424 R&R and PHS-398 Specific
Table Of Contents

PI: **REES, JUDITH R**	Title: Participant beliefs and bias in a randomized controlled trial	
	FOA: PAR12-144	
	FOA Title: NCI SMALL GRANTS PROGRAM FOR CANCER RESEARCH (NCI OMNIBUS R03)	
1 R03 CA178272-01		
	Organization: DARTMOUTH COLLEGE	

RESEARCH & RELATED Other Project Information

1. * Are Human Subjects Involved? ☒ Yes ☐ No

1.a. If YES to Human Subjects

Is the Project Exempt from Federal regulations? ☐ Yes ☒ No

2. * Are Vertebrate Animals Used? ☐ Yes ☒ No

3. * Is proprietary/privileged information included in the application? ☐ Yes ☒ No

4.a. * Does this project have an actual or potential impact on the environment? ☐ Yes ☒ No

5. * Is the research performance site designated, or eligible to be designated, as a historic place? ☐ Yes ☒ No

6. * Does this project involve activities outside of the United States or partnerships with international collaborators? ☒ Yes ☐ No

Abstract

The randomized, placebo-controlled trial is often described as the gold standard for research involving humans, primarily due to the minimization of bias through randomization, prospective follow-up, and, where feasible, participant and investigator blinding. Effective allocation concealment, combined with blinding of participants and investigators, aims to minimize bias in patient selection, adherence, and ascertainment of outcomes. When blinding is attempted, it can be compromised by noticeable characteristics of the intervention (e.g. pill taste), by flawed protocol design or execution, by health effects (good or bad), or by participants' attempts to identify their intervention (1-3). Some investigators argue for better reporting of blinding effectiveness and others have devised statistical methods to do so (1, 4-11). Their arguments are countered by the idea that a broader effort is needed to identify *any* sources of belief-related bias (13), recognizing that even a successfully blinded subject may hold strong beliefs about which intervention they received and its likely efficacy. Unblinding may be just one of several belief-related sources of bias in RCTs, all of which are poorly understood (14, 15). We already know that without allocation concealment and blinding, we have the potential to exaggerate (16-23), or even reverse a trial's conclusions (24); this bias may affect subjective outcomes more than objective ones (20). Building on previous work that revealed specific expectations among participants at the start of a RCT (12), we propose to investigate the biases associated with participants' beliefs during a large, multi-center, randomized, placebo-controlled trial of calcium and vitamin D in colorectal adenoma chemoprevention. We asked 2813 subjects how effective they thought the study treatments were for specified health effects, and which treatment they would prefer if given the choice. At randomization, and at the middle and end of the trial, we asked them to guess which treatment they had been given, and the reasons for their guess. In parallel, we collected longitudinal symptom and adherence data. We now propose to explore how those beliefs affect subsequent reporting of subjective and objective health outcomes, adherence, and attrition, after adjustments for such factors as randomized treatment, prior symptoms, reasons given by participants for their beliefs, demographic and medical factors. This will be a detailed, longitudinal analysis of the impact of individual subjects' expectations, preferences and beliefs about their assigned treatment on health outcomes and adherence during a large RCT. Our goal is to shift the current paradigm away from its focus on unblinding, by describing directly whether subjects' beliefs about the study treatment can generate bias. If so, we will explore strategies to correct or prevent these biases in future trials. Whether negative or positive, our results will significantly advance RCT methodology, and trigger similar research in other trials. There is currently little published research along the lines that we propose, although the potential importance of biases due to expectations and hunches about treatment efficacy and assignment is increasingly being recognized.

Narrative

Although unblinding is increasingly discussed as a cause of bias in randomized controlled trials, there is very little specific information about the mechanisms through which unblinding and related beliefs cause this bias, and whether it might be possible to avoid or offset these effects in future trials. We will address the important methodologic question of whether a participant's belief about which randomized treatment they were given, along with their beliefs about its efficacy, can cause bias in the reporting of subsequent subjective and objective health outcomes and adherence during a long term colorectal adenoma chemoprevention study.

Facilities and other resources

The project team is based within the Polyp Prevention Study Group (PPSG), an experienced team of epidemiologists, statisticians, analysts, programmers and project coordinators currently engaged in their fourth, multi-center, randomized, controlled, chemoprevention trial. The group is administered through the Section of Biostatistics and Epidemiology in the Department of Community and Family Medicine, at the Geisel School of Medicine at Dartmouth. The team of faculty and staff on the project interact frequently with epidemiologists and biostatisticians within the PPSG and the Department, as well as clinicians and laboratory scientists at the Dartmouth-affiliated Norris Cotton Cancer Center and Dartmouth-Hitchcock Medical Center. The project team already benefits from the expertise and resources of the clinical trials consortium based at the PPSG, which has been based at Geisel for about 25 years. The group includes individuals with substantial expertise in biostatistics, epidemiology, bioinformatics, statistical analysis, data cleaning and data entry.

Computer: The computer systems and facilities to support this study are maintained by the BioInformatics Service Center at the Geisel School of Medicine at Dartmouth, located on the 3rd floor of the EverGreen Center adjacent to the Dartmouth Hitchcock Medical Center. These systems include state-of-the art software and server infrastructure including approximately 45 servers running a variety of operating systems (e.g. Microsoft, Linux). These servers support the projects as well as the administrative functions of the BioInformatics center, including domain name services, email and calendaring services, project management tools, and back-up and archive services. The BioInformatics software infrastructure includes a suite of reusable metadata-driven tools that facilitate rapid development of the study system and good computing practice (e.g. case report form version change auditing, data change auditing, dynamic form generation, data validation). Each staff member working on the project has a personal computer with appropriate software for word processing and data analysis. All machines are password protected and run behind the BioInformatics firewall. Computing support provided by the BioInformatics enter includes server maintenance and security, personal computer maintenance, software maintenance and upgrading, networking services, high speed internet access, back-up and offsite archive of programs and documents, programming support and general system support. Data analysis for the project, conducted primarily by Ms Leila Mott, will use these computer services

Teleconferencing: The project team has worked together successfully to prepare manuscripts describing other work conducted within the Polyp Prevention Study, and has done so very successfully using weekly Skype video conferencing with Professor Peacock who is based in London, UK. The Skype Premium software allows multiple video participants to see each other at once, and is extremely effective.

