Available online at www.sciencedirect.com

SCIENCE DIRECT®

ELSEVIER

Contemporary Clinical Trials xx (2005) xxx–xxx

Contemporary Clinical Trials

www.elsevier.com/locate/conclintrial

Investigating placebo effects in irritable bowel syndrome: A novel research design

Lisa A. Conboy [a], Rachel H. Wasserman [a], Eric E. Jacobson [b], Roger B. Davis [c], Anna T.R. Legedza [c], Min Park [a], Andrea L. Rivers [a], Elizabeth B. Morey [a], Bong Hyun Nam [a], Louis Lasagna [d,*], Irving Kirsch [e], Anthony J. Lembo [f], Ted J. Kaptchuk [a], Catherine E. Kerr [a,*]

[a] *Osher Institute, Division for Research and Education in Complementary and Integrative Therapies, Harvard Medical School, 401 Park Drive, Suite 22A, Boston, MA 02115, United States*

[b] *Department of Social Medicine, Harvard Medical School, Boston, MA, United States*

[c] *Division of General Medicine and Primary Care, Beth Israel Deaconess Medical Center, Boston, MA, United States*

[d] *Sackler Graduate School for Biomedical Sciences, Tufts University, Boston, MA, United States*

[e] *School of Applied Psychosocial Studies, University of Plymouth, Plymouth, UK*

[f] *Department of Gastroenterology, Beth Israel Deaconess Medical Center, Boston, MA, United States*

Received 31 March 2005; accepted 22 November 2005

Abstract

Little is known about placebo effects with scientific precision. Poor methodology has confounded our understanding of the magnitude and even the existence of the placebo effect. Investigating placebo effects presents special research challenges including: the design of appropriate controls for studying placebo effects including separating such effects from natural history and regression to the mean, the need for large sample sizes to capture expected small effects, and the need to understand such potential effects from a patient's perspective. This article summarizes the methodology of an ongoing NIH-funded randomized controlled trial aimed at investigating whether the placebo effect in irritable bowel syndrome (IBS) exists and whether the magnitude of such an effect can be manipulated to vary in a manner analogous to "dose dependence." The trial also uses an innovative combination of quantitative and qualitative methods.

Keywords: Placebo effect; Randomized clinical trial; Research methodology; Irritable bowel syndrome; Placebo run-in; Patient–practitioner relationship; Qualitative methods

1. Introduction

Controversy and debate have characterized recent scientific considerations of placebo effects [1]. A well-performed meta-analysis produced evidence that placebo effects have little clinical significance and may not even exist [2]. Yet,

* Corresponding author. Tel.: +1 617 384 8583; fax: +1 617 384 8580.
E-mail address: Catherine_Kerr@hms.harvard.edu (C.E. Kerr).

* Deceased.

1551-7144/$ - see front matter © 2005 Elsevier Inc. All rights reserved.
doi:10.1016/j.cct.2005.11.012

several sophisticated mechanistic studies, including neuro-imaging and neuro-endocrine studies, have found clear evidence for a placebo effect operating through several different mechanistic pathways. This evidence, however, is not definitive because it is gathered in experimental settings over a short time span, of hours, or at the most, several days [3,4]. One of the biggest problems concerning knowledge of placebo effects is confusion over definition and conceptualization. Is the placebo effect the outcome of the patient–practitioner interaction and the entire medical context or is it simply the result of a bogus treatment that mimics an active treatment? Researchers have recently described these two distinct frameworks of placebo as a *broad* definition and a *narrow* definition [5,6]. Understandably, a recent trans-NIH conference has called placebo research an urgent priority [7–9].

While mechanistic research on placebo has progressed steadily, clinical research into placebo effects has, until recently, been hampered by a lack of rigorous research design and even a general lack of interest. Attention in the research community, however, has increased dramatically as pharmaceutical companies have become more concerned with the difficulty of separating drug effects from large effect sizes observed in the placebo arms of randomized controlled trials (RCTs) [10,11]. A methodological difficulty encountered in clinical research on placebo is that of confounding the outcome of the placebo control in a RCT with what is clinically called the "placebo effect." The placebo outcome in a RCT can include a possible placebo effect–the broad effects such as practitioner attention and/or the narrow effects of a dummy treatment–but may also include natural history, natural variability of the illness and regression to the mean [1]. Progress in clarification of these questions, however, poses serious methodological challenges including: the design of appropriate controls, the design of trials with sufficient sample size to be able to detect potentially small placebo effects, the development of methods to investigate the experience of a placebo intervention from a patient's point of view.

This paper describes the methodology of an ongoing NIH funded RCT titled "Enhancing the placebo effect in irritable bowel syndrome" specifically designed to examine whether a placebo effect exists beyond natural history in this patient population when given a placebo treatment, and whether this effect (if it exists) can be enhanced by a combination of positive cognitive, emotional and symbolic patient–practitioner interactions. The trial is designed to simultaneously compare a "broad" versus a "narrow" placebo intervention with a "wait-list" control, in which

