta-analysis in clinical trials. *Control Clin Trials* 1986; **7**: 177–88.
- 10 Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ* 2003; **327**: 557–60.
- 11 Egger M, Smith GD, Schneider M, Minder C. Bias in meta-analysis detected by a simple, graphical test. *BMJ* 1997; **315**: 629–34.
- 12 Hróbjartsson A, Gøtzsche PC. Placebo treatment for all conditions (Cochrane Review). In: *The Cochrane Library*, Issue 3. Chichester, UK: John Wiley & Sons Ltd, 2004.
- 13 Vase L, Riley JL III, Price DD. A comparison of placebo effects in clinical analgesic trials versus studies of placebo analgesia. *Pain* 2002; **99**: 443–52.
- 14 Rawlinson MC. Truth-telling and paternalism in the clinic: philosophical reflections on the use of placebos in medical practice. In: White L, Tursky B, Schwartz GE, eds. *Placebo. Theory, Research, Mechanisms*. New York: Guilford Press, 1985; pp. 403–18.
- 15 Ho VMS. The placebo effect: can we use it better? *BMJ* 1994; **309**: 69–70.
- 16 Cochrane AL. *Effectiveness and Efficiency. Random Reflections on Health Services*. UK: Cambridge University Press for British Medical Journal & The Nuffield Provincial Hospitals Trust, 1989; p. 31.
- 17 Hróbjartsson A, Norup M. The use of placebo interventions in medical practice – a national questionnaire survey of Danish clinicians. *Eval Health Prof* 2003; **26**: 153–65.
- 18 Gøtzsche PC. Sensitivity of effect variables in rheumatoid arthritis: a meta-analysis of 130 placebo controlled NSAID trials. *J Clin Epidemiol* 1990; **43**: 1313–8.
- 19 Alfano AP, Taylor AG, Foresman PA *et al*. Static magnetic fields for treatment of fibromyalgia: a randomised controlled trial. *J Altern Compl Med* 2001; **7**: 53–64.
- 20 Alkaiisi A, Stålnert M, Kalman S. Effect and placebo effect of acupressure (P6) on nausea and vomiting after outpatient gynaecological surgery. *Acta Anaesthesiol Scan* 1999; **43**: 270–4.
- 21 Alkaiisi A, Evarsson K, Johnsson VA, Ofenbartl L, Kalman S. P6 acupressure may relieve nausea and vomiting after gynaecological surgery: an effectiveness study in 410 women. *Can J Anesth* 2002; **49**: 1034–9.
- 22 Anderson B, Brackett J, Ho J, Laffel LMB. An office-based intervention to maintain parent-adolescent teamwork in diabetic management. Impact on parent involvement, family conflict, and subsequent glycemic control. *Diabetes Care* 1999; **22**: 713–21.
- 23 Antonio J, Colker CM, Torina G, Shi Q, Brink W, Kalman D. Effects of standardised guggulsterone phosphate supplement on body composition in overweight adults: a pilot study. *Curr Ther Res* 1999; **60**: 220–7.
- 24 Asmar R, Boutelant S, Chaignon M *et al*. Repeated measurements of non-invasive ambulatory blood pressure: distinction between reproducibility and the proper effect of placebo. *Blood Press Monit* 1996; **1**: 283–8.
- 25 Blades KJ, Patel S, Aidoo KE. Oral antioxidant therapy for marginal dry eye. *E J Clin Nutr* 2001; **55**: 589–97.



- 26 Blanchard EB, Appelbaum KA, Radnitz CL *et al.* Placebo-controlled evaluation of abbreviated progressive muscle relaxation and of relaxation combined with cognitive therapy in the treatment of tension headache. *J Consult Clin Psychol* 1990; **58**: 210–5.
- 27 Blanchard EB, Appelbaum KA, Radnitz CL *et al.* A controlled evaluation of the thermal biofeedback and thermal biofeedback combined with cognitive therapy in the treatment of vascular headache. *J Consult Clin Psychol* 1990; **58**: 216–24.