Office: Office space for the staff, who are part of the Biostatistics and Epidemiology section of the Department of Community and Family Medicine at Geisel, is located on the 3rd floor of the Evergreen Center adjacent to the Dartmouth-Hitchcock Medical Center and the Norris Cotton Cancer Center. The staff has sufficient office space in this building, with access to high speed internet connections, filing space, fax and Xerox machines, and all necessary office supplies. Locked cabinets, locked offices and a locked suite provide security for sensitive data.

PHS 398 Cover Page Supplement

OMB Number: 0925-0001

1. Project Director / Principal Investigator (PD/PI)

Prefix:
* First Name: Judith
Middle Name:
* Last Name: Rees
Suffix: Ph.D.

2. Human Subjects

Clinical Trial? ☒ No ☐ Yes

* Agency-Defined Phase III Clinical Trial? ☐ No ☐ Yes

3. Applicant Organization Contact

Person to be contacted on matters involving this application

Prefix:
* First Name: Jill
Middle Name:
* Last Name: Mortali
Suffix:

* Phone Number: 603-646-3007 Fax Number: 603-646-3670
Email: sponsored.projects@dartmouth.edu

* Title: Director

* Street1: 11 Rope Ferry Rd #6210
Street2:
* City: Hanover
County/Parish: Grafton
* State: NH: New Hampshire
Province:
* Country: USA: UNITED STATES * Zip / Postal Code: 03755-1404

PHS 398 Cover Page Supplement

4. Human Embryonic Stem Cells

* Does the proposed project involve human embryonic stem cells? ☒ No ☐ Yes

If the proposed project involves human embryonic stem cells, list below the registration number of the specific cell line(s) from the following list: http://stemcells.nih.gov/research/registry/. Or, if a specific stem cell line cannot be referenced at this time, please check the box indicating that one from the registry will be used:

Cell Line(s): ☐ Specific stem cell line cannot be referenced at this time. One from the registry will be used.

PHS 398 Research Plan

1. Application Type:

From SF 424 (R&R) Cover Page. The response provided on that page, regarding the type of application being submitted, is repeated for your reference, as you attach the appropriate sections of the Research Plan.

*Type of Application:

☒ New ☐ Resubmission ☐ Renewal ☐ Continuation ☐ Revision

SPECIFIC AIMS

The randomized, placebo-controlled trial has been described as the gold standard for research involving human subjects primarily due to the minimization of bias through randomization, prospective follow-up, and, where feasible, participant and investigator blinding. Allocation concealment, combined with blinding of participants and investigators aims to minimize bias in patient selection, adherence, and ascertainment of outcomes (16-23). Beliefs about the assigned treatment and its efficacy, as well as treatment preference could similarly affect adherence, health behaviours, attrition, and self-reported, subjective health endpoints (25, 26). However, the inter-relationships between these beliefs, blinding and bias are not straightforward. A participant may be blinded but still hold strong beliefs about which intervention they received (12) and how effective it might be (27). Further, participants may change their beliefs during a study (12). For these reasons, the biases associated with blinding and belief are poorly understood, yet are of paramount importance because they have the potential to exaggerate or reverse a trial's conclusions (24).

Controversy surrounds the question of whether trials of an effective medication can ever be successfully blinded (8, 10, 27, 28), and whether unblinding in trials can and should be formally assessed (7, 8, 10, 27-29). **The novelty of our study will be in setting aside the idea of studying unblinding, and instead focusing directly on how participants' beliefs about the randomized treatment (12) affect subsequent symptom reporting, adherence, and attrition, taking into account participants' expectations and the factors underlying their beliefs about the randomized treatment.** The understanding that we gain from this work will help develop strategies to prevent and/or correct for belief-related biases in future RCTs.

Building on our experience from a longitudinal assessment of blinding effectiveness in a previous trial (12), we will investigate the biases associated with subjects' beliefs about treatment allocation and efficacy in a large, randomized, placebo-controlled trial of calcium and vitamin D in colorectal adenoma chemoprevention. At enrolment, we asked how effective subjects thought the study agents were for various health outcomes ("efficacy beliefs"), and which intervention they would prefer if given the choice ("preference"). At the start, middle and end of randomized treatment, we asked participants to guess which intervention they had been given ("allocation belief") and why, and questions about adherence to pill-taking. Our proposed analyses will begin by describing the beliefs data, including correlations between each type of belief, and associations with demographic data. We will then address three specific aims:

Aim 1. Identify symptoms and other predictors of subsequent allocation belief at randomization, year 2, and end of treatment (EOT). We will use regression analyses to predict allocation belief at each time point in relation to preceding symptoms, the reasons given for allocation belief (e.g. symptoms, pill characteristics or hunches), efficacy belief, preference, and interactions e.g. [efficacy belief x symptom]. Randomized treatment will be included as an explanatory variable for allocation beliefs at year 2 and EOT. Strong predictors of allocation belief will be used as adjustment/stratification factors in Aims 2 and 3.

Aim 2. Examine how subjective and objective health outcomes are affected by allocation beliefs, efficacy beliefs, and preference. Whereas Aim 1 assessed predictors of beliefs, Aim 2 assesses how beliefs affect subsequent symptoms. There is evidence that inadequate allocation concealment and blinding cause bias preferentially in trials of subjective outcomes (20). We will use multivariable regression to explore the effects of allocation beliefs on four subjective symptoms (e.g. constipation) and five SF-36 scores (e.g., bodily pain), and on the objective trial endpoint, colorectal adenoma. Explanatory variables will include randomized treatment, efficacy belief, preference, preceding symptoms, reasons for beliefs, and interactions e.g. [allocation x efficacy] beliefs. Examples of our results could include risk of joint pain for those who believed they received Active v. Placebo, restricted to those who at baseline thought the Active pill is effective against joint pain.

Aim 3. Examine how adherence is affected by allocation beliefs, efficacy beliefs, and preference. Adherence in RCTs is associated with better health outcomes due to an intervention, but is also an independent predictor of health (30, 31). We will test whether subjects' beliefs and preferences are related to (i) percentage of pills taken during run-in; (ii) average self-reported pill-taking per week during the study, and (iii) failure to complete the study on treatment. Using an approach similar to Aim 2, we will quantify the independent and combined effects of efficacy beliefs, preference, and allocation beliefs on adherence and drop-out risk, while taking into account predictors of belief such as symptoms identified in Aim 1.