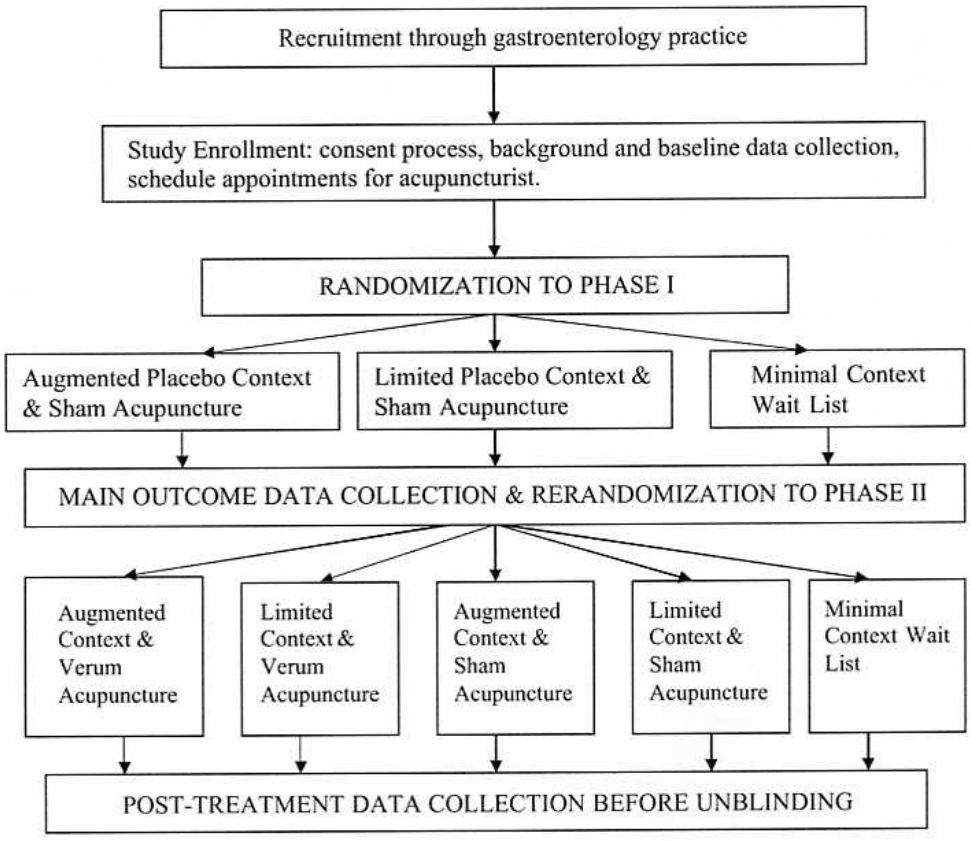

Fig. 1. Flow chart of experimental design.

patients enter the trial but receive no treatment until the trial's conclusion. The addition of a wait-list condition allows us to control for the effects of natural history, including routine improvement that may be due to the natural course of the disease. The trial also examines placebo responses in a clinical population over a much longer period of time than previous studies. The primary aim of this RCT is to determine whether placebo effects across two sham treatment arms and a no treatment arm can be modulated so as to behave monotonically in a manner analogous to "dose dependence." The paper also discusses an innovative combination of quantitative and qualitative methods aimed at understanding the patient's experience of being in a clinical trial and receiving a placebo intervention (Fig. 1).

2. Methodological difficulties with studying placebo effect

Despite the urgent need to demonstrate whether clinically significant placebo effects exist and whether they can be augmented by certain types of patient–practitioner interactions or are limited to the expectation associated with receiving a dummy treatment, there are few research studies addressing these questions. Contributing to the absence of well-designed placebo studies is the fact that the execution of a placebo focused RCT presents several inherent challenges. The three major problems we address in this paper are: (1) the modeling of appropriate controls, (2) dealing with the likelihood of relatively small effect sizes, and (3) understanding the patient's experience of being in an RCT and receiving a placebo treatment. In this trial we have developed innovative methods for creating appropriate controls and generating sufficient sample size to combat the relatively small effect size of placebo interventions as opposed to classical pharmacological effects.

2.1. The modeling of appropriate controls

Historically, most, if not all, clinical studies of the placebo effect have not paid adequate attention to the roles that regression to the mean and natural history may play in patient outcomes. In a RCT in particular, where a placebo treatment is given as a control for an "active" intervention, the effect size of the placebo treatment is, in fact, an aggregate of the placebo effect elicited by the patient–practitioner interaction and surrounding treatment context *and* regression to the mean/natural history. In order to accurately determine the size and direction of a placebo effect, these elements need to be studied as factors themselves. Natural history/regression to the mean can have a dramatic impact on outcomes in many illnesses including IBS [12–17]. In our trial, we have controlled for these factors by including a no treatment group (i.e., wait-list). Ideally, the most valid way to separate natural history from the effects of study enrollment status would be to include a fourth arm of patients uninformed that they are in a RCT and follow them without their knowledge. For ethical reasons, this design is untenable [18]. We have settled for the only ethically possible comparison—a no treatment arm. It should be noted, however, that patients in a natural history or no treatment group also receive increased clinical attention; for example, Hawthorne effects and expressions of support may accompany the mere process of entering a study and contribute to outcome [18–23]. In this trial we are using the wait list condition as an appropriate scientific and ethical means to control for natural history.

Our trial also includes two different placebo treatment arms, each of which receives a different "dose" of ritual, contextual factors to see whether these factors can be delivered in a manner analogous to dose-dependence. These two groups will also allow us to disentangle the relative influence of the broad sense of placebo (as a product of the patient–practitioner interaction) versus the narrow sense of placebo (as a dummy treatment being administered in a RCT).

By regarding the wait list group as an approximate indicator of the most minimal effects of study participation and attention, our design tests whether the context effects embedded in our three arms behave in a manner analogous to dose-dependence. The three groups are (1) 'augmented' patient–practitioner interaction + dummy placebo, (2) 'limited' patient–practitioner interaction + dummy placebo, (3) wait-list.

2.2. The problem of statistical power

A second major challenge for placebo research using a RCT is finding statistically significant results. This challenge has two interrelated components. First, the effect size for a placebo treatment is generally smaller than an active drug; achieving sufficient power to detect an effect is difficult. Second, in a traditional RCT design, part of the sample will

receive a verum treatment at all points in the trial. If one is studying placebo, placing a portion of one's sample size in verum treatment reduces sample size and resulting power. We have two solutions to combat these problems with statistical power: our choice of study population, and our choice of study design.