- 28 Bova JG, Bhattacharjee N, Jurdi R, Bennet WF. Comparison of no medication, placebo and hyoscyamine for reducing pain during a barium enema. *AJR* 1999; **172**: 1285–7.
- 29 Carbajal R, Chauvet X, Couderc S, Olivier-Martin M. Randomised trial of analgesic effects of sucrose, glucose, and pacifiers in term neonates. *BMJ* 1999; **319**: 1393–7.
- 30 Colker CM, Kalman DS, Torina GC, Perlis T, Street C. Effects of *Citrus aurantium* extract, caffeine, and St John's wort on body fat loss, lipid levels, and mood states in overweight healthy adults. *Curr Ther Res* 1999; **60**: 145–53.
- 31 Cupal DD, Brewer BW. Effects of relaxation and guided imagery on knee strength, reinjury anxiety, and pain following anterior cruciate ligament reconstruction. *Rehabil Psychol* 2001; **46**: 28–43.
- 32 Etringer BD, Cash TF, Rimm DC. Behavioral, affective and cognitive effects of participant modeling and an equally credible placebo. *Behav Ther* 1982; **13**: 476–85.
- 33 Etter J, Lazlo E, Zellweger J, Perrot C, Perneger TV. Nicotine replacement to reduce cigarette consumption in smokers who are unwilling to quit: a randomized trial. *J Clin Psychopharmacol* 2002; **22**: 487–95.
- 34 Fanti L, Gemma M, Passaretti S *et al.* Electroacupuncture analgesia for colonoscopy: a prospective, randomised, placebo-controlled study. *Am J Gastroenterol* 2003; **98**: 312–6.
- 35 Hossmann V, FitzGerald GA, Dollery CT. Influence of hospitalization and placebo therapy on blood pressure and sympathetic function in essential hypertension. *Hypertension* 1981; **3**: 113–8.
- 36 Kilman PR, Milan RJ, Boland JP *et al.* Group treatment of secondary erectile dysfunction. *J Sex Marital Ther* 1987; **13**: 168–82.
- 37 Kober A, Scheck T, Greher M *et al.* Prehospital analgesia with acupuncture in victims of minor trauma: a prospective, randomized, double-blinded trial. *Anesth Analg* 2002; **95**: 723–7.
- 38 Kotani N, Kushikata T, Suzuki A, Hashimoto H, Muraoka M, Matsuki A. Insertion of intradermal needles into painful points provides analgesia for intractable abdominal scar pain. *Reg Anesth Pain Med* 2001; **26**: 532–8.
- 39 Leibing E, Leonhardt U, Köster G *et al.* Acupuncture treatment of chronic low-back pain – a randomized, blinded, placebo-controlled trial with 9-months follow-up. *Pain* 2002; **96**: 189–96.
- 40 Lin JG, Lo MW, Wen YR, Hsieh CL, Tsai SK, Sun WZ. The effect of high and low frequency electroacupuncture in pain after lower abdominal surgery. *Pain* 2002; **99**: 509–14.
- 41 Molsberger AF, Mau J, Pawelec DB, Winkler J. Does acupuncture improve the orthopedic management of chronic low back pain – a randomized, blinded, controlled trial with 3 months follow up. *Pain* 2002; **99**: 579–87.
- 42 Rawling MJ, Wiebe ER. A randomized controlled trial of fentanyl for abortion pain. *Am J Obstet Gynecol* 2001; **185**: 103–7.
- 43 Ristikankare M, Julkunen R, Mattila M *et al.* Conscious sedation and cardiorespiratory safety during colonoscopy. *Gastrointest Endosc* 2000; **52**: 48–54.
- 44 Robinson R, Darlow S, Wright SJ *et al.* Is transcutaneous electrical nerve stimulation an effective analgesia during colonoscopy. *Postgrad Med J* 2001; **77**: 445–6.
- 45 Roongpisuthipong C, Panpakdee O, Boontawee A, Kulapongse S, Tanphaichitr V. Possible thermogenesis with dexfenfluramine. *J Med Assoc Thai* 1999; **82**: 150–9.
