Active Albuterol or Placebo, Sham Acupuncture, or No Intervention in Asthma

Michael E. Wechsler, M.D., John M. Kelley, Ph.D., Ingrid O.E. Boyd, M.P.H., Stefanie Dutile, B.S., Gautham Marigowda, M.B., Irving Kirsch, Ph.D., Elliot Israel, M.D., and Ted J. Kaptchuk

BACKGROUND
In prospective experimental studies in patients with asthma, it is difficult to determine whether responses to placebo differ from the natural course of physiological changes that occur without any intervention. We compared the effects of a bronchodilator, two placebo interventions, and no intervention on outcomes in patients with asthma.

METHODS
In a double-blind, crossover pilot study, we randomly assigned 46 patients with asthma to active treatment with an albuterol inhaler, a placebo inhaler, sham acupuncture, or no intervention. Using a block design, we administered one each of these four interventions in random order during four sequential visits (3 to 7 days apart); this procedure was repeated in two more blocks of visits (for a total of 12 visits by each patient). At each visit, spirometry was performed repeatedly over a period of 2 hours. Maximum forced expiratory volume in 1 second (FEV<sub>1</sub>) was measured and patients’ self-reported improvement ratings were recorded.

RESULTS
Among the 39 patients who completed the study, albuterol resulted in a 20% increase in FEV<sub>1</sub>, as compared with approximately 7% with each of the other three interventions (P<0.001). However, patients’ reports of improvement after the intervention did not differ significantly for the albuterol inhaler (50% improvement), placebo inhaler (45%), or sham acupuncture (46%), but the subjective improvement with all three of these interventions was significantly greater than that with the no-intervention control (21%) (P<0.001).

CONCLUSIONS
Although albuterol, but not the two placebo interventions, improved FEV<sub>1</sub> in these patients with asthma, albuterol provided no incremental benefit with respect to the self-reported outcomes. Placebo effects can be clinically meaningful and can rival the effects of active medication in patients with asthma. However, from a clinical-management and research-design perspective, patient self-reports can be unreliable. An assessment of untreated responses in asthma may be essential in evaluating patient-reported outcomes. (Funded by the National Center for Complementary and Alternative Medicine.)
Placebo effects (i.e., benefits resulting from simulated treatment or the experience of receiving care) are reported to improve signs and symptoms of many diseases in clinical trials and in clinical practice. On this basis, the accepted standards for clinical-trial design specify that the effects of active treatment should ideally be compared with the effects of placebo. Despite this common practice, it is unclear whether placebo effects observed in clinical trials (or those that presumably occur in clinical care) influence both objective and subjective outcomes and whether placebo effects differ from the natural course of disease or regression to the mean.

In patients with asthma, the administration of an inhaled bronchodilator can result in rapid increases in expiratory airflow that can be measured with spirometry. Since repeated lung-function assessments can be performed over short periods of time, asthma is an excellent model for the study of placebo effects. Although many studies suggest that such effects occur in patients with asthma, these studies have generally not controlled for the effects of variability that can occur over the period of observation without treatment.

In this pilot study, we compared acute changes in lung function that occurred after repeated administration of four interventions: a masked bronchodilator (inhaled albuterol), two different types of placebo (an inert inhaler and a validated sham acupuncture needle), and a period of no intervention. By using different placebos and a no-intervention control, we were able to determine whether placebo interventions in asthma can lead to objective changes in airway caliber, self-reported subjective improvements, or both beyond the changes in lung function and symptoms that are attributable to the natural history of the disease.

Methods

Patients and Study Design
Between January 2007 and December 2008, we conducted a randomized, double-blind, crossover pilot study with the use of a block design to determine the acute response to an inhaled bronchodilator in patients with stable asthma. At the initial screening visit, 79 patients completed questionnaires and, having had short-acting bronchodilator therapy withheld for a minimum of 8 hours and long-acting bronchodilator therapy withheld for at least 24 hours, underwent bronchodilator reversibility testing with open-label inhaled albuterol. The 46 patients who had a response, defined as a forced expiratory volume in 1 second (FEV₁) that was at least 12% higher than the baseline value, were eligible to continue in the study (Fig. 1). (Details about the inclusion and exclusion criteria, structure of the study, and study visits can be found in the Supplementary Appendix, available with the full text of this article at NEJM.org.)

