Both reviewers independently selected outcomes by referring only to the methods sections of articles; any disagreements were resolved by discussion. As the primary outcome, we selected the main objective or parallel-group trial (i.e., using the between-group standard deviations and total number of participants for both groups).

Conclusions
We found little evidence in general that placebos had powerful clinical effects. Although placebos had no significant effects on objective or binary outcomes, placebo effects were more consistent in studies with continuous subjective outcomes and for the treatment of pain. Outside the setting of clinical trials, there is no justification for the use of placebos.

Placebos have been reported to improve subjective and objective outcomes in up to 30 to 40 percent of patients with a wide range of clinical conditions, such as pain, asthma, high blood pressure, and even myocardial infarction. In his 1955 article "The Powerful Placebo," Beecher noted that placebos have a high degree of therapeutic effectiveness in treating subjective responses, decided improvement, interpreted under the unknowns technique as a real therapeutic effect, being produced in 35±2.2±2.2% of cases.

Back
Background

We conducted a systematic review of clinical trials in which patients were randomly assigned to either placebo or no treatment. A placebo could be pharmacologic (e.g., a tablet), physical (e.g., a manipulation), or psychological (e.g., a conversation).

Results
We identified 130 trials that met our inclusion criteria. After the exclusion of 16 trials without relevant data on outcomes, there were 32 with binary outcomes (involving 37,915 patients, with a median of 51 patients per trial) and 82 with continuous outcomes (involving 47,305 patients, with a median of 27 patients per trial). As compared with no treatment, placebo had no significant effect on binary outcomes, regardless of whether these outcomes were subjective or objective. For the trials with continuous outcomes, placebo had a beneficial effect, but the effect decreased with increasing sample size, indicating a possible bias related to the effects of small trials. The pooled standardized mean difference was significant for the trials with objective outcomes but not for those with subjective outcomes. In 27 trials involving the treatment of pain, placebo had a beneficial effect, as indicated by a reduction in the intensity of pain of 6.5 mm on a 100-mm visual-analogue scale.

We calculated the pooled relative risk of an unwanted outcome for trials with binary outcomes and the pooled standardized mean difference for those with continuous outcomes.13

Methods
Definition of Placebo
Placebo is difficult to define satisfactorily. In clinical trials, placebos are generally control treatments with a similar appearance to the study treatments but without their specific activity. We therefore defined placebo practically as an intervention labeled as such in the report of a clinical trial.

Literature Search
We searched Medline, EMBASE, PsycLIT, Biological Abstracts, and the Cochrane Controlled Trials Register for trials published before the end of 1998. The search was developed iteratively for synonyms of "placebo," "no treatment," and "randomized clinical trial" (the exact search strategy is available as Supplementary Appendix 1 with the full text of this article at http://www.nejm.org and was based on a published protocol). We systematically read the reference lists of included trials and selected books and review articles. We also asked researchers in the field to provide lists of relevant trials.

Selection of Studies
We included studies if patients were assigned randomly to a placebo group or an untreated group (often there was also a third group that received active treatment). We excluded studies if randomization was clearly not concealed — that is, if group assignment was unpredictable (e.g., patients were assigned to treatment groups according to the day of the month). We also excluded studies if participants were paid or were healthy volunteers, if the person who assessed objective outcomes was aware of group assignments, if the dropout rate exceeded 50 percent, or if it was very likely that the alleged placebo had a clinical effect not associated with the treatment used (e.g., patients were assigned to treatment groups according to the day of the month). We also excluded studies if participants were paid or were healthy volunteers, if the person who assessed objective outcomes was aware of group assignments, if the dropout rate exceeded 50 percent, or if it was very likely that the alleged placebo had a clinical effect not associated with the treatment used (e.g., patients were assigned to treatment groups according to the day of the month).

Extraction of Data
Data were extracted from the report of each trial with the use of forms tested in pilot studies. We contacted the authors of the included studies when reported outcome data were inadequate for meta-analysis. We noted how the randomization was conducted and whether the therapist responsible for the administration of placebo (as distinct from the observer) was unaware of group assignments. Furthermore, we noted the purpose of the trial, the dropout rate, whether the placebo was given in addition to the standard treatment, and whether the main outcome was clearly indicated.