We have chosen IBS as a target disease because, in addition to its clinical importance, we believe that this disorder will give us sufficient power to test our hypotheses. The placebo effect in IBS, like other chronic conditions such as chronic pain and depression, in which subjective self-monitoring, selective attention and psycho-social interactions have been observed to play a critical role in symptom expression, is reported to be especially large. A recent systematic review of RCTs for IBS found global improvement of 41% (range 7–75%) in placebo controls [24,25]. Some investigators have reported consistent effects on the order of 90% [26]. A recently published placebo controlled RCT in IBS performed by our team had a placebo response on the order of magnitude of 57% [27]. Another recently completed placebo controlled RCT in functional constipation, which has significant overlap with IBS, had a placebo response rate of approximately 45% [28]. In addition, other placebo controlled RCTs in IBS, which we have participated in, have had placebo response rates between approximately 40% [29] and 50% [30]. Obviously, these rates are likely to contain natural history and regression to the mean, but the average length of illness for our studies has been about 10 years which makes it appear unlikely to be entirely natural history. Precisely for this reason, our study, unlike previous studies, will include a natural history no treatment arm, which allows for calculation of the magnitude of placebo effect as separate from natural history and regression to the mean.

Because there are no biological markers in IBS, the severity of the disease is defined by subjective factors. While in most scientific experiments this is a handicap, it may be a benefit for studying the possibility of dose-dependent placebo effect. As a recent Lancet article noted, "context-effects are in all probability small" and "the more clear-cut [objective] the disease studied, the smaller the effects are likely to be" [31]. In conditions such as IBS, depression, and pain, that are primarily subjectively defined, the mechanism of selective attention probably plays a significant modulating role in placebo phenomena [6,32,33]. The greater the subjectivity the more room for the perception and attribution of improvement due to treatment [34]. As one review has noted, placebo effects are most likely to be high in IBS and other disorders where symptoms "have a fluctuating course" that rooted, at least partially, in the patient's subjective experience of illness [35].

Second, our application of the traditional run-in design allows us to obtain relevant placebo effect data from every person enrolled in the study while still offering a customary and ethically appropriate chance of verum treatment. Traditionally, the placebo run-in is a research design that places all participants in a RCT in single-blind placebo treatment for between 1 and 3 weeks [36–38] before randomization to verum or continuation of placebo treatment. The goal of this procedure is to monitor placebo responses before randomization, and to only randomize those subjects who do not 'respond' in the placebo run-in to the main phase of the study. Sometimes the technique is used as a check for adherence before randomization. This technique of removing initial placebo responders has traditionally been thought to increase the usability of the post-run-in sample by decreasing the magnitude of eventual overall placebo response in the sample and making detection of genuine versus placebo treatment more efficient [36,38]. Recently, there has been debate whether this practice actually reduces overall placebo response in RCTs. Research on this has offered some positive evidence (e.g. [39,40]) and some negative(e.g. [41–44]). Debate has also occurred as to whether such run-ins affect the external validity of RCTs [45].

In our IBS trial, we have adopted the run-in for entirely different reasons than previous RCT researchers. Our primary aim is to investigate the effects of different kinds of placebos, and the run-in design maximizes our ability to detect differences in the effects of different placebos by allowing us to use the entire sample during the run-in phase. This aim is situated within the context of a RCT of acupuncture for IBS. The placebo effect is studied using a validated sham needle device that has an identical appearance and sensation to real acupuncture but instead of penetrating the skin retracts up the shaft of the needle like a retractable knife [46–49].

The run-in phase (Phase I of our two phase study) thus maximizes our statistical power. Use of the run-in also allows us to assess several secondary questions including: does dropping placebo responders from the analysis after the run-in phase affect the verum-placebo difference in Phase II? After the run in, we re-randomize the sample so that half of those patients receive real acupuncture. This Phase II re-randomization confers three benefits: it allows us to continue collection of placebo data from those subjects who are randomized to continue on placebo, so that we can understand the placebo over a longer duration. It allows us to offer the possibility of a real treatment to the entire sample as is required by the ethical norms of clinical research. As a further benefit of this design, we will also be able to conduct a preliminary assessment of acupuncture's efficacy as a treatment for IBS.

3. Trial design and methods

Below we describe the specific methods we used to address the three problems outlined above: designing appropriate controls, maximizing statistical power for placebo comparisons, and the need to understand the patient's perspective on such potential effects. We also discuss general considerations about recruitment, eligibility and treatment implementation.

Our three-armed RCT is designed to test whether the placebo effect can be enhanced, and if so in a predictable manner. We designed three study arms to provide varying amounts and qualities of interaction with study staff. The groups are (1) a patient–practitioner interaction augmented with positive cognitive, emotional and symbolic interactions such as 'holistic' intake, positive expectation, warm-friendly manner, empathy, attentive listening, touch, thoughtful shared silence plus sham needle treatment; (2) a 'limited' business-like patient–practitioner interaction plus the identical sham needle treatment given in group 1; and (3) wait-list no treatment control. With this design we can consider if the 'non-specific' effects across two different placebo arms and a wait list arm behave monotonically in a manner analogous to 'dose dependence,' i.e. is 'augmented' arm (A)>'limited' arm (L)>wait-list no treatment (W)? Or, is A>L; L>W; A>W?

Our recruitment began in 2004 and is expected to be completed in early 2006 yielding a sample of 287 subjects randomized equally to all three arms. The 6-week subject participation time includes Phase I, in which all subjects in a treatment group receive "acupuncture" with a sham acupuncture device. Here we are comparing the effects of three different levels or 'doses' of the patient–practitioner relationship context, operationalized in a sham acupuncture treatment. Phase II begins with a re-randomization of the treatment arm subjects to continue receiving sham acupuncture or to start receiving verum acupuncture; subjects continue in their originally assigned patient–practitioner context group. Subjects, research assistants and nurse assessors are blinded to subjects' group placement in both phases and the seamlessness of our transition from sham to verum is assessed through the use of a blinding questionnaire asking patients to guess which treatment they are receiving. A research assistant, not involved in measurement collection, completes the re-randomization and contacts the treating acupuncturist.