RESEARCH STRATEGY

A. SIGNIFICANCE

The randomized, placebo-controlled trial (RCT) has been described as the gold standard for research involving human subjects (14), primarily due to the minimization of bias through randomization, prospective follow-up, and, where feasible, participant and investigator blinding. When blinding is attempted, it can be compromised by noticeable characteristics of the intervention (*e.g.,* side effects, taste of the pill); flawed protocol design or execution, or participants' efforts to identify their intervention by taste or other types of test.

The rationale for participant blinding is that they may behave or respond differently if they know they are taking the active or placebo treatment, and have expectations based on their confidence in the effectiveness of the treatment or the individual prescribing it (32-37). Knowledge of treatment assignment may bias the experience or reporting of a health event under study (38), it may affect an individual's adherence to treatment, and through these mechanisms may bias the trial's measure of effect. Trials that do not attempt to use double-blinding exaggerate treatment effects by up to 25% compared with trials that attempt to double-blind and trials with inadequate concealment of treatment allocation yield up to 40% higher estimates of treatment effect than trials that make reasonable attempts at concealment (16-23), but this bias may affect subjective, rather than objective health outcomes (20). Unblinding of the investigators who assess health outcomes can also cause bias large enough to reverse a study's results (24). While there is legitimate concern that participant unblinding may affect subsequent health events and adherence, we most likely do not need to worry about bias if unblinding results from treatment efficacy (27). However, subjects may also be influenced by health events unrelated to the primary study outcome, particularly if they expect the study treatment to have certain health effects. These issues complicate the already challenging task of measuring whether a trial was properly blinded, and a standard to report such measurements in the Consolidated Standards for Reporting Trials (CONSORT) was recently removed for lack of a sound methodologic approach (39). Several investigators argue that we should pay more, rather than less attention to the issue of unblinding e.g. (4, 6, 8, 10) and others highlight the various related biases in RCTs, of which unblinding is just one (11, 12, 14, 15).

Our previous analyses of the relations between health endpoints and allocation beliefs measured serially in a smaller trial (12) highlighted the need to consider allocation beliefs in the broader context of preferences and perceived treatment efficacy, and with specific attention to the timing at which symptoms occurred and beliefs were measured. We now propose secondary analyses of data collected during a large, multi-center RCT. Other from a small number of related studies in pain research (25, 26), we are not aware of any detailed, longitudinal analysis of the impact of participants' expectations, preferences and beliefs about their assigned treatment on health outcomes and adherence in a large, long-term RCT.

Our study has the potential to shift the current paradigm away from its focus on unblinding, by examining directly whether subjects' beliefs about the study treatment generate bias. If they do, we will explore strategies to correct or prevent these biases in future trials. A critical task will be to explore the direction of the associations between health effects and participants' allocation beliefs using longitudinal symptom data, and the reasons participants give for their belief. Whether negative or positive, our results will significantly advance RCT methodology. If we do not find significant belief-related bias on subjective and objective health outcomes or adherence, this would provide the first specific evidence from a RCT relevant to the removal of the blinding assessment standard from the CONSORT statement. If we have positive findings, i.e. significant bias due to participants' beliefs, trialists would be obliged to consider methods to prevent or offset these biases in future studies; in the proposed work we will explore such strategies. To maximize generalizability, we will study three outcomes (subjective and objective health endpoints and adherence). We recognize that the results of one study cannot necessarily be generalized to others. However, because there is currently *no* published research along the lines that we propose, a positive finding would be the first of its kind, highlighting the need for trialists to consider the potential for belief-related bias in any trial, an issue that is now largely overlooked.

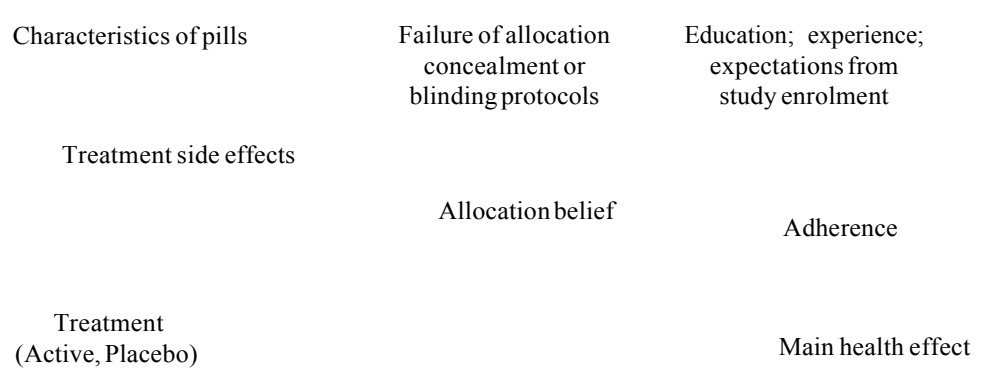
B. INNOVATION

Unblinding may affect adherence, attrition, self-reported subjective symptoms and related health behaviors. Participants' expectations may affect their experience of subjective symptoms, a phenomenon that is part of the "placebo effect". The stronger the belief in a treatment's efficacy, the greater these biases may be, but susceptibility to these effects varies in different individuals and under different conditions.

It has been argued that "a double blind design can work only if the subject is clearly free from the influence of suggestion resulting from accurate information about his medication" (40). This statement illustrates a common misconception about unblinding. Bias does not result simply because some subjects can identify their intervention group correctly. Bias is generated when differences in belief about treatment assignment lead to distorted estimates of treatment efficacy, either directly, or through adherence (Table 1). Some investigators

have tried to measure unblinding in RCTs and others have developed statistical indices based on the correctness of guesses about the assigned treatment (1, 4, 5). We argue that we should be concerned not about unblinding, but about any beliefs about treatment (whether correct or incorrect, differential or non differential, consistent or changing during the trial) that may influence adherence or health endpoints. Others agree that biases in RCTs are not fully understood (15): we "need more and better measures of these bias-generating consequences, whether they arise from the loss of blindness, the development of hunches, or any other cause" (Sackett, 2007) (13). Our novel analyses will consider the independent and combined effects of baseline preferences, expectations of benefits from study pills and allocation beliefs during the trial. Configuring the problem based on beliefs (rather than unblinding) will simplify it and allow us to investigate related biases systematically without worrying that we cannot know if a correct guess reflects true unblinding or random chance (12).