These patients returned within a week and were assigned to a randomly ordered series of four interventions — active albuterol inhaler, placebo inhaler, sham acupuncture, or no-intervention control — administered on four separate occasions, 3 to 7 days apart (block 1) (Fig. 2). This procedure was repeated in two more blocks of four visits each (blocks 2 and 3), during which the interventions were again randomly ordered and administered. Thus, each subject received a total of 12 interventions. As before, short-acting and long-acting bronchodilator therapy was withheld for 8 hours and 24 hours, respectively, before each intervention. The no-intervention control condition differs from the natural history of asthma, since it controls for nonspecific factors such as attention from study staff, responses to repeated spirometry, regression to the mean, natural physiological variation, and any effects arising from the hospital setting. Nonetheless, no-intervention controls are the best approximation of no treatment in an experimental design. The study was conducted in accordance with the protocol (available at NEJM.org).

Objective and Subjective Outcomes
At each of the 12 visits, spirometry was used to obtain a baseline measurement of FEV₁, after which patients received the intervention for that particular visit (as randomly assigned within the four visits of that block of visits). Spirometry was then repeated every 20 minutes for 2 hours. Also at each visit, patients were asked to score any perceived improvements in asthma symptoms on a visual-analogue scale, with scores ranging from 0 (no improvement) to 10 (complete improvement), and were also asked whether they thought they had received a genuine therapy or placebo (to assess blinding). These subjective responses were then converted to percent improvement in FEV₁, during the 2 hours by multiplying each score by 10.
Statistical Analysis
Drug and placebo effects were assessed by means of repeated-measures analysis of variance. If significant main effects were found, we compared each intervention with the use of two-tailed, paired t-tests. We used a Bonferroni correction to control type I error, and only those effects with P values of less than 0.008 were considered to be significant. (See the Supplementary Appendix for details.) The magnitudes of the effects were assessed with the use of Cohen’s d statistic, which provides a measure of the differences in the mean values of changes in symptom severity between groups in relation to the pooled standard deviation.12

Results
Patients
Seventy-nine patients were screened, of whom 46 with mild-to-moderate asthma met the entry criteria, gave written informed consent, and were randomly assigned to the study interventions (Fig. 1). The demographic characteristics, baseline spirometric values, and baseline asthma medications are shown in Table 1.

Objective Physiological Outcome
Figure 3 shows the mean physiological responses to each double-blind intervention (albuterol inhaler, placebo inhaler, sham acupuncture, and no intervention) across the three study visits. At the initial screening visit, the mean (±SE) percent improvement in FEV₁ was 21.9±1.6%, and all patients had an improvement in FEV₁ of at least 12%. During the double-blind test series, the mean percent improvement in FEV₁ was 20.1±1.6% with inhaled albuterol, as compared with 7.5±1.0% with inhaled placebo, 7.3±0.8% with sham acupuncture, and 7.1±0.8% with the no-intervention control. There were no significant differences between the three inactive interventions, none of which resulted in the degree of improvement observed with active albuterol. The difference in drug effect between the albuterol inhaler and the placebo inhaler, as indexed by the difference in mean percent improvement in FEV₁, was significant (P<0.001) and large (d = 1.48). In contrast, the placebo effects did not differ significantly between the two placebo interventions and the no-intervention control (P = 0.65 for the comparison of placebo inhaler with no intervention, and P = 0.75 for the comparison of sham acupuncture with no intervention). In addition, the sizes of these effects were negligible (d = 0.07 for placebo inhaler and d = 0.04 for sham acupuncture). With the use of the standard definition of treatment response (≥12% improvement in FEV₁), patients assigned to the active albuterol inhaler had a response 77% of the time, whereas patients assigned to the placebo inhaler, those assigned to sham acupuncture, and those assigned to no intervention had a response 24%, 20%, and 18% of the time, respectively (Table 2 in the Supplementary Appendix).

Subjective (Patient–Reported) Outcome
As shown in Figure 4, patients reported substantial improvement not only with inhaled albuterol (50% improvement) but also with inhaled placebo (45%) and with sham acupuncture (46%). In con-
trast, the improvement reported with no intervention was only 21%. The difference in the subjective drug effect between the active albuterol inhaler and the placebo inhaler was not significant (P = 0.12), and the observed effect size was small (d = 0.21). With respect to the placebo effects, however, the difference between the two placebo interventions and no intervention was large (d = 1.07 for placebo inhaler and d = 1.11 for sham acupuncture) and significant (P < 0.001 for both comparisons). Treatment credibility was high, and most patients believed that they had received active treatment (73% for double-blind albuterol, 66% for double-blind placebo inhaler, and 85% for sham acupuncture). The two double-blind conditions did not differ significantly from each other, but sham acupuncture was significantly more credible than both inhaler conditions (P < 0.05).