3.1. Inclusion

We diagnosed IBS based on the Rome II criteria published in 1999 [50], and the exclusion of organic disease by evaluating for 'warning signs'. Use of this approach has been associated with a specificity of >99% for diagnosing IBS [51] and includes the collection of a standard history and physical evaluation for presence of the alarm features of rectal bleeding, weight loss, recent change in symptoms, and family history of colon cancer or inflammatory bowel disease. The Rome II criteria require the presence of abdominal pain or discomfort to be present for 12 weeks or more, which need not be consecutive, in the past 12 months that has at least 2 of the following 3 features: (1) relieved with defecation; (2) onset associated with a change in frequency of stool; and (3) onset associated with a change in form (appearance) of stool. Table 1 offers more specific information.

3.2. Interventions

Our interventions in the "augmented" and "limited" context arms are designed to maximize the difference between two placebo treatment conditions in order to address the hypothesis that placebo effects are modifiable and can be administered in a manner analogous to dose response. To do this we randomized patients into one of three possible groups: augmented, limited, and wait-list.

3.2.1. Augmented

The 'augmented' arm is constructed to provide a broad array of the cognitive, emotional, and symbolic factors thought to enhance the patient–practitioner relationship and presumably the placebo effect. It has all the elements of a realistic 'holistic' acupuncture intake. The components of our 'augmented' arm are derived from a review of placebo effects in clinical trials and the related literature. In particular we have utilized the results of Di Blasi's systematic review of the last 50 years of RCTs concerning the effects of context on verum and placebo effects [52].

Table 1
Inclusion and exclusion criteria used to determine eligibility for participation

Inclusion

1. Provide signed and dated informed consent and understand the nature of the study sufficiently to allow completion of all study assessments.
2. Be ambulatory, community dwelling, 18–80 years, inclusive.
3. Meet ROME II diagnostic criteria for IBS, specifically: At least 12 weeks or more, which may be consecutive, in the preceding 12 months, of abdominal discomfort or pain that has two of the following three features: (a) relieved with defecation; (b) onset associated with change in frequency of stool; (c) onset associated with a change in form (appearance) of stool.
4. Have IBS of at least "moderate" severity score at screening visit (based on the IBS Severity Score, IBSSS).
5. If the patient is on medications which affect the gastrointestinal tract or visceral sensation (e.g., tricyclic antidepressants, opioids), they must be on a stable dose for at least 1 month prior to entering the study and for the duration of the study.

Exclusion

1. Have a history of severe or intractable IBS, defined as continuous, unremitting and severe abdominal pain (>12 h per day).
2. Have history of acupuncture treatment.
3. Be pregnant or lactating.
4. Score >150 on the IBSSS.
5. Have an established diagnosis of any concomitant bowel disturbance that would interfere with the assessment of efficacy or safety in the study, e.g., Hirschsprung's disease, Chagas' disease, lactose, fructose or sorbitol intolerance or caffeine sensitivity.
6. Have a history of laxative abuse.
7. Have undergone previous abdominal surgery (with the exception of uncomplicated appendectomy, hysterectomy, or cholecystectomy >6 months prior to enrollment).
8. Have a history of metabolic or inflammatory disease that may affect bowel motility, e.g., inflammatory bowel disease, diabetes mellitus, sarcoidosis, connective tissue disease, amyloidosis, or poorly controlled hypo/hyperthyroidism.
9. Have a history of significant concomitant psychiatric, neurological, metabolic, hepatic, renal, infectious, hematological, cardiovascular, gastrointestinal, or pulmonary illness. If there is a history of such disease but the condition has been stable for more than 1 year and is judged by the physician/investigator not to interfere with the patient's participation in the study, the patient may be included. Staff will document such cases.
10. Have a history of drug or alcohol abuse within 2 years of entry into the study or test positive for opiates on the initial visit drug screen.
11. Exhibit abnormalities on physical examination, or have abnormal vital signs, or clinical laboratory values unless these abnormalities are judged to be clinically insignificant by the physician/investigator. Such cases will be noted.
12. Be unable or unwilling to cooperate with the study protocol or considered by the Investigator to be unsuitable for the study.
13. Have insufficient knowledge of English to be able to participate in the study.

The principal difference between the augmented and limited treatment conditions comes in the first session in which augmented patients receive an initial 45–60 min in-depth "holistic" intake interview. In this initial intake interview, patient and practitioner discuss the patient's physical symptoms, emotional experiences with illness, lifestyle and more general thoughts about what the illness "means" to the patient. As part of the intake, we have implemented a series of naturalistic interactions which the acupuncturist carries out at key moments in the interview, including: definite communications of empathy, behaviors that demonstrate active listening, the conveyance of a warm and friendly manner, the delivery of a clearly defined period of thoughtful shared silence, and an expression of confidence in the treatment. After the holistic intake, as part of the first session, augmented patients receive a brief (+/−20 min) treatment with sham acupuncture. The first session, which establishes the patient–practitioner context, is followed by five shorter follow-up sham treatment sessions, each lasting 20 min, for a total of six treatments over a 3-week period.

Originally we thought that it would be more scientific to "isolate" different components of our augmented arm in order to look for a more refined proximal causality. After much discussion, our team felt that demonstrating whether such effects can be elicited at all, through the use of a broad-based amalgam of patient–practitioner intervention, was a more critical scientific question at this time than trying to isolate the particular components.

The contextual factors being studied in the augmented condition are not specific to any single therapeutic approach. We have included only factors that are broadly believed to enhance the patient's experience of receiving treatment. These factors, summarized by DiBlasi, *explicitly do not include elements of specific behavioral interventions* which may have efficacy in IBS. These interventions include relaxation-based therapies [53], cognitive behavioral therapy [54] and/or educational interventions [55]. While these behavioral therapies, including cognitive-behavioral psychotherapy—an especially promising treatment for IBS, are defined by their specific therapeutic approaches to the problem of IBS (e.g., "training in illness-related cognitive coping strategies" [56]), our augmented patient–practitioner package includes factors such as empathy, warmth, and attentive listening that are common to what is

considered to be good practice in a range of medical and behavioral specialties [52,57]. We chose to load our patient–practitioner amalgam with these common, non-specific practices in order to enhance the generalizability of our findings.