Figure 1. Relations between beliefs, adherence and outcomes in RCTs



In a double-blinded, randomized, placebo-controlled trial of water filtration on gastrointestinal illness, we found that only two weeks after enrolment, 76% of subjects in both treatment arms thought they had been assigned the active treatment (12). "Forced guesses" from those who said they did not know were similarly distributed, suggesting that baseline expectations or hopes exist both among subjects with a clearly stated belief, and among those who initially expressed uncertainty. We also demonstrated differential changes of allocation belief in the treatment groups, but did not explore how symptoms had affected beliefs or vice versa, an issue we realized should be addressed. We used our experience to plan data collection for the proposed study, including baseline preferences and efficacy beliefs, allocation beliefs, adherence measures, subjective outcomes and the trial's primary health endpoint. Our data offer a rare opportunity to explore the potential for participant beliefs to influence participation and generate bias in a large RCT.

Table 1. Hypothetical effects of allocation beliefs on risk ratio

Belief scenarios	Observed RR (95% CI) [1]
1. Non differential beliefs	
a. No beliefs	1.00 (0.92 - 1.09)
b. All believe A[2]	1.00 (0.90 - 1.11)
c. All believe P[2]	1.00 (0.93 - 1.07)
2. Differential beliefs	
a. Group A believes A; Group P believes P	0.67 (0.61 - 0.73)
b. Group A believes A[3]	0.80 (0.73 - 0.88)
c. Group P believes P[3]	0.83 (0.77 - 0.90)
d. Group A believes P; Group P believes A	1.50 (1.37 - 1.64)
e. Group A believes P[3]	1.20 (1.11 - 1.30)
f. Group P believes A[3]	1.25 (1.13 - 1.38)

[1]Risk ratio (95% confidence interval) in Group A (Active) v. Group P (Placebo)
[2]Believe A: believe they received the Active pills; Believe P: believe they received Placebo
[3]Where beliefs in only one group are stated, those in the other hold no beliefs

In a hypothetical example, 2000 participants are randomized with equal probability to identical-looking pills that we will call Active and Placebo. Neither pill has any biological effect on depression, which usually will affect half the participants, but those who think they have the Active pill are 20% less likely to report depression and those who think they have the Placebo are 20% more likely to do so. Table 1 shows how allocation beliefs might cause bias under various conditions e.g. study-wide unblinding (2a), makes the Active pill seem protective. Non-differential beliefs (1a-c) did not bias the (already null) RR estimate. Generally, if beliefs cause non-differential under-reporting of outcomes, the RR is unbiased, and if beliefs cause non-differential over-reporting of outcomes, the RR is biased towards the null. These situations are analogous to non-differential misclassification of the outcome with perfect specificity/imperfect sensitivity and perfect sensitivity/imperfect specificity, respectively (41, 42). Differential beliefs may generate bias in either direction.

C. APPROACH

C1. Parent study: We will use data from the on-going Polyp Prevention Trial (Table 2). This is a randomized, placebo-controlled, multi-center chemoprevention trial. Participants, enrolled 2004-2008, had at least one large bowel adenoma removed and no remaining polyps in the bowel after a recent, complete colonoscopic examination. Eligible subjects were 45-75, in good general health with no contraindications to study treatment, and no familial colorectal cancer syndromes or serious gastrointestinal disease. A detailed health questionnaire was administered at enrolment, followed by a 56-84 day blinded, placebo run-in phase to identify and exclude, before randomization, subjects who were unlikely to adhere to study procedures. Eligibility for randomization was determined after the run-in by blinded coordinators via a web-based system. Consumption of ≥80% of pills via self-reported pill count was a requirement for randomization. Study treatment was given for 3 or 5 years depending on the colonoscopy screening interval prescribed by each participant's gastroenterologist. Block randomization was stratified by center, surveillance interval (3 or 5 years), and sex in a modified 2x2 factorial design to vitamin D_3 (1,000IU/day), calcium carbonate (1,200mg/day elemental calcium), both, or placebo (4-arm study). Women who declined to forego calcium supplementation were randomized to calcium + vitamin D or calcium (2-arm study). Data were collected with web-based surveys at enrolment, randomization, every 6 months, and at the end of treatment (EOT). Extremely high completion rates were recorded. At year 2 and EOT, subjects were asked to mail back paper "quality control" questionnaires including questions on beliefs; 72 and 70% of these surveys were returned, respectively. SF-36 surveys were completed at the start and EOT.

Table 2. Number of completed questionnaires for study

Time	Enrollment	[1]Random-ization	[2]Every 6 months	Year 2	[3] treatment
Beliefs:					
Preference	2813				
Efficacy	2813				
Allocation		2502		1626	>1580
Reasons		2502		1626	
Health Outcomes					
[4]SF-36 scores	2813				>1870
Constipation	2813	2502	>2120	>2120	>1921
Chronic fatigue	2813	2502	>2120	>2120	>1921
Muscle pains	2813	2502	>2120	>2120	>1921
Bone pains	2813	2502	>2120	>2120	>1921
Adenoma	2813				>1921
Adherence:					
Pill count	2813				
Pills taken/week			>2120	>2120	>1921
Surveys done:	2813	2502	>2120	1626	>1580

[1]2259 were randomized after completing this questionnaire
[2]numbers varied from 2120 to 2204 in years 1-2
[3]study is ongoing; 234 more subjects are expected to give data through Jul 2013

C2. General approach: Our broad goal is to provide information about the relative impacts of participants' beliefs in the causal pathway shown in Figure 1. To do this, we must first describe and understand the factors that predict allocation belief early in the study, including efficacy beliefs and preference, *prior* symptoms, and the reasons stated by subjects for their belief (Aim 1). We will then assess the impact of allocation beliefs on rates of *subsequent* symptom occurrence (Aim 2), and adherence (Aim 3), adjusting for important factors identified during those preliminary analyses, and randomized treatment effect (even though we chose subjective outcomes not proven susceptible to vitamin D or calcium supplementation).

