How Pilot Studies Improve Large-Scale Clinical Trials: Lessons Learned from the COMBINE Study*

STEPHANIE S. O'MALLEY, PH.D.,† DANIEL J. MARTIN, PH.D., JAMES D. HOSKING, PH.D.,† AND BARBARA J. MASON, PH.D.†

Department of Psychiatry, Yale University School of Medicine, CMHC – SAC202, 34 Park Street, New Haven, Connecticut 06519

ABSTRACT. Objective: The design of a clinical trial to evaluate a potential therapy requires decisions about issues that include safety, efficacy, measurement, feasibility and training. Experience from the COMBINE Study, which tests the combination of medications and behavioral therapies for alcohol dependence, is presented as an example of how pilot studies improve large-scale clinical trials. **Method:** The COMBINE Pilot 1 inpatient study was designed to inform the main trial about the safety and tolerability of the doses of acamprosate (3 g/day) and naltrexone (100 mg/day) selected for study, alone and in combination. Pilot 2 was conducted as a feasibility study for the main trial, with the goals of (1) assessing the length of and compliance with research assessments, (2) developing methods for subject recruitment and staff training and (3) assessing the safety of the medications under less controlled outpatient conditions. **Results:** Results from Pilot 1 provided safety information to support testing the medications in an outpatient study and contributed to the decision to incorporate dose reductions into the main trial protocol to manage adverse events. The results of Pilot 2 formed a basis for (1) reducing the length of the assessment battery, (2) having staff fully trained and recruitment procedures established for the main trial and (3) extending the drug safety results of Pilot 1 to outpatient conditions similar to those of the main trial. **Conclusions:** The COMBINE Study provides several examples of the successful application of pilot studies to inform the design of a clinical trial. (*J. Stud. Alcohol*, Supplement No. 15: 66-71, 2005)

THE EVALUATION of a new therapy for the treatment of disease in humans can rarely be based on the results of a single study. Before embarking on a definitive efficacy study, each key element of the study protocol must be decided and an evaluation made as to whether sufficient information is available on which to justify those decisions. Such decisions can extend to questions related to safety, efficacy, measurement, feasibility and training, among others.

When planning a trial to evaluate the efficacy of a new therapy, much of the information should already be available from the research literature. However, more often than not some aspect requires additional preliminary work in the form of a pilot study. Ultimately, the success of the trial may rest on the foundation of this preliminary work. The COMBINE Study, which tests the combination of medications and behavioral therapies for alcohol dependence, provides several examples of the kinds of questions that can be addressed through pilot studies.

*This work was supported by National Institute on Alcohol Abuse and Alcoholism Cooperative Agreements U10AA011787 (O'Malley), U10AA011721 (Hosking), U10AA011727 (Mason) and K05AA014715 (O'Malley).

†Correspondence may be sent to Stephanie S. O'Malley at the above address, or via email at: Stephanie.omalley@yale.edu. James D. Hosking is with the Department of Biostatistics (CSCC), University of North Carolina at Chapel Hill, Chapel Hill, NC. Barbara J. Mason is with the Pearson Center for Alcoholism and Addiction Research, Department of Neuropharmacology, The Scripps Research Institute, La Jolla, CA.

Medication development

The process of the development of new medical treatments (drugs, biologics, devices) has evolved into a standardized sequence of clinical trials following completion of preclinical testing of the drug in animals (http://www.fda.gov/cder/handbook/). The trials in this sequence are typically divided into four phases. In this process, initiation of trials in each phase is contingent on completion of the prior phase. In Phase 1 studies, the new drug is tested in humans for the first time to determine the drug's metabolism and pharmacological actions, to discover the side effects associated with increasing doses and to gain an early indication of effectiveness. These studies may be performed in healthy individuals and/or patients. Subsequently, controlled Phase 2 clinical studies are performed to evaluate the efficacy and safety of the drug for a particular indication in patients with the disorder under study and to determine the common short-term side effects and risks. Phase 3 studies are larger, controlled trials that are initiated after preliminary evidence suggesting effectiveness of the drug has been obtained. Phase 3 trials are intended to gather additional information to evaluate the overall benefit-risk relationship of the drug and provide an adequate basis for physician labeling (http://www.clinicaltrials.gov). The results of the Phase 1-3 trials are compiled and form the basis for evaluation of the drug by the Food and Drug Administration (FDA). Finally, Phase 4 trials provide additional

safety and efficacy information in patients, after drug approval. These studies often investigate the long-term efficacy and safety of the drug, delineate its optimal use and estimate efficacy in "standard clinical practice."