- 46 Roscoe JA, Morrow GR, Bushunow P, Tian L, Matteson S. Acustimulation wristbands for the relief of chemotherapy-induced nausea. *Altern Ther* 2002; **8**: 56–62.
- 47 Roschke J, Wolf C, Müller MJ *et al.* The benefit from whole body acupuncture in major depression. *J Affect Disord* 2000; **57**: 73–81.
- 48 Scharff L, Marcus DA, Masek BJ. A controlled study of minimal contact thermal biofeedback treatment in children with migraine. *J Pediatr Psychol* 2002; **27**: 109–19.
- 49 Shallreuter KU, Moore J, Behrens-Williams S, Panske A, Harari M. Rapid initiation of repigmentation in vitiligo with dead sea climatotherapy in combination with pseudocatalase (PC-KUS). *Int J Dermatol* 2002; **41**: 482–7.
- 50 Shen J, Wenger N, Glaspy J *et al.* Electroacupuncture for control of myeloablative chemotherapy-induced emesis. *JAMA* 2000; **284**: 2755–61.
- 51 Sinaiko AR, Gómez-Marín O, Prineas RJ. Effect of low sodium diet or potassium supplementation on adolescent blood pressure. *Hypertension* 1993; **21**: 989–94.
- 52 Straub WF, Spino MP, Alattar MM *et al.* The effect of chiropractic care on jet lag of Finnish junior elite athletes. *J Manipulative Physiol Ther* 2001; **24**: 191–8.
- 53 Sumaya IC, Rienzi BM, Deegan JF, Moss DE. Bright light treatment decreases depression in institutionalized older adults: a placebo-controlled crossover study. *J Gerontol A Biol Sci Med Sci* 2001; **56A**: M356–60.
- 54 Thomas VJ, Dixon AL, Milligan P. Cognitive-behaviour therapy for the management of sickle cell disease pain: an evaluation of a community based intervention. *Br J Health Psychol* 1999; **4**: 209–29.
- 55 Thomas KS, Muir KR, Doherty M, Jones AC, O'Reilly SC, Bassey EJ. Home based exercise programme for knee pain and knee osteoarthritis: randomised controlled trial. *BMJ* 2002; **325**: 752–7.
- 56 Tritrakarn T, Lertakyamane J, Koompong P *et al.* Both EMLA and placebo cream reduced pain during extracorporeal piezoelectric shock wave lithotripsy with the piezolith 2300. *Anaesthesiology* 2000; **92**: 1049–54.
- 57 Tsay SL, Chen ML. Acupressure and quality of sleep in patients with end-stage renal disease – a randomised controlled trial. *Int J Nurs Stud* 2002; **40**: 1–7.
- 58 Wang B, Tang J, White PF *et al.* Effect of the intensity of transcutaneous acupoint electrical stimulation on the postoperative analgesic requirement. *Anesth Analg* 1997; **85**: 406–13.
- 59 Werntoft E, Dykes A. Effect of acupressure on nausea and vomiting during pregnancy. A randomized, placebo-controlled, pilot study. *J Reprod Med* 2001; **46**: 835–9.

Correspondence: A. Hróbjartsson, The Nordic Cochrane Centre, Rigshospitalet, Department 7112, Blegdamsvej 9, DK 2100 Copenhagen Ø, Denmark.  
(fax: +45 35 45 70 07; e-mail: a.hrobjartsson@cochrane.dk).