C3. Statistical analyses: In general, our statistical analyses will begin with descriptive summaries of the data (means, medians, standard deviations, frequencies and proportions) relevant to our hypotheses (Table 2). These summaries will compare groups of subjects defined by allocation beliefs. Where useful, we will use graphical methods to illustrate apparent trends in the data (*e.g.*, box plots over time to compare the allocation belief groups). We will then fit appropriate regression models to the data to evaluate joint predictors of each outcome variable, adjusting for randomized treatment or, where possible, doing stratified analyses. For example, we will evaluate the association between allocation beliefs and quality of life SF-36 scores. Because many of our outcome variables will be measured serially, our statistical analyses will need to account for the longitudinal aspects of the data. Initially, we will evaluate associations at each time point separately. We will then consider how baseline factors affect changes in outcome from baseline to year 2 and EOT. Finally, we will fit models for longitudinal data to evaluate effect modification over time. We will use appropriate linear and/or logistic models that will be fit using generalized estimating equations (GEE).

The factorial design of this trial would allow us to consider the impact of allocation beliefs on outcomes/adherence separately for each intervention agent (calcium, vitamin D), but in the 4-arm study, we will pool those who believe they are receiving *either* calcium *or* vitamin D only, to optimize power in those smaller groups, and consider "either agent" as an Active treatment belief. For women already taking calcium supplementation in the 2-arm study, we will treat their calcium supplementation as a baseline medication that

is not considered as an Active treatment. In this way, allocation beliefs will be classified as "Believe Active" (A) for any of the active treatment options in the 4-arm study or calcium + vitamin D in the 2-arm study; "Believe Placebo" (P) for placebo in the 4-arm study; calcium in the 2-arm study; or "Don't Know (DK). We will use these classifications in the analyses of all subjects and repeat analyses separately for those in the 4-arm and 2-arm studies. We will assess the distribution of forced guesses of those who said DK. If, as seen previously (12), their distributions are similar to those who initially make a guess, we will conduct sensitivity analyses pooling forced guesses with the corresponding "forced" Active or Placebo guess, for greater statistical power. We will conduct other sensitivity analyses as needed, to consider the impact of including the discordant categories (people who believed they were given calcium only, vitamin D only) among those who believed Active.

In general, we will consider a 2-sided p-value less than 0.05 to be statistically significant; however, we will consistently present findings using estimates with confidence intervals and p-values. We will not make adjustments for multiple testing. When using efficacy beliefs as an adjustment factor, we will use the efficacy belief specific to the health condition of interest. For example, in an analysis relating to general health scores, we would choose only the efficacy belief concerning general health.

Most of our analyses in Aims 2 and 3 will follow the same general progression, and it is useful to consider an illustrative example in relation to beliefs and general health SF-36 score:

Step 1. We will initially look at the association between allocation beliefs and general health SF-36 score at baseline. We will conduct a descriptive analysis by estimating the mean general health score according to allocation belief group. The results may be tabulated or plotted graphically.

Step 2. We will then fit a multiple regression model to the baseline data. The dependent variable will be general health score. The independent variables will include allocation belief group and adjustment factors, including efficacy beliefs specific to general health, and preference.

Step 3. Steps 1 and 2 will be repeated using general health SF-36 score at EOT including, as predictors, changes in allocation belief up from randomization to EOT, and randomized treatment.

Step 4. We will look at the association between baseline factors and general health SF-36 score at EOT in two ways. First, we will assess changes in general health score from baseline. Second, we will model the follow-up measures directly while using the baseline measure as a covariate. In the event that these analyses yield conflicting results, we will focus on the latter. Explanatory variables will include allocation belief, time point, [time point x allocation belief] interaction, and adjustment factors. *For analyses of data after randomization (but not during run-in), we will also include randomized treatment in the models.*

Below we elaborate on the statistical methods for the specific hypotheses associated with each aim. Where we will repeat analyses using data from a different time point, the latter is shown in square brackets.

Aim 1. Identify predictors of allocation belief at randomization, year 2, and EOT.

We will begin by describing the beliefs data overall, and in relation to subject characteristics, and then explore the most important predictors of allocation beliefs at each time point (Table 2). While blinded to randomized treatment, we will create a systematic classification of the reasons subjects provided (in categorical and text response format) for their allocation beliefs. Each reason will be classified dichotomously for its relation to: health effects; pill characteristics; luck/hunches; and reasons inconsistent with a randomized trial's procedures (Examples from our data: "Because I am older, I should have the Active pill"; "Why would they give me the pill that wouldn't do anything?"). When possible, reasons will be further classified as negative, positive, and neutral/unspecified: e.g. gastrointestinal symptoms (adverse, absent/improved, or unspecified); pill taste (bad, good, or not specified). In addition to stated reasons, we will explore other "health reason" categories: (i) symptoms (constipation, chronic fatigue, bone pains and muscle pains) recorded every 6 months, (ii) hypovitaminosis D diagnosed independently of the study by the medical providers of 114 subjects. We will use these categories to understand the common themes underlying allocation beliefs during the placebo run-in phase and after randomization. For subjects whose reasons reveal a misunderstanding of equipoise and randomization, we will describe their group characteristics (age, sex, etc.), so that future trials involving such individuals might consider providing clearer explanations at enrolment to address this issue.

We will then use multivariable regression analysis to predict allocation beliefs at randomization [year 2, EOT]. Explanatory variables will include the reasons stated by subjects for those beliefs; preference; efficacy belief; *prior* symptoms recorded in a separate questionnaire; interactions e.g. (efficacy belief x prior symptom); randomized treatment; demographic factors; education; baseline use of multivitamins; calcium supplements, alcohol consumption, smoking, and center. Important predictors of allocation belief (e.g. symptoms during run-in) will be considered as potential adjustment or stratification factors in Aims 2 and 3.