3.2.2. Limited

The 'limited' arm provides a minimal business-like patient–practitioner interaction, and attempts to isolate the impact of a sham treatment from other context effects. In the initial limited interview, patients meet with the acupuncturist for not more than 10 min at intake during which time the acupuncturist tells the subject that the patient–practitioner interaction must to be kept to a minimum because of the scientific nature of a RCT. Subsequent treatment sessions of the 'limited' patient–practitioner interaction context are performed within the limits that define this first session so that the context continues to be restricted. The sham acupuncture treatment (and later verum acupuncture treatment in Phase II) takes another 20 min and will involve no additional communication between practitioner and patient other than that required to safely and convincingly administer the sham or verum acupuncture treatment. Identical to the augmented arm, as follow up, patients will receive five 20-min sessions of sham acupuncture treatment.

Time is a consideration for placebo effects. In our study, over the course of the first three weeks, the limited arm receives 28% less contact time with the practitioner than does the augmented arm (the total duration of time contact over the first 3 weeks is approximately 152.5 min in the 'augmented' arm and 110 min in the 'limited arm'). Over the course of 6 weeks, the limited arm receives approximately 16% less contact time (augmented receives 272.5 min of contact, limited receives 230 min). While we believe a difference between the augmented and limited arms is unlikely to be accounted for entirely by the difference in total time spent with a practitioner, we will conduct a post-hoc analysis of the effect of time by looking at whether, across both arms, time spent in the initial session is correlated with outcome. This analysis will be used to inform our follow-up study assessing the efficacy of different components, including the role of time, in our augmented patient–practitioner context.

3.2.3. Wait list

Our third arm is a wait-list, which allows us to control for the effects of natural history, regression to the mean, and any effects of attention consequent to enrolling and waiting for a RCT to begin. These subjects know that their treatment will begin 6 weeks after study enrollment. As an incentive for retention, all wait-list subjects (and participants who received only placebo treatment) are eligible to receive 6 courtesy acupuncture treatments after Phase II.

3.3. Implementation and fidelity of treatment

Because we are primarily manipulating the patient practitioner context surrounding acupuncture treatment, we felt it was important to use clinicians who are experienced in both acupuncture and RCT research. All have previous experience in multiple RCTs and understand the importance of maintaining fidelity to protocol and concealment.

Training in the 'augmented' and 'limited' context interventions proceeded according to the methodology used by Dr. Elvira Lang et al. [58,59] which has resulted in high level of adherence to protocol. A training manual was created which included the theories which inform the interventions, and a detailed list of prescribed and proscribed behavioral components. Further, investigators created a training video of a model session for each context, following the methods of a training tape that Dr. Lang has developed in many studies. In addition to the training video and training manual, training includes monthly supervision in which actual taped sessions have been used to discuss acupuncturists' experiences implementing different doses of interaction.

To monitor our implementation we follow recommendations by Moncher and Prinz [60] for achieving and monitoring fidelity of treatment in psycho-social and behavioral interventions. All interactions between patients and acupuncturists are videotaped and a random sample of 10% of each acupuncturist's videotapes is independently rated for compliance by two research assistants who are not otherwise connected with the trial. These two raters have been trained to recognize the various specific behaviors outlined in the Treatment Manual. The raters work to achieve an inter-rater reliability of 0.70 intra-class correlation [61]. (The inter-rater reliability achieved following this procedure in Dr. Lang's last RCT was 0.81.) Adherence is continually monitored and any corrective feedback is given to staff members during monthly supervisory meetings.

3.4. Considering patients' experiences

While patient expectation of benefit is thought to be a strong factor in eliciting placebo response, to date there have been few efforts to prospectively study the processes by which patients develop and/or change their expectations of treatment during the course of a clinical trial. Indeed, there is suggestive evidence that patients' perceptions of the treatment context (e.g., their regard for the practitioner, their reaction to the clinical setting and to the therapeutic procedures) may vary significantly over the course of the clinical trial [62]. This variability may provide some explanation of the prolonged failure of placebo researchers to identify either trait or initial state variables that predict placebo response [63].

To address this gap in the literature, one of our aims in this study is to investigate qualitatively how IBS patients describe their expectations of treatment over the course of the trial across the three different trial conditions. This line of inquiry builds on an earlier qualitative study by our group in which we found that some clinical trial participants used complex language to describe their clinical trial experiences including their experiences forming expectations about treatment assignment and success of treatment. For our parallel qualitative study, we are randomly selecting 27 additional subjects who will be distributed across the augments, limited and wait list conditions and will undergo the exact same treatment protocol as patients in the larger trial, except that they will receive three in-depth interviews, at baseline, midpoint and the conclusion of the study. These interviews will be carried out by our team of medical sociologists and anthropologists and will explore the relationship between subjects' interpretations and experiences of IBS and their subsequent experiences with treatment. This parallel qualitative study has been framed in ways analogous to studies of "illness experience" carried out in the field of medical anthropology [64]. By eliciting open-ended descriptions of patients' "clinical trial experience," our investigative team of medical anthropologists and sociologists will capture changes in their expectations in "real time." The qualitative interviews will also capture information about patients' immediate assessments of treatment and treatment assignment. This qualitative data will be used to inform quantitative data analysis and to generate prospective hypotheses for future studies.

3.5. Measures

We chose our measures to increase our understanding of placebo phenomena in this population (see Table 2 below for more detail). When given the choice, we chose to use continuous rather than dichotomous variables. Our preference for continuous scales is supported by the recent meta-analysis that compared placebo arms with wait-list controls, and found that studies using continuous variables were more likely to detect significant placebo effects [2].