## Appendix Table of included trials identified from 1999 to 2003 (with available data)

Trials with binary outcomes			
	Disease	<i>n</i>	RR (95% CI)
Fanti <i>et al.</i> [34]	Pain	20	0.89 (0.61–1.29)
Molsberger <i>et al.</i> [41]	Pain	111	1.16 (0.86–1.56)
Scharff <i>et al.</i> [48]	Pain	23	0.89 (0.64–1.24)
Alkaissi <i>et al.</i> [20]	Nausea	40	0.82 (0.44–1.53)
Alkaissi <i>et al.</i> [21]	Nausea	275	0.92 (0.75–1.12)
Shallreuter <i>et al.</i> [49]	Vitiligo	20	0.75 (0.22–2.52)
Trials with continuous outcomes			
	Disease	<i>n</i>	SMD (95% CI)
Alfano <i>et al.</i> [19]	Pain	38	–0.14 (–0.80 to 0.52)
Blanchard <i>et al.</i> [26]	Pain	24	–0.70 (–1.53 to 0.13)
Blanchard <i>et al.</i> [27]	Pain	42	–0.28 (–0.90 to 0.33)
Bova <i>et al.</i> [28]	Pain	70	–0.14 (–0.61 to 0.33)
Carbajal <i>et al.</i> [29]	Pain	50	0.10 (–0.45 to 0.66)
Cupal and Brewer [31]	Pain	20	0.00 (–0.88 to 0.88)
Kober <i>et al.</i> [37]	Pain	41	0.19 (–0.42 to 0.80)
Kotani <i>et al.</i> [38]	Pain	47	–0.55 (–0.14 to 0.03)
Leibing <i>et al.</i> [39]	Pain	79	–0.59 (–1.04 to –0.14)
Lin <i>et al.</i> [40]	Pain	50	–0.51 (–1.08 to 0.05)
Rawling and Wiebe [42]	Pain	185	–0.06 (–0.35 to 0.23)
Robinson <i>et al.</i> [44]	Pain	23	–0.11 (–0.93 to 0.72)
Thomas <i>et al.</i> [54]	Pain	40	–0.31 (–0.96 to 0.35)
Thomas <i>et al.</i> [55]	Pain	156	–0.27 (–0.59 to 0.04)
Thomas <i>et al.</i> [55]	Pain	233	0.12 (–0.14 to 0.37)
Tritrakarn <i>et al.</i> [56]	Pain	82	–0.72 (–1.17 to –0.27)
Wang <i>et al.</i> [58]	Pain	51	–0.40 (–0.96 to 0.15)
Asmar <i>et al.</i> [24]	Hypertension	68	–0.12 (–0.60 to 0.35)
Hossmann <i>et al.</i> [35]	Hypertension	24	–0.18 (–0.98 to 0.62)
Sinaiko <i>et al.</i> [51]	Hypertension	87	0.21 (–0.24 to 0.67)
Roscoe <i>et al.</i> [46]	Nausea	54	–0.15 (–0.69 to 0.38)
Shen <i>et al.</i> [50]	Nausea	67	–0.25 (–0.73 to 0.23)
Werntoft and Dykes [59]	Nausea	40	–0.26 (–0.88 to 0.37)
Antonio <i>et al.</i> [23]	Obesity	12	0.56 (–0.60 to 1.72)
Colker <i>et al.</i> [30]	Obesity	11	0.29 (–0.95 to 1.53)
Roongpisuthipong <i>et al.</i> [45]	Obesity	37	–0.11 (–0.75 to 0.54)
Roschke <i>et al.</i> [47]	Depression	48	–0.18 (–0.75 to 0.38)
Sumaya <i>et al.</i> [53]	Depression	20	0.21 (–0.67 to 1.09)
Tsay and Chen [57]	Insomnia	64	–0.08 (–0.57 to 0.41)
Etter <i>et al.</i> [33]	Smoking	658	–0.41 (–0.57 to –0.26)
Etringer <i>et al.</i> [32]	Phobia	25	–0.33 (–0.12 to 0.46)
Straub <i>et al.</i> [52]	Jet lag	10	0.27 (–0.98 to 1.52)
Kilmann <i>et al.</i> [36]	Erectile dysfunction	8	0.04 (–1.35 to 1.42)
Blades <i>et al.</i> [25]	Dry eye	80	0.00 (–0.44 to 0.44)
Anderson <i>et al.</i> [22]	Diabetic care	57	0.08 (–0.44 to 0.60)
Ristikankare <i>et al.</i> [43]	Difficulty with colonoscopy	122	0.20 (–0.16 to 0.55)

A list of included trials without extractable data, and selected excluded trials, can be seen in the Cochrane version of the review, or by contacting the authors. *n*, number of participants; RR, relative risk; SMD, standardized mean difference.