We noted whether the placebo was pharmacologic (e.g., a tablet), or psychological (e.g., a conversation); whether clinical problems reported by the patients could have been observed by others (i.e., whether the symptoms were observable outcomes such as cough); and whether objective outcomes were laboratory data, were derived from examinations that required the cooperation of the patients (i.e., objective outcomes such as forced expiratory volume), or did not require such cooperation (e.g., edema).

Both reviewers independently selected outcomes by referring only to the methods sections of articles; any disagreements were resolved by discussion. As the primary outcome, we selected the main objective or subjective outcome of each trial (preferably a characteristic symptom). If a main outcome was not indicated, we used the outcome that we felt was most relevant to patients. Binary outcomes (e.g., the proportions of smokers and nonsmokers) were preferred to continuous ones (e.g., the mean number of cigarettes smoked). Data recorded immediately after the end of treatment were preferred to follow-up data, although end-of-treatment data were not always available. For crossover trials, we extracted data from the first treatment period only; if that was not possible, we used the summary data as if they had been derived from a parallel-group trial (i.e., using the between-group standard deviations and total number of participants for both groups).

Synthesis of Data
For each trial with binary outcomes, we calculated the relative risk of an unwanted outcome, which was defined as the ratio of the number of patients with an unwanted outcome to the total number of patients in the placebo group, divided by the same ratio in the untreated group. Thus, a relative risk below 1.0 indicates a beneficial effect of placebo.

For trials with continuous outcomes, we calculated the standardized mean difference, which was defined as the difference between the mean value for an unwanted outcome in the placebo group and the corresponding mean value in the untreated group divided by the pooled standard deviation. A value of 1.0 indicates that the mean in the placebo group was 1 SD below the mean in the untreated group, indicating a beneficial effect of placebo.

We calculated the pooled relative risk of an unwanted outcome for trials with binary outcomes and the pooled standardized mean difference for those with continuous outcomes. Because of the different clinical conditions and settings, we expected that the data sets would be heterogeneous — that is, that the effects of individual trials would vary more than expected by chance alone. The variance and statistical significance of the differences were therefore assessed with the use of random-effect calculations. We calculated the pooled effects for subjective and objective outcomes and for specific clinical problems that had been investigated in at least three trials by different research groups.

We performed preplanned analyses of subgroups to see whether our findings were sensitive to the type of placebo or the type of outcome involved. Furthermore, for each trial, we plotted the effect against the inverse of its standard error (which increases with the number of trial participants). Since the variation in the estimated effect decreases with increasing sample size, the plot is expected to resemble a symmetrical funnel. If there is significant asymmetry in such funnel plots, it is usually caused by small trials' reporting greater effects, on average, than large trials, which can reflect publication bias.

In trials with continuous outcomes, we used F tests to check whether the standard deviations of the placebo group and the untreated group were significantly different. We regarded the distributions of either group as
We identified 727 potentially eligible trials. We subsequently excluded 597 trials for the following reasons: 404 were nonclinical or nonrandomized, 129 were missing a placebo group or an untreated group, 29 were reported in more than one publication, 11 had clearly unblinded assessment of objective outcomes, and 24 met other criteria for exclusion, such as dropout rates over 50 percent. No relevant outcome data were available for 16 of the remaining 130 trials. The analysis therefore included 114 trials. 

There were 10 crossover trials, of which 7 (which included a total of 182 patients) were handled as parallel trials. In 112 trials, there was a third group assigned to active treatment in addition to the placebo and the untreated groups. In 88 of these, determining the effect of placebo was not mentioned as an objective of the study. The trial reports were published in five languages between 1946 and 1998. The outcomes were binary in 32 trials, continuous in 26, both binary and continuous in 37, and continuous in 8.