3.6. Ethical considerations

The placebo run-in design continues to be used in research, especially where large placebo responses in the placebo arm seem to threaten the ability to detect the effect of the intervention (e.g. trials for depression, hypertension, pain) [11]. Nonetheless, ethical reservations and objections have begun to be raised concerning the methods of placebo run-ins. It is clear that participants in a RCT run-in are rarely informed of "all procedures to be followed" including the fact that there is a preliminary period during which they all will receive a placebo. Some have argued that many informed consents in studies using run-ins do not adhere to the *Federal Register* [65] not to mention the Nuremberg Code and the Declaration of Helsinki [66]. More recently, some have suggested that new wordings or procedures that mandate absolute transparency of informed consent leave no room for an element of concealment or deception [67–69]. Our team considered many options in our informed consent. At the time of ethical review, our institution's Institutional Review Board decided to consider the run-in a legitimate device of concealment rather than an unethical deception. The ethics of the run-in are in a state of flux, and future placebo research may have to follow some of the more stringent requirements being considered at the present time [67–69].

4. Discussion

We have discussed our efforts to address and improve upon current shortcomings in clinical trials geared toward detecting placebo effects; specifically to counteract three major problems of placebo research: (1) the design of

Table 2
Measures to capture change in symptom evaluation, psychological state, quality of life, and psychosocial aspects are described below

Scale	Description
Symptom evaluation	
Irritable Bowel Syndrome Symptom Severity Scale (IBS-SSS) [1]	This 5-item questionnaire provides a simple way to scale IBS symptoms and the progress of the disease. Items consider pain, distention, bowel dysfunction, and quality of life/global well being. Scores showed good reliability and sensitivity to change.
IBS Global Improvement Scale [2]	This single question records a rating of IBS Global Improvement. "Compared to the way you usually felt during the 3 months before you entered the study, are your IBS symptoms over the past 3 weeks: substantially worse, moderately worse, slightly worse, no change, slightly improved, moderately improved, substantially improved?" This scale is unusual in the IBS literature as most measures of improvement are binary; we chose this scale for its potential ability to allow us to see more variation in our sample than would be allowed by a binary scale.
Adequate Relief of IBS Symptoms [3]	A single yes or no question makes up this simple global measure which has been extensively tested for responsiveness and validity in IBS. It examines whether patients experienced adequate relief of IBS pain and discomfort in the last week, "In the past seven days have you had adequate relief of your irritable bowel syndrome pain and discomfort?"
Psychological measures	
Hospital Anxiety and Depression Scales (HAD).[4–6]	This 14-item measure is validated for detecting mild mood disorders in non-psychiatric outpatients, and has been used in patients with IBS. This measure has been well validated and is able to distinguish between depression and anxiety.
Quality of life	
Irritable Bowel Syndrome Quality of Life (IBS-QOL) [7,8]	This 35-item scale is designed to assess the impact of IBS on 8 dimensions of health status including: dysphoria, interference with activity, body image, health worry, food avoidance, social reaction, and sexual relationships. It shows high internal consistency, reproducibility and good convergent and discriminant validity.
Psychosocial measures	
Social Support [9–12]	Our expert study team will add to the Sarason Social support questionnaire, which is designed to measure respondent's number of social supports and the degree of satisfaction with them. Our additional 6 items will measure subjects' perception of the three main types of supportive social interactions: emotional, informational, instrumental.
Implicit Models of Illness Questionnaire (IMIQ) [13]	This 38-item self-report measure assesses respondents' evaluation of their illness along 4 dimensions: seriousness of illness, personal responsibility for illness, controllability of illness, and its changeability. The IMIQ scales have discriminant validity and good reliability.

appropriate controls, (2) dealing with relatively small effect sizes, and (3) investigating the patients' experience of being in an RCT and receiving a placebo. Our interventions are designed to maximize the placebo response, while at the same time addressing the hypothesis that the magnitude of placebo effects associated with patient–practitioner context are modifiable and can be systematically varied in a manner analogous to dose-dependence. A positive finding that placebo can be delivered at different levels in a manner analogous to dose dependence would have important clinical and research implications that would extend beyond IBS and its treatment by acupuncture. A finding that patients in the 'augmented' group had better outcomes than patients in the 'limited' group would have significant implications for how medicine is practiced; it would demonstrate that patients may benefit from stronger patient–practitioner relationships. Further research could then address whether patients who receive this augmented treatment would use fewer visits and less costly care.

A finding that patients in the 'limited' group had better outcomes than the 'wait list' group would provide strong evidence for the scientific validity of placebo effects and would address Hrobjartsson and Gotzsche's meta-analytic [2] finding that there was "little evidence that placebo had powerful clinical effects" beyond natural history and regression to the mean. An absence of such a result would support their findings. Finally, the discovery of a 'dose-dependent' relation between the three arms would demonstrate for researchers a specific method of modulating placebo effects within a RCT, and could have important methodological implications for the design of clinical trials. A negative finding in our primary aim, i.e. the failure to find augmented superior to limited or limited superior to waitlist, would also be significant suggesting that age-old clinical beliefs need to be re-examined. Last, regardless of our main study outcomes,

our parallel qualitative sub-study will provide several novel types of data on patients' experience of IBS acupuncture. This data will not only be useful in the design of future trials and treatments, but will inform the current debates surrounding the ethics of conducting placebo research.

Acknowledgements

This paper is funded by the following grants from the National Center for Complementary and Alternative Medicine (NCCAM) at the NIH: 1R01 AT01414-01, 1R21 AT002860-01 and 1R21 AT002564-01. Several other individuals have contributed to the design and execution of this trial including Mary Quilty, William Stason, and Oriana Rodgriues.