**DISCUSSION**

In this repeated-measures pilot study in which active-drug and placebo effects were assessed in patients with asthma, two different types of placebo had no objective bronchodilator effect beyond the improvement that occurred when patients received no intervention of any kind and simply underwent repeated spirometry (no-intervention control). In contrast, the subjective improvement in asthma symptoms with both inhaled placebo and sham acupuncture was significantly greater than the subjective improvement with the no-intervention control and was similar to that with the active drug. Thus, even though there was a large, objective drug effect (mean percent improvement in FEV₁, 20%) that was nearly three times the effect of the two placebos and the no-intervention control (mean percent improvement in FEV₁, approximately 7% for all three), patients could not reliably detect the difference between this robust effect of the

---

**Table 1. Demographic and Clinical Characteristics of the 46 Patients with Asthma.**

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Demographic</strong></td>
<td></td>
</tr>
<tr>
<td>Age</td>
<td>41.5±17.0</td>
</tr>
<tr>
<td>Female sex (%)</td>
<td>80</td>
</tr>
<tr>
<td>Race or ethnic group (%)†</td>
<td></td>
</tr>
<tr>
<td>White</td>
<td>62</td>
</tr>
<tr>
<td>Black</td>
<td>23</td>
</tr>
<tr>
<td>Other</td>
<td>15</td>
</tr>
<tr>
<td><strong>Clinical</strong></td>
<td></td>
</tr>
<tr>
<td>FEV₁ at baseline</td>
<td>2.1±0.8</td>
</tr>
<tr>
<td>FEV₁ (% of predicted)</td>
<td>67.0±16.9</td>
</tr>
<tr>
<td>Improvement in FEV₁ (%)</td>
<td>21.9±9.9</td>
</tr>
<tr>
<td>Baseline score on questionnaire‡</td>
<td>1.8</td>
</tr>
<tr>
<td>Asthma maintenance medication (% of patients)</td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>49</td>
</tr>
<tr>
<td>Inhaled glucocorticoid only</td>
<td>23</td>
</tr>
<tr>
<td>LABA</td>
<td>0</td>
</tr>
<tr>
<td>Leukotriene modifier only</td>
<td>0</td>
</tr>
<tr>
<td>Inhaled glucocorticoid and LABA</td>
<td>13</td>
</tr>
<tr>
<td>Inhaled glucocorticoid and leukotriene modifier</td>
<td>2</td>
</tr>
<tr>
<td>Inhaled glucocorticoid, LABA, and leukotriene modifier</td>
<td>13</td>
</tr>
</tbody>
</table>

* Plus–minus values are means ±SD. FEV₁ denotes forced expiratory volume in 1 second, and LABA long-acting beta agonist.
† Race or ethnic group was self-reported.
‡ The scores on the Asthma Control Questionnaire ranged from 0 to 10, with 0 indicating no improvement and 10 indicating complete improvement.

---

**Figure 2.** Screening and Randomization.
active drug and the effects of inhaled placebo and sham acupuncture (mean subjective improvement reported by patients in all groups ranged between 45 and 50%).

For the objective physiological outcome (change in FEV₁), there was a powerful medication effect (drug vs. placebo) but no placebo effect (no difference between placebo and the no-intervention control). For the subjective outcome, the placebo effects were equivalent to the drug effect, and all were greater than the no-intervention effect. The two placebo interventions had a strong effect on the patient-reported outcome but had no effect on the objective outcome; the active drug had a strong effect on the objective outcome but had no incremental benefit with respect to the subjective outcome.

Most randomized, controlled trials and laboratory experiments have not included a no-intervention control. Our inclusion of a no-intervention control — the control for the placebos — allowed us to detect subjective placebo effects. We found that the results of placebo interventions did not differ from those of the no-intervention control when an objective measure of airflow was used (FEV₁). However, for the subjective outcome, both placebos had a greater effect than no intervention. This may have been due to the effect of the patient’s expectation on the patient-reported outcome or to reporting bias (e.g., the wish to please the investigator). We consider the latter influence unlikely because the patients in the no-intervention group also reported subjective improvement, even though they presumably had no expectation of improvement and their wish to please the investigator would have made a report of no improvement more likely. Our findings might have been influenced by possible weaknesses in the scale used to assess subjective responses (which lacked prior formal validation). However, it is unlikely that the findings were due to the instrument used, since the no-intervention group did not show such an effect.