Behavioral therapy development

In the development of a new behavioral therapy, three sequential stages of the scientific process have been outlined that take initial clinical conceptualization up through effectiveness research (Onken et al., 1997; Rounsaville et al., 2001). In the first stage, manual writing, the development of training programs and adherence and competency measures for the new treatments are developed in conjunction with pilot testing. The goal of Stage 1 research is to specify the elements required to test the efficacy of the new therapy in a randomized clinical trial (Rounsaville et al., 2001). In concept, this goal corresponds to the Phase 1 and 2 objectives of defining a safe and effective dose of a drug for definitive efficacy testing. Stage 2 consists of randomized controlled clinical trials to evaluate the efficacy of the manual-based treatment that showed promise in earlier Stage 1 pilot testing. Successful Stage 2 studies can be followed by additional Stage 2 studies designed to examine mechanisms of action or effective components of the treatments. Finally, Stage 3 studies are those that evaluate the transportability of efficacious treatments into real practice, similar to Phase 4 studies in drug development.

The COMBINE Study

The COMBINE Study can be considered a Phase 3/4 study with regard to the medications being tested and a Stage 2 study with regard to the behavioral therapies under investigation. As described previously (COMBINE Study Research Group, 2003a), COMBINE is evaluating a variety of combinations of two medications and two behavioral interventions. The status of prior research on each intervention is as follows:

- Naltrexone was approved by the FDA in 1994 for the treatment of alcohol dependence. It has also been approved by a number of international regulatory authorities.
- Acamprosate was not approved by the FDA when the study was undertaken, although it was recently approved for use in alcohol dependence (July 2004). It had already been approved by a number of international regulatory authorities.
- Medical Management (MM) is a manual-based treatment designed to approximate a primary care approach to alcohol dependence. Although consisting of standard, tested approaches to promote abstinence and enhance medication compliance, its efficacy has not been formally tested.
- Combined Behavioral Intervention (CBI) is individualized psychotherapy suitable for delivery by psychotherapists with specialty training and/or experience in alcoholism treatment. It is composed of elements of the successful behavioral interventions from Project MATCH, a nationwide, multisite, patient-treatment matching study (Project MATCH Research Group, 1997). Again, the efficacy of this specific package has not been formally tested.

The primary objective of COMBINE is to determine if improvements in treatment outcome for alcohol dependence can be achieved by combinations of pharmacotherapies and psychotherapies. The COMBINE main study is a randomized, double-blind clinical trial. Essentially, the design is a $2 \times 2 \times 2$ factorial, with naltrexone vs placebo, acamprosate vs placebo and MM + CBI vs MM alone as the three factors. A ninth cell is evaluating CBI without MM or any pills.

As with any single intervention, the basic issues in evaluating these combined treatments are selecting optimal doses of each component, maximizing compliance and patient acceptance and then evaluating the efficacy, safety and patient acceptance of the selected regimen. The investigators felt the evidence for the efficacy of each pharmacotherapy was substantial, but not overwhelming. There was a consensus in favor of using each drug at the high end of the doses studied for alcohol dependence (100 mg/day for naltrexone and 3 g/day for acamprosate).

Evidence for the safety and tolerability of each of the medications individually at these doses was judged adequate. However, data on the safety and tolerability of the combination were quite limited. Information was available from only one trial, and that was performed in healthy subjects and at standard therapeutic doses (50 mg/day naltrexone and 2 g/day acamprosate; Mason et al., 2002). Therefore, evaluating the safety and tolerability of the combination of naltrexone and acamprosate at higher doses, in alcohol-dependent individuals, was the first major objective of the pilot studies.

Although the behavioral treatments to be tested in COMBINE were based on existing treatments, they were modified for the purposes of COMBINE through extensive Stage 1 work. For example, CBI, although rooted in the three manual-based treatments tested in Project MATCH (Longabaugh et al., this supplement), represented an amalgamation of the purported active ingredients of these treatments in addition to new components that were developed to address perceived gaps. Similarly, the MM counseling was based on several existing manuals (Carroll and O'Malley, 1996; Fleming et al., 1996; Mason and Goodman, 1997; Volpicelli et al., 2001); however, the treatment needed to be modified to fit the specifics of the medications being tested in COMBINE (Pettinati et al., this supplement). Thus, manuals for the two approaches specifying the techniques to be used and to be excluded, a plan for training and measures to evaluate adherence and competence had to be developed.