In 76 trials, the outcome in the data we extracted was identified as a main outcome by the authors of the trials. If only patients in the placebo and untreated groups were counted, the trials with binary outcomes included 3795 patients with a median of 31 patients per trial (interquartile range, 26 to 72), and the trials with continuous outcomes included 4730 patients with a median of 27 patients per trial (interquartile range, 20 to 52).

The typical pharmacologic placebo was a lactose tablet. The typical psychological placebo was a nondirectional, neutral discussion between the patient and the treatment provider, referred to as an "attention placebo." No treatment typically entailed observation only or standard therapy; in the latter case, all patients in the trial received standard therapy, and the placebo was additional.

The results for the individual trials are available as Supplementary Appendix 2 and Supplementary Appendix 3 with the full text of this article at http://www.nejm.org. The trials investigated 40 clinical conditions: hypertension, asthma, anemia, hyperglycemia, hypercholesterolemia, headaches, Raynaud's disease, alcohol abuse, smoking, obesity, poor oral hygiene, herpes simplex infection, bacterial infection, common cold, pain, nausea, ileus, infertility, cervical dilation, labor, menopause, prostatism, depression, schizophrenia, insomnia, anxiety, phobia, compulsive nail biting, mental handicap, marital discord, stress related to dental treatment, organic difficulties, fecal soiling, enuresis, epilepsy, Parkinson's disease, Alzheimer's disease, attention-deficit-hyperactivity disorder, carpal tunnel syndrome, and undiagnosed ailments.

There was significant heterogeneity among the trials with binary outcomes (P=0.003), indicating that the variation in the effect of placebo among trials was larger than would be expected to result from chance alone. The heterogeneity was not due to small trials showing more pronounced effects of placebo than large trials (P=0.56).

Three clinical problems had been investigated in at least three independent trials with binary outcomes: nausea, relapse after the cessation of smoking, and depression. Placebo had no significant effect on these outcomes, but the confidence intervals were wide (Table 2).

There was significant heterogeneity among the trials with continuous outcomes (P=0.001). The magnitude of the effect of placebo decreased with increasing sample size (P=0.05), indicating a possible bias related to the effects of small trials.

Small trials involving the treatment of pain did not have significantly greater effects than large trials (P=0.2), but the power of the test was low. There was no significant heterogeneity among the nine sets of data on specific clinical problems (P=0.10), but the power of these analyses was also low.

Sensitivity Analyses

The number of trials compared in the sensitivity analyses was in most cases nine or more, and they included more than 1000 patients. There was no difference in the effect of placebo between subcategories of objective and subjective binary outcomes (Table 3). The effect of placebo among subcategories of continuous outcomes did not differ significantly, except for a negative effect of placebo in four trials with laboratory data (data not shown).

For both continuous and binary outcomes, there were no significant differences among the various types of placebo trials (Table 4).

The effect of placebo on continuous or binary outcomes was not influenced by the dropout rate (≤15 percent vs. >15 percent) or by whether the observers were aware of group assignments, but only two trials with binary objective outcomes (involving 316 patients) showed a significant effect of placebo as compared with no treatment (pooled standardized mean difference, –0.27; 95 percent confidence interval, –0.40 to –0.15). There was no significant effect of placebo on the other outcomes of interest.

Discussion

We did not detect a significant effect of placebo as compared with no treatment in pooled data from trials with subjective or objective binary or continuous objective outcomes. We did, however, find a significant difference between placebo and no treatment in trials with continuous subjective outcomes and in trials involving the treatment of pain.
Several types of bias may have affected our findings. Blinded evaluation of subjective outcomes was not possible in the trials we reviewed. Patients in an untreated group would know they were not being treated, and patients in the placebo group might think they had received treatment. It is difficult to distinguish between reporting bias and a true effect of placebos on subjective outcomes, since a patient may tend to try to please the investigator and report improvement when none has occurred. The fact that placebos had no significant effect on objective continuous outcomes suggests that reporting bias may have been a factor in the trials with subjective outcomes.