Hypothesis 1. Allocation beliefs change during the course of a double-blinded trial. Having described the distributions of beliefs, we will next identify *changes* in allocation beliefs from randomization to year 2 [EOT]. If beliefs change, do they change differentially by randomized treatment, and do changes in beliefs result primarily from health effects, pill characteristics, or non-specific factors like guesswork and luck? We will define beliefs as consistent Active (A-A), consistent Placebo (P-P), switching to Active (P-A, DK-A), or switching to Placebo (A-P, DK-P), and examine these by treatment assignment using chi-square tests. We will use regression models to identify the factors associated with these changes of belief categories as described previously. This will help elucidate the factors driving the causal pathway in Figure 1.

Aim 2. Examine how subjective and objective health outcomes are affected by allocation beliefs, efficacy beliefs, and preference.

Whereas Aim 1 assessed how symptoms affect beliefs, here we will assess how beliefs affect subsequent symptoms. We will use multivariable regression to quantify the effects of allocation beliefs on selected subjective symptoms and on the objective trial endpoint, colorectal adenoma. Where needed, we will stratify on or adjust for predictors of allocation belief such as randomized treatment.

Hypothesis 2. Allocation beliefs at randomization [year 2, EOT] predict the occurrence of each of five specified categorical health effects, after adjustment for a history of that same health effect. We will begin by assessing as dependent variables in separate regression models, any report in the two 6-monthly questionnaires after randomization of: **constipation, chronic fatigue, bone pains** and **muscle pains**. Explanatory variables tested in each model will include baseline efficacy beliefs for that symptom; preference; allocation belief, changes in belief, interaction terms e.g. (efficacy belief x allocation belief); randomized treatment; reasons for allocation belief; prior symptom during run-in; adherence during run-in; pre-trial use of multivitamins, age, sex, education and race. We will present odds ratios for each symptom-treatment association, stratified by belief (Active, Placebo, Don't Know). Similarly, we will develop models describing the role of allocation beliefs at these three time points in predicting **adenoma** recurrence. Hypothesis 2 will provide critical information about the importance of belief-related bias on subjective and objective health outcomes.

Hypothesis 3. Allocation beliefs at randomization [year 2, EOT] predict a significant change in each of five SF-36 scores during the study, independently, or by interaction with efficacy beliefs about that health effect, after adjusting for baseline score and randomized treatment. Hypothesis 3 is analogous to Hypothesis 2, but for continuous variables from SF-36 data on **general health**, **physical functioning, bodily pain, vitality,** and **mental health,** at baseline and EOT, will allow us to analyze cross-sectional SF-36 scores, and changes in scores. There is little consensus on the definition of a clinically significant change in SF-36 score (43-45). Because many quality of life investigators disagree on what constitutes a clinically relevant change, we will base our analyses on the scores themselves rather than their clinical relevance, presenting statistical "score" data that can be readily interpreted using an investigator's own preferred definition of clinically relevant change. We will also illustrate our results using a two-step approach that has been described as one solution to this issue (46). We will analyze the selected SF-36 scores directly and as changes from baseline to EOT. When assessing the measure directly, the baseline score will be used as a covariate. Explanatory variables will be included as described for Hypothesis 2.

Aim 3. Examine how adherence is affected by allocation beliefs, efficacy beliefs, and preference.

Adherence in RCTs is associated with better health outcomes due to an intervention, but is also an independent predictor of health. We will develop regression models to predict subjects' adherence, using, as explanatory variables, preferences and efficacy beliefs (both potentially modifiable factors), allocation beliefs at different time points, and demographic and medical factors. Our adherence measures are self-reported pill-counts after run-in, estimates of pills taken per week, and duration of any gaps in pill-taking (recorded every 6 months); some validation will be attempted using the start dates for each new pill bottle.

Hypothesis 4. The proportion of pills taken during the blinded, placebo run-in is associated with allocation beliefs at randomization, after adjustment for efficacy beliefs and preferences at baseline. The number of pills left in the bottle after the placebo run-in was ascertained via self-report, to establish eligibility for randomization. We will begin with descriptive summaries, including the proportion of assigned pills taken. This outcome is likely to have a skewed distribution; if so, we will attempt to fit a linear model either to the outcome directly, or if appropriate, to a transformation of adherence (e.g., log transform). If it is not possible to achieve a good linear fit, we will investigate other techniques such as smoothing methods. The dependent variable will be proportion of pills taken. Independent variables will include allocation belief at randomization; baseline efficacy

beliefs and preferences; randomized treatment and others listed in Aim 1.

We will repeat these analyses while deriving the dependent variable from treatment adherence averaged over the first two years of treatment. Randomized treatment will be included as an adjustment factor. Whereas allocation beliefs at randomization likely reflect subjects' experiences with blinded placebo consumption during run-in and may predict future adherence, allocation belief at year 2 will reflect the change from placebo to randomized treatment, and health over two years. Finally, we will develop models for a dichotomous dependent variable for completion of the trial on study pills.

C4. Sample Size Justification

In the parent study, 2,813 subjects were enrolled and 2,259 randomized. For analyses of adherence during run-in, we will use all 2,813 enrollees. For other analyses, we will use the 2,259 randomized subjects. In the 4-arm study, 26% had no preferred treatment assignment; 62% preferred both calcium + vitamin D, 10% preferred one agent, and 2% preferred placebo. Of those enrolled (21%) in the 2-arm study, 77% would prefer both calcium + vitamin D; 17% had no preference. In baseline efficacy beliefs, between 10-42% said they did not know if the study agents would prevent polyps or bodily pain, improve mood or general health, or cause constipation; 30-88% believed they were very or somewhat likely to be effective; and 20% believed they were likely to cause constipation. After run-in, 45% subjects in the 4-arm study guessed Active (23% both agents, 22% one agent), 30% Placebo, and 25% DK and, in the 2-arm study, 37% Active, 37% Placebo and 26% DK. Existing data are shown in Table 2, with up to 234 additional questionnaires expected by July 2013.

For continuous outcome measures, we will have 90% statistical power to detect a difference of 0.09 standard deviations (SD) at baseline, and 0.11 SD at year 2. These calculations are based on t-test comparisons of allocation beliefs. For multivariable analyses sample size calculations are approximate but assuming that each additional predictor variables requires a 10% increase in sample size, then these sample sizes are sufficient to detect small adjusted differences of 0.15 SD at baseline and 0.20 SD at Year 2 or EOT in models containing 20 additional predictors (degrees of freedom). For binary outcomes, at baseline, we will have 90% statistical power to detect an odds ratio <1.3 at baseline, year 2 and EOT. For multivariable analyses assum-ing the outcome has prevalence of just 10% of the total sample, model estimates will be reliable with 28 predictors for baseline data models and 16 predictors for year 2 models. Hence our study sample is sufficient for the analyses described.