Finally, the investigators were concerned about the feasibility of the protocol, in terms both of participant and staff burden. The pharmacotherapy protocol required participants to take eight pills per day (four in the morning, two at midday and two at night). The psychotherapy protocol required as many as three sessions a week (one for MM and up to two for CBI), in addition to a session for research assessments in some weeks. In addition to the concerns about participant recruitment that are ubiquitous in clinical trials, the "CBI, no pills" condition raised some special concerns. Would participants who volunteered to participate in a protocol testing two pharmacotherapies, in which they had an 8 in 9 chance of being assigned to "pills" (albeit, possibly to double-placebo), accept assignment to the no pill condition?

To address these objectives, the COMBINE investigators designed two pilot studies: an inpatient trial focusing on safety and tolerability and an outpatient trial designed as a feasibility study of the complete protocol proposed for the main trial. A description of each trial and of the lessons learned from each are detailed below.

Pilot 1: Dose-Ranging Kinetics and Behavioral Pharmacology of Naltrexone and Acamprosate, Both Alone and Combined

The first pilot study was designed to answer several questions about the combination of naltrexone and acamprosate, including whether the medication doses to be combined are safe (Johnson et al., 2003). The study tested the pharmacological and behavioral safety and tolerability of low vs high doses of naltrexone (50 mg/day vs 100 mg/day) and acamprosate (2 g/day vs 3 g/day), independently and combined, among 23 nontreatment-seeking, alcohol-dependent individuals, in a protocol spanning 23 days. This study was conducted on an inpatient unit to provide careful medical monitoring and support to determine the upper limit of compliance with the medication schedule and the course and severity of adverse events. The inpatient setting also provided an environment that supported the carefully controlled execution of the protocol. The details of the protocol and the results are delineated by Johnson et al. (2003). Briefly, participants were randomly assigned to one of four sequences of drug administration, with each sequence incorporating four phases. In Phase 1 (Days 1-3), all participants received double placebos. In Phase 2 (Days 4-9), participants either received 2 g acamprosate, 3 g acamprosate, 50 mg naltrexone or 100 mg naltrexone and placebo for the other medication. Participants continued on their original medication throughout the remainder of the experiment. In Phase 3 (Days 10-15), the other medication was added at the lower dose and was increased to the higher dose in Phase 4 (Days 16-21). For example, the sequence for one of the four groups was as follows: placebo (Phase 1), 2 g acamprosate (Phase 2), 2 g acamprosate + 50 mg naltrexone (Phase 3) and 2 g acamprosate + 100 mg naltrexone (Phase 4). To evaluate the safety of the medications, an extensive battery of assessments was obtained daily, including adverse events, vital signs, mood, target abuse liability assessments, signs and symptoms of withdrawal, sleep quality and measures of attention and learning. At the end of each phase, memory was evaluated, a neurological examination was conducted, and a comprehensive panel hematology and biochemistry indices were analyzed. Trough levels of 6-β-naltrexone and acetylhomotaurine were measured daily to examine potential pharmacokinetic interactions.

A priori specified criteria for pathological changes from baseline were analyzed, and clinical summaries were reviewed. The results revealed that the medications were generally well tolerated and the attrition rate from the complex and lengthy study was modest (approximately 15%). These data provided important safety information to support testing the combination in an outpatient setting. The finding that escalating the dose of acamprosate from 2 g/day to 3 g/day was associated with increases in nervousness and fatigue helped inform the decision to incorporate dose reductions into the final protocol to help manage adverse events (COMBINE Study Research Group, 2003a,b).