If patients in the untreated groups sought treatment outside the trials more often than patients in the placebo groups, the effects of placebo might be less apparent. Very few trials provided information on concomitant treatment. The risk of bias is expected to be larger in trials in which placebo is the only treatment and is not given in addition to standard therapy. We did not, however, find a difference in effect between the two types of trials.

There was some evidence that placebos had greater effects in small trials in controlled trials than in large trials. This could indicate that some small trials with negative outcomes have not been published or that we did not identify them. It is difficult to identify relevant trials in this field; another systematic search for trials involving placebo groups versus untreated groups found only 12 studies. We identified 114 trials from which the outcomes could be extracted, but 88 of these trials investigated the effect of active treatment in a third group of patients and did not explicitly study the effect of placebo. Because the publication of negative trials is often associated with the effect of studies with positive outcomes, the existence of unreported trials could explain the higher effects reported in small studies.

Poor methodology in small trials could also explain the larger effects of placebo. It is possible that we found no association between measures of the quality of a trial and placebo effects. However, the power of our sensitivity analyses may have been too low. Furthermore, it is possible that trials tended to investigate clinical conditions in which placebo may have had greater effects. Thus, although we found an effect of placebo on subjective continuous outcomes, the inverse relation between trial size and effect size implies that the estimated of pooled effect should be interpreted cautiously.

It will also be difficult to determine whether a placebo standardized mean difference is large enough to be clinically meaningful. Some individual trials reported clinically relevant effects with standardized mean differences of less than -0.6, but such “outlier” values may be spurious. If the possible biases we have discussed are disregarded, the pooled effect of placebo on pain corresponds to one third of the effect of nonopioid analgesics with placebo, in double-blind trials. It is uncertain whether such an effect is important for patients.

Our study has other limitations. We did extensive analyses of predefined subgroups according to the type of placebo, disease, and outcome without identifying a subgroup of trials in which the effect of placebo was large. However, we cannot exclude the possibility that, in the pooling of heterogeneous trials, the existence of such a subgroup was obscured. Our conclusions are also limited to the clinical conditions and outcomes that were investigated. It should be noted that few trials reported on the quality of life or patients’ well-being.

We reviewed the effect of placebos but not the effect of the patient-provider relationship. We could not rule out a psychological therapeutic effect of this relationship, which may be largely independent of any placebo intervention.

Moreover, the use of placebos in blinded, randomized trials is a precaution directed against many forms of bias and not only a way of controlling for the effects of placebos. Patients who are aware of their assignment may differ from unaware patients in their way of reporting beneficial and harmful effects of treatment, in their tendency to seek additional treatment outside the study, and in their risk of dropping out of the study. Furthermore, staff members who are aware of treatment assignments may differ from each other in their way of using alternative forms of treatment and in the assessment of outcomes. Thus, even if there was no true effect of placebo, we would expect to find differences between placebo and untreated groups because of bias associated with a lack of double-binding.

We were unable to detect any such significant difference in trials with subjective or objective binary or continuous outcomes. This surprising finding can possibly be explained by changes in the design and conduct of trials. Since our objective was to study the clinical effect of placebos, we reduced the influence of observer bias and bias due to dropouts by excluding trials with clearly unblinded objective outcomes and by attempting to analyze post-treatment data instead of follow-up data. In addition, since most trials we included did not primarily address the effect of a placebo but, rather, evaluated that of an active treatment, our study may have underestimated the clinical effects of placebos. Since the design of our review precludes estimation of the overall effect of bias due to lack of double-binding, our results do not imply that control groups that receive no treatment can be substituted for control groups of placebo receiving placebo without creating a risk of bias. This result is in accordance with an empirical study of 33 meta-analyses, which found that randomized trials that were not double-blinded yielded larger estimates than blinded trials, with odds ratios that were exaggerated by 17 percent.

In conclusion, we found little evidence that placebos in general have powerful clinical effects. Placebos had no significant pooled effect on subjective or objective binary or continuous objective outcomes. We found significant effects of placebo on subjective continuous outcomes and for the treatment of pain but also bias related to larger effects in small trials. The use of placebo outside the aegis of a controlled, properly designed clinical trial cannot be recommended.