References

[1] Kaptchuk T. Powerful placebo: the dark side of the randomised controlled trial. Lancet 1998;351(9117):1722–5.
[2] Hrobjartsson A, Gotzsche P. Is the placebo powerless? An analysis of clinical trials comparing placebo with no treatment. N Engl J Med 2001;344(21):1594–602.
[3] Finnis D, Benedetti F. Mechanisms of placebo response and their impact on clinical trials and clinical practice. Pain 2005;114:3–6.
[4] Wager T, Rilling JK, Smith EE, et al. Placebo-induced changes in fMRI in the anticipation and experience of pain. Science 2004;303 (5661):1162–7.
[5] Hrobjartsson A. What are the main methodological problems in the estimation of placebo effects? J Clin Epidemiol 2002;55:430–5.
[6] Kaptchuk T. The placebo effect in alternative medicine: can the performance of a healing ritual have clinical significance? Ann Intern Med 2002;136(11):817–25.
[7] http://placebo.nih.gov.
[8] http://grants1nih.gov/grants/guide/rfa-files/RFA-AT-02-002.html, Departments of Health and Human Services: Elucidation of the underlying mechanism of placebo effect.
[9] Guess HA, Kleinman A, Kusek JW, Engel LW. The science of the Placebo: toward an interdisciplinary research agenda. London: BMJ Books; 2002.
[10] Temple R. Where are clinical trials of a given agent vs. placebo no longer appropriate or feasible. Control Clin Trials 1997;18:613–20.
[11] Temple RaSSE. Placebo-controlled trials and active-control trials in the evaluation of new treatments: Part 1. Ethical and scientific issues. Ann Intern Med 2000;133:455–63.
[12] Davis CE. The effect of regression to the mean in epidemiologic and clinical studies. Am J Epidemiol 1976;104:493–8.
[13] Kelkar P, Ross MA. Natural history of disease and placebo effect. Perspect Biol Med 1994;37:244–6.
[14] McDonald C, Mazzuca SA, McCabe GP. How much of the placebo "effect" is really statistical regression? Stat Med 1983;2:417–27.
[15] Senn S. Estimating treatment effects in clinical trials subject to regression to the mean. Biometrics 1985;41:555–60.
[16] Whitney C, von Korff M. Regression to the mean in treated versus untreated chronic pain. Pain 1992;50:281–5.
[17] Yudkin P, Stratton IM. How to deal with regression to the mean in intervention studies. Lancet 1996;347:241–3.
[18] Bouchet C, Guillemin F, Briancon S. Nonspecific effects in longitudinal studies: impact of quality of life measures. J Clin Epidemiol 1996;49:15–20.
[19] Campbell J, Maxey VM, Watson WA. Hawthorne effect: implications for prehospital research. Ann Emerg Med 1995;26:590–4.
[20] Dalession D. Studies on pain and the Hawthorne effect. Headache 1974;11:109–10.
[21] Gillespie R. Manufacturing knowledge: a history of the Hawthorne experiments. Cambridge: Cambridge University Press; 1991.
[22] Miller F. Is the placebo powerless [letter]. N Engl J Med 2001;345:1277.
[23] Parsons H. What happened at Hawthorne? Science 1974;183:922–32.
[24] Jailwala J, Imperiale TF, Kroenke K. Pharmacologic treatment of the irritable bowel syndrome: a systematic review of randomized, controlled trials. Ann Intern Med 2000;133:136–47.
[25] Patel M, Stason WB, Legedza A, et al. The placebo effect in irritable bowel syndrome trials: a meta-analysis. Neurogastroenterology Motility 2005;17(3):332–40.
[26] Fielding J. Clinical recognition of the irritable bowel syndrome. Ir Med J 1981;74:134.
[27] Lembo A, Camilleri M. Chronic constipation. N Engl J Med 2003;349(17):1360–8.
[28] Parkman H, Miner P, Rao SC, et al. NT-3 improves functional constipation. Am J Gastroenterol 2001;96.
[29] Camilleri M, Williams DE. Economic burden of irritable bowel syndrome. Proposed strategies to control expenditures. Pharmacoeconomics 2000;17:331–8.
[30] Mueller-Lissner S, Fumagalli I, Pace F, Nault B, Lefkowitz MP. Tegaserod, a 5-HT4 reception partial agonist, relieves key symptoms of irritable bowel syndrome. Gastroenterology 2000;118(2).
[31] Van Weel C. Examination of context of medicine. Lancet 2001;357:733.
[32] Leventhal H, N. D, Straus A. Self-regulation and mechanisms for symptom appraisal. Psychosocial Epidemiology. Symptoms, Illness Behavior and Help SeekingNew York: Neale Watson; 1982.
[33] Whitehead WE, Palsson OS. Is rectal pain sensitivity a biological marker for irritable bowel syndrome: psychological influences on pain perception. Gastroenterology 1998;115:1263–71.