Pilot 2: Testing Combined Pharmacotherapies and Behavioral Interventions for Alcohol Dependence: A Pilot Feasibility Study

Pilot 2 was conducted as a feasibility study that tested out every aspect of the proposed protocol for the main trial, with the exception of the follow-up interviews that occurred after treatment termination (Combine Study Research Group, 2003b). In this pilot study, 108 individuals across 11 sites were randomized to receive placebo, naltrexone or acamprosate alone or in combination. With the exception of one group who received CBI alone without pills, the remaining participants also received either MM alone or MM in combination with CBI for 16 weeks. Participants completed all of the proposed research assessments, which included measures of health status (physical examination, physiological and laboratory assessments), adverse events, treatment-related expectancies, alcohol consumption, alcohol and drug involvement, motivation, craving, psychological symptoms and diagnostic information, social support and quality of life. The decision to conduct this pilot study was based on questions about medical and practical issues in the design of the main trial. For example, we needed to assess the willingness of participants to engage in and adhere to a study requiring multiple treatment appointments and research assessments while taking more than eight pills per day in a less controlled environment than studied in Pilot 1. Training of staff in the behavioral interventions and the research procedures was another important objective.

One decision made in the design of this pilot study that differs from what would typically be done in the development of a new drug was the decision not to obtain preliminary estimates of efficacy. There were two major motivations for this choice. First, all the existing data suggested that the efficacy of the various interventions as monotherapies would be modest (e.g., perhaps an improvement of 10% in percentage of days abstinent in comparison with placebo). The investigators saw no reason to expect the incremental efficacy of the combined therapies to exceed that of the monotherapies by any larger amount. Power computations suggested that even the main study (with a target sample size of more than 1,300) was marginally powered to detect such an effect (COMBINE Study Research Group, 2003a). Second, data from Project MATCH (Project MATCH Research Group, 1998) suggested that the full effect of the behavioral therapies might not be apparent at the end of therapy but, rather, emerge only during the posttreatment follow-up period. Thus we concluded that efficacy estimates from the pilot had a large risk of being misleading, rather than informative. For this reason, whereas outcome assessments were conducted to model the participant and staff burden in the main protocol and to provide staff training, the pilot protocol specifically excluded analysis of those data.

Training of staff

Prior to the main trial, staff needed to become familiar with the many procedural issues that were required. Pilot 2 served the objective of providing training to the clinicians and research assistants in the study, thereby ensuring that staff would be fully trained by the onset of the main study. The opportunity to gain experience with the study procedures during Pilot 2 also benefited the main trial in other ways. For example, the rate of completion of endpoint data at Week 16 is very high in the main trial, due in part to the experience gained in methods to follow-up participants who discontinued treatment early in Pilot 2.

Subject recruitment and eligibility

A key reason that many clinical trials fail is an inability to meet recruitment goals. Because of this, during the planning stages of the trial, much attention was paid to establishing thorough recruitment plans for each site. During Pilot 2, participants were recruited by a variety of advertisements and from referrals from local clinics across all 11 participating study sites. Sites tracked each interested participant's recruitment source to gain knowledge about which recruitment methods were more or less successful.

It was expected that the eligibility criteria were not so restrictive as to make recruitment overly difficult. Although results from Pilot 2 suggested that no specific criterion unduly affected eligibility, a significant number of participants were excluded due to concurrent treatment with antidepressants. As a result, there was considerable discussion regarding whether to open up enrollment to those on stable antidepressant therapy in the main trial. On the one hand, doing so would facilitate recruitment and enhance the generalizability of the results. On the other hand, enrolling these participants might introduce another source of variability into the study, potentially as a result of adverse events from the combination of naltrexone, acamprosate and antidepressants. Given that we did not have safety or efficacy data on the combination of naltrexone, acamprosate and antidepressants, we decided against modifying this exclusion criterion immediately. However, a small sample of patients who were excluded from the main study because of current antidepressant treatment were offered participation in an outpatient safety study of combined treatment using procedures similar to those of Pilot 2, except that individuals were provided with open label acamprosate and naltrexone while being maintained on their originally prescribed antidepressant. Ultimately, recruitment for the main trial proceeded well, and we did not need to use these data for the purpose of informing further discussions about enrolling patients on antidepressants. However, the data will be compared to data from matched individuals from Pilot 2 who received naltrexone and acamprosate but were not on an antidepressant, with the goal of providing initial information on safety and drinking outcomes.