The subjective responses to placebo were equivalent to the subjective responses to the active drug, even though the active drug produced a marked increase in FEV₁. Thus, the administration of a placebo did not affect the objective measure (placebo as compared with the natural history of asthma), and the effect of the active medication did not exceed that of the ritual of the treatment itself (albuterol as compared with either placebo). The fact that the patient-reported outcome was independent of the physiological outcome suggests either that patients with asthma poorly perceive changes in FEV₁ or that use of subjective assessment may have some limitations in the interpretation of physiological outcomes in asthma and may have upper limits, possibly explaining why asthma symptoms in many patients remain uncontrolled.

It is notable that the two placebos had similar effects on both the objective measure and the subjective measure. Since all the patients had prior experience with active inhalers, one might have expected better outcomes with the placebo inhaler than with sham acupuncture, owing to classical conditioning. One possible reason for the apparent equivalence of the two placebo interventions is that the patients may have become conditioned to the setting and personnel at a well-known hospital as much as to the inhaler itself. Another possible explanation is that the remarkably high credibility of sham acupuncture in our study (85%, vs. 66% for the placebo inhaler), which is consistent with the findings in previous studies, might have resulted in a greater expectation of improvement with sham acupuncture.

Our findings complement the results of a re-
A randomized, controlled trial that examined the effects of optimistic drug presentation (enhanced positive expectations) on outcomes with placebo or active medication (montelukast) in 610 patients with asthma: Placebo given with enhanced expectations significantly increased subjective outcomes but had no effect on objective measures, whereas enhanced expectations for medication influenced neither subjective nor objective outcomes. Although another study of asthma reported objective improvement with placebo, it lacked a no-intervention comparison, so it is not known whether the reported improvement reflected an actual effect of placebo or simply the natural history of asthma. Our findings strongly contrast with a series of studies in which placebo interventions plus strong suggestion resulted in marked changes in FEV₁ in patients with asthma. In these studies, the patients were deceptively told that they were receiving “powerful” medication, whereas our study was conducted with double-blind instructions. Most of these other studies lacked no-intervention controls, and the two studies that included them showed no placebo effect.

Although placebo effects may differ according to the specific disease, our study has implications for understanding placebo effects in general. Our findings are consistent with those of a meta-analysis involving multiple conditions, in which the placebos, as compared with no-intervention controls, had no significant effect on objective measures but did have significant effects on subjective outcomes (e.g., pain). Also, our data support recent systematic reviews of studies that involved specific conditions (which, unfortunately, lacked natural-history controls) suggesting that placebo effects are primarily detectable in subjective outcomes; when objective changes occur, they tend to be well within the range of the natural history of the condition. Furthermore, our findings do not contradict recent laboratory studies showing that placebo treatment elicits quantifiable changes in neurotransmitters and regionally specific brain activity that influence symptoms. The bifurcation of placebo effects between objective and subjective outcomes that we observed in this pilot study may represent the distinction that social scientists make between treating disease (objective physiology) and treating illness (subjective perceptions). Although effective medications target and modulate objective biologic features, the mere ritual of treatment may affect patients’ self-monitoring and subjective experience of their disease.

Our subjective measure deserves comment. Since there were no preexisting subjective measures for the acute asthma response, we constructed our own metric for global subjective assessment of improvement in dyspnea; as a result, its reliability and validity can be questioned. However, our measure had good face validity. Even though similar measures are common in medicine and have been shown to have good reliability and validity (e.g., the Borg scale, which is used to assess dyspnea), none have been validated for use in assessing either asthma or the acute bronchodilator response. In addition, patients used the entire range of the measure, and there were no ceiling or floor effects. The broad range of responses and roughly normal distribution argue against the existence of strong acquiescence (tendency to agree regardless of the content of a question) or central-tendency biases. Our subjective scale did not encompass worsening of symptoms (i.e., the scale measured improvement, from none to complete); thus, it could indicate a perceived lack of improve-
ment but not a perceived worsening. This limitation could potentially create a floor effect and underestimate the degree of subjective deterioration for some patients. However, there was no floor effect observed in the distribution of assessment scores for the active or placebo interventions. In contrast, there was, as expected, a floor effect with the no-intervention control, since patients assigned to this control overwhelmingly reported no improvement. This floor effect serves to strengthen our findings concerning the discrepancy between subjective and objective outcomes. It does so because the only significant difference with respect to the subjective outcome was the lower degree of improvement in the no-intervention condition as compared with the other three conditions. Thus, the floor effect for the no-intervention control would, if anything, have served to diminish the difference between no intervention and the active and placebo interventions and would have decreased our ability to detect such a difference.