[34] Lundh L. Placebo beliefs and health: a cognitive-emotional model. Scand J Psychol 1987;28:128–43.
[35] Quitkin F. Placebos, drug effects, and study design: a clinician's guide. Am J Psych 1999;156:829–36.
[36] Pablos-Mendez A, Barr RG, Shea S. Run-in periods in randomized controlled trials: implications for the application of results in clinical practice. JAMA 1998;279:222–5.
[37] Senn S. Are placebo run-in justified. Br Med J 1997;314:1191–3.
[38] Knipschild P, Leffer P, Feinstein AR. The qualification period. J Clin Epidemiol 1991;63:279–88.
[39] Blackwelder W, Hasting BK, Lee MLF, Deloria MA. Value of a run-in period in a drug trial during pregnancy. Control Clin Trials 1990;11:187–98.
[40] Kopec J, Abrahamowicz M, Esdaile JM. Randomized discontinuation trials: utility and efficacy. J Clin Epidemiol 1993;46:959–71.
[41] Davis C, Applegate WB, Gordon DJ, Curtis RC, McCormick M. An empirical evaluation of the placebo run-in. Control Clin Trials 1995;16:41–50.
[42] Reimherr F, Ward MF, Beverley WF. The introductory placebo washout: a retrospective evaluation. Psych Rev 1989;30:191–9.
[43] Trivedi M, Rush AJ. Does a placebo run-in or a placebo treatment cell affect the efficacy of antidepressant medications? Neuropsychopharmacology 1994;11:33–43.
[44] Walsch B, Seidman SN, Sysko R, Gould M. Placebo responses in studies of major depression: variable, substantial and growing. JAMA 2002;287:1840–7.
[45] Verger V, Rezvani A, Makaarewicz VA. Direct effect on validity of response run-in selection in clinical trials. Control Clin Trials 2003;24:156–66.
[46] Kong J, Fufa DT, Gerber AJ, et al. Psychophysical outcomes from a randomized pilot study of manual, electro, and sham acupuncture treatment on experimentally induced thermal pain. J Pain 2005;6:55–64.
[47] Pariente J, White P, Frackowiak RS, Lewith G. Expectancy and belief modulate the neuronal substrates of pain treated by acupuncture. Neuroimage 2005;25:1161–7.
[48] Streitberger K, Kleinhenz J. Introducing a placebo needle into acupuncture research. Lancet 1998;352(9125):364–5.
[49] White P, Lewith G, Hopwood V, Prescott P. The placebo needle, is it a valid and convincing placebo for use in acupuncture trials? A randomised, single-blind, cross-over pilot trial. Pain 2003;106:401–9.
[50] Thompson WG, Longstreth GF, Drossman DA, Heaton KW, Irvine EJ, Muller-Lissner SA. Functional bowel disorders and functional abdominal pain. Gut 1999;45(2):1143–7.
[51] Vanner S, Depew WT, Paterson WG, et al. Predictive value of the Rome criteria for diagnosing the irritable bowel syndrome. Am J Gastroenterol 1999;94:2912–7.
[52] Di Blasi Z, Harkness E, Ernst E, Gerogiou A, Kleijnen J. Influence of context effects on health outcomes: a systematic review. Lancet 2001;357:757–62.
[53] Keefer L, B.E.. The effects of relaxation response meditation on the symptoms of irritable bowel syndrome: results of a controlled treatment study. Behav Res Ther 2001;39(7):801–11.
[54] Drossman D, Toner BB, Whitehead WE, et al. Cognitive-behavioral therapy versus education and desipramine versus placebo for moderate to severe functional bowel disorders. Gastroenterology 2003;125(1):19–31.
[55] Bengtsson M, Ulander K, Borgdal EB, Christensson AC, Ohlsson B. A course of instruction for women with irritable bowel syndrome. Patient Educ Couns 2005.
[56] Heymann-Monnikes I, A. R, Florin I, Herda C, Melfsen S, Monnikes H. The combination of medical treatment plus multicomponent behavioral therapy is superior to medical treatment alone in the therapy of irritable bowel syndrome. Am J Gastroenterol 2000;95 (4):981–94.
[57] Marvel MK, D. W, Weiner E. Medical interviewing by exemplary family physicians. J Fam Pract 1998;47(5):343–8.
[58] Lang EV, Benotsch EG, Fick LJ, et al. Adjuctive non-pharmacological analgesia for invasive medical procedures: a randomized trial. Lancet 2000;355(9214):1486–90.
[59] Morag E, Liegerman G, Shaffer K, Novelline R, Volkan K, Lang EV. Objective Structured Clinical Examination (OSCE) for medical students in radiology: comparison with traditional clerkship evaluations. Acad Radiol 2001;8:74–81.
[60] Moncher F, Prinz RJ. Treatmet fidelity in outcome studies. Clin Psychol Rev 1991;11:247–66.
[61] Shrout P, Fleiss JL. Intraclass correlations; use in assessing rater reliability. Psychol Bull 1979;86:420–8.
[62] Stone D, Kerr CE, Jacobson E, Conboy LA, Kaptchuk TJ. Patient expectations in placebo-controlled randomized clinical trials. J Eval Clin Pract 2005;11(1):77–84.
[63] Kaptchuk T. Acupuncture: theory, efficacy, and practice. Ann Intern Med 2002;136(5):374–83.
[64] Kleinman A. Patients and healers in the context of culture: an exploration of the borderland between anthropology, medicine and psychiatry. Berkeley: University of California Press; 1981.
[65] Protection of human subjects: informed consent. Federal Register; 1981.
[66] Kaptchuk TJ. Concerns about run-in periods in randomized trials. J Am Med Assoc 1998;279:1526–7.
[67] Evans M. Justified deception? The single blind placebo in drug research. J Med Ethics 2002;26:188–93.
[68] Ramsey L. Commentary: placebo run ins have some value. Br Med J 1997;314:1193.
[69] Wendler D, Miller FG. Deception in the pursuit of science. Arch Intern Med 2004;164:597–600.

Table references

[1] Francis CY, Morris J, Whorwell PJ. The irritable bowel severity scoring system: a simple method of monitoring irritable bowel syndrome and its progess. Aliment Pharmacol Ther 1997;11:395–402.

[2] Lembo AJ, Camilleri M. Chronic constipation. N Engl J Med 2003;349(17):1360–8.
[3] Camilleri M, Williams DE. Economic burden of irritable bowel syndrome. Proposed strategies to control expenditures. Pharmacoeconomics 2000;17:331–8.
[4] Snaith RP, Zigmond AS. The hospital anxiety and depression scale. Br Med J 1986;292:344.
[5] Snaith RP, Taylor CM. Rating scales for depression and anxiety: a current perspective. Br J Clin Pharmacol 1985;19:17s–20s.
[6] Zigmond AS, Snaith RP. The hospital and anxiety scale. Acta Psychiatr Scand 1983;67:361–70.
[7] Drossman DA, Patrick DL, Whitehead WE, et al. Further Validation of the IBS-QOL: a disease specific quality-of-life questionnaire. Am J Gastroenterol 2000;95:999–1007.
[8] Patrick DL, Drossman DA, Frederick IO, Dicesare J, Puder K. Quality of life in persons with irritable bowel syndrome: development and validation of a new measure. Dig Dis Sci 1998;43:400–11.
[9] Cohen S, Lichtenstein E. Partner behaviors that support quitting smoking. J Consult Clin Psychol 1990;58:304–9.
[10] Cohen S. Psychosocial models of the role of social support in the etiology of physical diseases. Health Psychol 1988;73:269–97.
[11] Helgeson VS, Cohen S. Social support and adjustment to cancer: reconciling descriptive, correlational, and intervention research. Health Psychol 1996;15:135–48.
[12] Sarason J. Sarason Social support questionnaire (brief form). Soc Personal Relation 1986;4:497–510.
[13] Turk DC. Implicit models of illness. J Behav Med 1986;9:453–74.