Length of assessments

It is difficult to determine in advance the actual length of time that a battery of assessments will take to complete. In designing the protocol, emphasis had been placed on limiting the number of intake assessments to minimize the likelihood that clinical improvement due to assessment reactivity would outweigh the effects of the treatments under study. It was estimated that the initial assessment battery would take approximately 4 hours to complete. However, Pilot 2 revealed that the battery took an average of about 6 hours to complete. To ease staff and participant burden, the number of assessments was reduced for the main trial. The resulting baseline evaluation (including informed consent, a physical examination, laboratory testing and formal assessments) was limited to one that takes approximately 4.75 hours to complete (COMBINE Study Research Group, 2003a).

Subject acceptance of and adherence to the protocol

At the time of trial planning, it was unclear whether participants would be willing to take part in a trial that offered multiple treatments and research assessments. In addition, for those receiving medication, the protocol specified that eight pills per day be taken over a 16-week

period, which might be cumbersome for patients. Conversely, questions were raised about whether randomization to a behavioral therapy only condition would be acceptable to participants being recruited into a study with a strong emphasis on medications.

Our experience during the Pilot 2 study indicated that participants were willing to cooperate with all aspects of the trial, including the assessments, the possibility of being randomized to the therapy only condition (CBI only) and the medication regimen. With respect to medication adherence, Pilot 2 indicated that participants took about 65% of the total pills prescribed by the protocol, which was considered acceptable. Although compliance was significantly better for participants who received both behavioral interventions (MM + CBI) than those that received only MM, there was no difference between the medication conditions, which provided additional reassurance on the combination. With regard to overall adherence with the protocol, the results of Pilot 2 suggested that more than 69% of participants completed the study, and about three quarters of subjects provided endpoint data at Week 16, with no significant differences on the variables among treatment groups. These results suggested that participants would adhere to study requirements during the main trial and would not differentially drop out of treatment across groups.

Safety

As mentioned earlier, concern existed about how well participants would tolerate the medications, especially the combination of acamprosate and naltrexone, in an outpatient setting. Although Pilot 1 and an earlier study (Mason et al., 2002) suggested that the medications would be well tolerated in combination, both studies were conducted in a highly controlled environment. A recent single-site outpatient clinical trial likewise found the combination of naltrexone and acamprosate to be well tolerated in the standard therapeutic doses (50 mg/day and 2 g/day, respectively; Kiefer et al., 2003). Pilot 2 evaluated the higher doses specified in the COMBINE trial (100 mg/day naltrexone and 3 g/day acamprosate) in an outpatient environment where participants had access to alcohol. To gather as much experience with the combination as possible, the number of participants assigned to the combined active medication group was doubled relative to the other treatment conditions. In the main trial, randomization was equal to all nine treatment conditions.

The Pilot 2 study results showed that, although there was a significant difference in dose reductions between individuals in the placebo vs the active medication groups, there were no significant differences between any of the active medication groups, including the combined acamprosate/naltrexone group. Paralleling this finding, the rate of adverse events for the combination group was similar to that of the monotherapy groups. The results also suggested that there was not an increased risk of liver or kidney toxicity for participants taking both medications and suggested that few individuals would experience an increase in liver enzymes during the main trial. This was an important question because the study planned to use a higher dose of naltrexone (100 mg/day) than had been used in previous studies. Only one patient, who had relapsed to drinking and had tested positive for hepatitis C, was discontinued due to an elevation in liver enzymes. The planned frequency of monthly monitoring of liver function tests was deemed appropriate for the main trial. Consistent with the results of Pilot 1, the results of Pilot 2 supported the safety of the combination of naltrexone and acamprosate and extended this conclusion to outpatient treatment under the conditions that would be similar to the main trial.

Summary

Pilot studies may be mounted in advance of a main trial to clarify research design issues (e.g., doses selected for study, criteria for safety and efficacy, estimates of compliance and overall feasibility). The COMBINE Study provides several examples of the successful application of pilot studies to resolve research design questions to inform the design of a main trial. Results from the COMBINE Pilot 1 inpatient drug interaction safety study provided support for the safety and tolerability of the higher doses of acamprosate (3 g/day) and naltrexone (100 mg/day) selected for study in the main COMBINE trial. Pilot 1 data also informed the main trial study procedures for dose reduction as a method to manage adverse events.

Results from the COMBINE Pilot 2 outpatient feasibility study verified the safety and tolerability of the medication conditions studied in Pilot 1 in an outpatient setting and identified a need to reduce the quantity of research assessments in the main trial. Importantly, the standard operating procedures developed for staff training, subject recruitment and follow-up in Pilot 2 resulted in increased efficiency in executing the main trial and higher rates of completion of endpoint data.