This study has several other limitations. First, we studied acute asthmatic responses, so it remains unclear whether our findings would apply to chronic asthma or to other conditions. Even with respect to the treatment of acute asthma, it is important that the findings from our study be replicated to assess their reliability and robustness. In addition, we measured outcomes using a single subjective measure and a single objective measure (FEV₁). Future research should investigate whether our findings can be generalized to other subjective and objective measures of acute asthma. Finally, we did not assess subjective symptoms before each visit’s intervention; therefore, the severity of subjective symptoms before each treatment remains unclear. Assessing subjective measurements before and after interventions could have yielded other differences. Although it is possible that the degree of physiological deficit in these patients was not sufficient for them to have symptoms at rest, it is increasingly recognized that not all patients with asthma who have deficits in lung function fully appreciate the degree to which their asthma limits airflow until they are given bronchodilators that result in improvement in lung function, symptoms, or both. In this study, there was a significant improvement in lung function with the genuine bronchodilator (about 20%) that coincided with an improvement in symptoms, whereas treatment with placebo had no effect on measurable biologic factors but was indistinguishable from medication with regard to subjective outcomes.

Our research has important implications both for the treatment of asthma and for clinical-trial design in general. Many patients with asthma have symptoms that remain uncontrolled, and the discrepancy between objective pulmonary function and patients’ self-reports noted in this study suggests that subjective improvement in asthma should be interpreted with caution and that objective outcomes should be more heavily relied on for optimal asthma care. Indeed, although improvement in objective measures of lung function would be expected to correlate with subjective measures, our study suggests that in clinical trials, reliance solely on subjective outcomes may be inherently unreliable, since they may be significantly influenced by placebo effects. However, even though objective physiological measures (e.g., FEV₁) are important, other outcomes such as emergency room visits and quality-of-life metrics may be more clinically relevant to patients and physicians. Although placebos remain an essential component of clinical trials to validate objective findings, assessment of the course of the disease without treatment, if medically appropriate, is essential in the evaluation of patient-reported outcomes.

Supported by grants (R21-AT002793-01 and K24-AT004095) from the National Center for Complementary and Alternative Medicine, National Institutes of Health.

Dr. Wechsler reports receiving fees for consulting and advisory-board membership from GlaxoSmithKline, Merck, Schering-Plough, AsthmaX, Genentech, Sunovion (formerly Sepracor), Cephalon, MedImmune, NKT Therapeutics, and Boehringer Ingelheim; lecture fees from Merck, Novartis, AstraZeneca, and GlaxoSmithKline; and grant support from GlaxoSmithKline and AsthmaX. Dr. Israel reports receiving fees for consulting and advisory-board membership from Abbott, Amgen, Astellas Pharma US, BI Worldwide, Cowen, Gilead Sciences, GlaxoSmithKline, Icagen, Johnson & Johnson, MedImmune, Merck, newMentor, NKT Therapeutics, Novartis, Ono Pharmaceutical, PDL BioPharma, Pulmatrix, Schering-Plough, Sunovion (formerly Sepracor), and Teva Pharmaceuticals; lecture fees from Merck; fees for expert testimony from Beksm and Conley; Prince, Lobel, Glovsky, and Iye; Campbell, Campbell, Edwards, and Conroy; Diedrich and Donohue; Sulloway and Hollis; and Ryan Ryan Deluca; and grant support from Aerovance, Amgen, Biota, Ception, Genentech, Icagen, Johnson & Johnson, MedImmune, and Novartis. No other potential conflict of interest relevant to this article was reported.

Disclosure forms provided by the authors are available with the full text of this article at NEJM.org.

We thank Dr. Aaron Deykin for assistance in protocol development; and the following clinical research coordinators, who assisted in study procedures and recruitment: Erika Line-Nitu, Lawrence Mollo, Jared Preston, Suzanne Vogt, Crystal Wesockes, and Emily Zarookian.
REFERENCES


Copyright © 2011 Massachusetts Medical Society.