C5. Limitations

(i) The main challenge in our proposed work is the complexity of the causal pathway including beliefs, symptoms, outcomes and adherence. We will address this by studying important sections of the causal pathway separately; using statistical adjustments e.g. for treatment effect; taking into account the timing of beliefs and symptoms; and conducting stratified analyses where possible. (ii) We may not obtain accurate data from subjects about the reasons for their beliefs; we recognize that subjects may try to give "helpful" replies, and are unlikely to volunteer that they tested their study pills (1). Our approach addresses the limitations of the "reasons" data, by using broad categories of reasons such as symptoms or pill characteristics. (iii) Our adherence measures, self-reported pill counts and weekly consumption estimates, are less accurate than methods like electronic monitoring in pill packaging (47), but logistics in the parent study precluded more complex options. We will partially validate the data through comparisons with dates bottles were started. Errors in pill counts would reduce power to identify a relation between adherence and beliefs, should one exist, but our large sample size gives sufficient power to detect very small changes in adherence related to differences in beliefs. (iv) Calcium and vitamin D have few, if any, side effects. This will prevent us studying how true side effects cause bias but it will facilitate the study of perceived, subjective side effects. An early subset of our data on reasons for allocation belief showed that 319/884 (36%) of "text" (non-categorized) responses were health related, indicating the likely importance of health effects. (v) Generalizability is a potential issue, but we argue that in the absence of any detailed studies on this topic, positive findings will highlight the issue of belief-related bias in *any* trial. A finding that beliefs are unrelated to symptoms or adherence in our study would be important negative evidence in support of the RCT as gold standard, indicating that allocation beliefs due to symptoms or pill characteristics, etc., cause no significant bias in this chemoprevention trial. This would serve as a baseline for future studies of a similar kind, using agents whose side effects or efficacy are very obvious to participants.

C6. Timeline

Months 1-6: Data cleaning, classify reasons for beliefs, descriptive data analyses, begin regression models
Months 7-19: Regression analyses; sensitivity analyses; stratified analyses as needed
Months 20-24: Manuscript preparation & submission; presentation at meeting

HUMAN SUBJECTS RESEARCH

Scenario B. Non-exempt Human Subjects Research.

1. Protection of Human Subjects

1.1. Risks to Human Subjects

a. Human Subjects Involvement, Characteristics and Design

This research will use epidemiologic data collected from subjects enrolled in the Vitamin D/Calcium Polyp Prevention Study (VCPPS). This on-going multi-centered clinical trial (described in more detail below) is coordinated at the Geisel School of Medicine at Dartmouth under the direction of the principal investigator, Dr. John A. Baron, and is funded by the National Cancer Institute. Enrollment began in July 2004 and ended in July 2008. A total of 2,813 subjects were enrolled and 2,259 subjects were randomized in the trial. As of October 2012, all but 234 participants have completed their follow-up colonoscopy and ended their participation in the trial; the remaining subjects are anticipated to complete their follow-up colonoscopies and end their participation by July 2013. For the current proposal, existing data from the trial will be analyzed by the PPS team at its headquarters at Dartmouth College.

The Vitamin D/Calcium Polyp Prevention Study is a double-blind, placebo-controlled trial of vitamin D and/or calcium supplementation for the prevention of large bowel adenomas that is currently on-going. The subjects were recruited from the clinical services and associated practices of the eleven participating institutions: the Cleveland Clinic, the Dartmouth-Hitchcock Medical Center, the University of Colorado School of Medicine, the University of Iowa School of Medicine, the University of Minnesota School of Medicine, the University of North Carolina School of Medicine, the University of Southern California School of Medicine, Emory University, the University of South Carolina, the University of Texas and the University of Puerto Rico. In order to qualify for the trial, potential participants must have had one or more histologically verified neoplastic polyps removed from the large bowel within 4 months of study entry with the entire bowel visualized endoscopically and judged free of remaining polyps. In addition, they had to have anticipated colonoscopic follow-up either 3 or 5 years after the qualifying colonoscopy. Subjects had to be between the age of 45 and 75 and in good general health, with no diagnosed narcotic or alcohol dependence and no conditions interfering with absorption of the study agents or increasing the risk of toxicity from supplemental intake of calcium or vitamin D. These conditions include a history of kidney stones, granulomatous diseases, hyperparathyroidism or elevated levels of serum calcium, vitamin D or creatinine. They must also not have any conditions indicating a need for calcium or vitamin D supplementation, including osteoporosis or low levels of serum calcium or vitamin D. Other exclusions included familial colonic polyposis syndromes, ulcerative colitis, Crohn's disease or a history of invasive carcinoma in any colonic polyp removed. No subpopulations were excluded from recruitment, and no vulnerable populations were involved.

Subjects were randomized in a modified 2 x 2 factorial design to vitamin D (1000 IU/day), calcium carbonate (1200 mg elemental calcium/day), both agents, or placebo only. Women who declined to forego calcium supplementation were randomized only to calcium alone or to calcium plus vitamin D. Randomization was stratified by gender, study center of recruitment, and anticipated follow-up interval (see below), and was conducted separately for female subjects randomized only to vitamin D. Subjects agreed to avoid taking study agents outside the trial and were initially be observed in a 3-month placebo run-in period to identify (and exclude before randomization) subjects likely not to adhere to the study regimen. Blood levels of calcium and creatinine were obtained at baseline and 1 year after randomization, as well as 3 years after randomization for subjects with a 5-year surveillance cycle. In addition, 25-hydroxy-vitamin levels were measured at baseline and 1 year after randomization, 3 years after randomization (for subjects with a 5-year surveillance cycle), and at end of treatment (near the time of the follow-up colonoscopy). Every six months after randomization subjects completed a questionnaire regarding compliance with study agents, use of medications and vitamin/mineral supplements, illnesses, hospitalizations, and dietary intake of calcium and vitamin D. In addition, a paper survey was sent at the end of year 2 and the end of treatment, to ask quality assurance questions that included the allocation belief data to be used in this study. All activities described here are approved by the IRBs at Dartmouth and each of the 11 clinical centers. The endpoint of the study will be new adenomas detected on

follow-up colonoscopy, scheduled to occur either 3 years or 5 years after the qualifying examination, depending on the follow-up interval recommended by each patient's endoscopist. In the primary analyses of the parent study, the occurrence of new neoplastic polyps in the interval between randomization and the follow-up exam will be compared between treatment groups.