Acknowledgments

The authors thank all the members of the COMBINE Steering Committee as well as the staff who participated in the pilot studies.

Disclosure

Dr. O'Malley has been a consultant to and received research support from Alkermes, Inc., and from Forest Laboratories. Dr. O'Malley has received research support from Lipha Pharmaceuticals, Inc., a subsidiary of Merck Sante s.b.s. and from Dupont. She has been a consultant to Johnson & Johnson, Ortho-McNeil and Pfizer and has received research support from Ortho-McNeil, Bristol-Myers Squibb and Pfizer. Dr. Mason

has served as a consultant to and/or received research support from DuPont Pharma; Lipha Pharmaceuticals, Inc., a subsidiary of Merck Sante s.b.s.; and Forest Laboratories, Inc. She has also served as an investigator in a clinical trial sponsored by Alkermes, Inc.

References

Carroll, K.M. and O'Malley, S.S. Compliance Enhancement: A Manual for the Psychopharmacotherapy of Alcohol Dependence, New Haven, CT: School of Medicine, Yale University, 1996, unpublished treatment manual.

COMBINE Study Research Group. Testing combined pharmacotherapies and behavioral interventions in alcohol dependence: Rationale and methods. Alcsm Clin. Exp. Res. **27:** 1107-1122, 2003a.

COMBINE Study Research Group. Testing combined pharmacotherapies and behavioral interventions for alcohol dependence (The COMBINE Study): A pilot feasibility study. Alcsm Clin. Exp. Res. **27:** 1123-1131, 2003b.

Fleming, M., Zweben, A., Barrett, D. and Manwell, L. Brief Motivational Enhancement Therapy: Conducting BMET in a Combined Psychosocial and Pharmacotherapy Clinical Trial, Madison, WI: University of Wisconsin, 1996, unpublished treatment manual.

Johnson, B.A., O'Malley, S.S., Ciraulo, D.A., Roache, J.D., Chambers, R.A., Sarid-Segal, O. and Couper, D. Dose-ranging kinetics and behavioral pharmacology of naltrexone and acamprosate, both alone and combined, in alcohol-dependent subjects. J. Clin. Psychopharmacol. **23:** 281-293, 2003.

Kiefer, F., Jahn, H., Tarnaske, T., Helwig, H., Briken, P., Holzbach, R., Kampf, P., Stracke, R., Baehr, M., Naber, D. and Wiedemann, K. Comparing and combining naltrexone and acamprosate in relapse prevention of alcoholism: A double-blind, placebo-controlled study. Arch. Gen. Psychiat. **60:** 92-99, 2003.

Mason, B.J. and Goodman, A.M. Brief Intervention and Medication Compliance Procedures: Therapist's Manual, New York: Lipha Pharmaceuticals, 1997.

Mason, B.J., Goodman, A.M., Dixon, R.M., Hameed, M.H., Hulot, T., Wesnes, K., Hunter, J.A. and Boyeson, M.G. A pharmacokinetic and pharmacodynamic drug interaction study of acamprosate and naltrexone. Neuropsychopharmacology **27:** 596-606, 2002.

Onken, L.S., Blaine, J.D. and Battjes, R. (Eds.) Behavioral Therapy Research: A Conceptualization of a Process, Washington, DC: American Psychiatric Press, 1997.

Project MATCH Research Group. Matching alcoholism treatments to client heterogeneity: Project MATCH posttreatment drinking outcomes. J. Stud. Alcohol **58:** 7-29, 1997.

Project MATCH Research Group. Matching alcoholism treatments to client heterogeneity: Project MATCH three-year drinking outcomes. Alcsm Clin. Exp. Res. **22:** 1300-1311, 1998.

Rounsaville, B.J., Carroll, K.M. and Onken, L.S. A stage model of behavioral therapies research: Getting started and moving on from stage I. Clin. Psychol. Sci. Pract. **8** (2): 133-142, 2001.

Volpicelli, J.R., Pettinati, H.M., McLellan, A.T. and O'Brien, C.P. Combining Medication and Psychosocial Treatments for Addictions: The BRENDA Approach, New York: Guilford Press, 2001.