Supported by a grant from the Faculty of Health Sciences at University of Copenhagen.

We are indebted to Henrik R. Wulff, Jos Kleijnen, and Iain Chalmers for valuable comments on previous versions of the manuscript; to Gunvor Kienle, Andrew Vickers, Harald Walach, Clive Adams, and Iain Chalmers for lists of relevant trials; and to the numerous placebo-trial researchers for access to additional data.

Source Information

From the Department of Medical Philosophy and Clinical Theory, University of Copenhagen, Panum Institute, and the Nordeco Centre, Rigshospitalet — both in Copenhagen, Denmark.

Address reprint requests to Dr. Hróbjartsson at the Department of Medical Philosophy and Clinical Theory, University of Copenhagen, Panum Institute, Blegdamsvej 3, DK-2200 Copenhagen N, Denmark, or at a.hrobjartsson@cochrane.dk

References

1. Beecher HK. The powerful placebo. JAMA 1955;159:1402-10. [ISI]
8. Oh YMS. The placebo effect: can we use it better? BMJ 1994;309:696-700. [Full Text]
NEJM -- Is the Placebo Powerless?- An Analysis of Clinical Trials C... http://content.nejm.org/cgi/content/full/344/21/1594?maxto show=&H...
Neither the indexation of clinical trials nor the reporting in abstracts in PsycLIT was helpful for the reliable identification of randomized trials. With the purpose of minimizing the number of missed randomized trials, any search terms aimed at identifying clinical trials were omitted. In a later filtering process, abstracts were read in full.

Comment:
It was not possible to search for words containing less than three letters in the Cochrane Controlled Trials Register. This made searches for the essential term "no" impossible. The two-phase search strategy described above was therefore implemented.

This article has been cited by other articles:

- Devereaux, P. J., Bhandari, M., Montori, V. M., Manns, B. J., Ghali, W. A., Guyatt, G. H. (2002). Double blind, you are the weakest link -- goodbye!.
- Devereaux, P. J., Bhandari, M., Montori, V. M., Manns, B. J., Ghali, W. A., Guyatt, G. H. (2002). Double blind, you have been voted off the island!.
- Devereaux, P. J., Bhandari, M., Montori, V. M., Manns, B. J., Ghali, W. A., Guyatt, G. H. (2002). "Double blind, you are the weakest link -- goodbye!". Evid Based Nurs 5: 36-37 [Full Text]
- Devereaux, P. J., Bhandari, M., Montori, V. M., Manns, B. J., Ghali, W. A., Guyatt, G. H. (2002). "Double blind, you are the weakest link -- goodbye!". Evid Based Nurs 5: 36-37 [Full Text]
- Devereaux, P. J., Bhandari, M., Montori, V. M., Manns, B. J., Ghali, W. A., Guyatt, G. H. (2002). "Double blind, you are the weakest link -- goodbye!". Evid Based Nurs 5: 36-37 [Full Text]
- Devereaux, P. J., Bhandari, M., Montori, V. M., Manns, B. J., Ghali, W. A., Guyatt, G. H. (2002). "Double blind, you are the weakest link -- goodbye!". Evid Based Nurs 5: 36-37 [Full Text]
- Devereaux, P. J., Bhandari, M., Montori, V. M., Manns, B. J., Ghali, W. A., Guyatt, G. H. (2002). "Double blind, you are the weakest link -- goodbye!". Evid Based Nurs 5: 36-37 [Full Text]
- Devereaux, P. J., Bhandari, M., Montori, V. M., Manns, B. J., Ghali, W. A., Guyatt, G. H. (2002). "Double blind, you are the weakest link -- goodbye!". Evid Based Nurs 5: 36-37 [Full Text]
- Devereaux, P. J., Bhandari, M., Montori, V. M., Manns, B. J., Ghali, W. A., Guyatt, G. H. (2002). "Double blind, you are the weakest link -- goodbye!". Evid Based Nurs 5: 36-37 [Full Text]