b. Sources of Materials

Questionnaire data:
For the parent study, participants completed (or will complete) study-specific questionnaires at enrollment, and every six months until completion of randomized treatment. The questionnaires address demographic characteristics, lifestyle habits (including physical activity and smoking status), pill-taking compliance, health status and history, hospitalizations and other medical events or diagnoses, and the use of nutritional supplements and medications. Data was (or will be) collected by real-time data entry into web-based forms by study coordinators during in-person or telephone interviews with subjects. Participants completed a telephone survey after the 3 month run-in phase, to establish their eligibility for randomization. They also completed (or will complete) a Food Frequency Questionnaire at enrollment and an SF-36 Health Survey at enrollment and end of treatment. Beliefs data were collected during the enrolment survey and after run-in, according to the schedule described above. In addition, paper surveys containing quality assurance questions were completed and mailed back by participants at the end of year 2 and the end of treatment.

Pathology and medical records: Pathology and endoscopy reports were (or will be) collected by research coordinators at the each of the participating institutions to document polyp diagnoses and to obtain medical confirmation of endpoints. Coordinators also collected (or will collect) existing medical records to confirm cancer, stroke, coronary heart disease and other significant diagnoses or hospitalizations. Cause of death was (or will be) documented using discharge summaries from final hospitalizations and death certificates. Coordinators borrowed (or will borrow) pathology slides from colonoscopies from the local pathology labs and submitted them for review by the study pathologist. Copies of medical records were (or will be) collected on paper and identified either with subject names or subject ID numbers. Medical record data was (or will be) subsequently entered into the central study database.

Linkage to subjects: The central study database links all of the information collected about the subject to the subject's study ID number and to identifying information including the subject's name and address, which are used for the purpose of shipping the study agents to the participants from the study pharmacy located at Dartmouth Medical School. Access to the central study database is restricted to informatics and database management staff at the Dartmouth Medical School.

c. Potential Risks

Social risks could occur if subject confidentiality is violated. The likelihood of such events is very small; all identifying information is rigorously protected, and the analytic dataset for use in this study will contain no sensitive genetic data. It will contain demographic data, data on subjective health outcomes and polyp occurrence, data from the SF-36 questionnaire, and the participants' beliefs about treatment assignment.

1.2. Adequacy of Protection Against Risks

a. Recruitment and Informed Consent

Recruitment was completed in July 2008; a total of 2,813 subjects were enrolled in the trial. The informed consent process began during the first patient contact and continues throughout the subject's participation in the study. During the initial in-person appointment and during visits to discuss the study, a trained clinical center coordinator conducted an informed consent discussion, focusing on the research nature of the study, its purposes, expected duration of participation, study procedures, possible risks or discomforts, possible benefits, and confidentiality. It stressed the voluntary nature of participation and the right to withdraw. Willing patients documented their consent to participate by signing and dating an informed consent document presenting this information in written form. The discussions took place at the clinical center in a private, comfortable environment free of interruptions. Written documentation of informed consent was obtained from all subjects at

enrollment.

The protocol for the study was reviewed and approved by the research ethics committees (IRBs) of the participating institutions. Ongoing IRB review ensures that the informed consent obtained from subjects during their participation in the trials remains adequate for the new research that is proposed or, if inadequate, that new consent is obtained.

b. Protection Against Risk

Protection from social risks: To protect subject confidentiality, the following steps are taken:

- All analysis datasets generated for this research will be de-identified and will not contain any subject names or addresses.
- All identifying study data will be stored in locked filing cabinets or in secure, password-protected computer files.
- All communications and data transfers regarding this research will use only subject or specimen study identification numbers.
- All staff members will be able to access only the minimum necessary study data required to complete their tasks.
- The results of this study will be analyzed in a statistical way only, and no individuals will be identified in any reports or publications.

1.3. Potential Benefits of the Proposed Research to Human Subjects and Others

This research will investigate the effects of beliefs in generating bias during a randomized trial. If we can identify scenarios in which blinding leads to bias, we will explore strategies to avoid or offset those biases in future trials, a result that would provide general benefits by improving the quality of future trials. This study may also identify subgroups of individuals who have misunderstood the concepts of randomization; if so, knowledge of the people affected and the issues that were not well understood would offer an opportunity to improve the process of informed consent in future trials.

1.4. Importance of the Knowledge to be Gained

The improvement of trial methodology as described in 1.3 will benefit many future trials, from the perspective of their providing more reliable data to the research community to benefit public health, prevention activities and clinical medicine.

Inclusion of Women and Minorities

All eligible adult women were invited to participate in the clinical trial at each of the participating clinical sites, yielding 37% women in the study. The high proportion of male participants is partly attributable to the participation of Veterans Affairs Hospitals at some of the clinical centers. In addition, both autopsy and screening studies show a somewhat higher prevalence of adenoma among men. Women with child-bearing potential who agreed to practice effective birth control were included in the study.

All eligible individuals from all minority groups were invited to participate in the clinical trial at each of the participating clinical sites. The demographic breakdown of study participants is 81% white, 10% black, 2% Asian, and 6% other. In addition, 7% of enrolled subjects are of Hispanic ethnicity.

Little epidemiological research has been conducted on population distribution of colon adenomas, but the available data suggest roughly similar prevalence in whites and blacks. Our group's previous polyp prevention studies also indicate similar polyp recurrence rates by ethnic group. Because of the trial's eligibility criteria, requiring that potential subjects undergo a complete colonoscopy within three or four months before the intake appointment, it was not practical to target the general community for recruitment. Study coordinators focused their efforts on the population receiving care from collaborating gastroenterology clinics.

As we do not expect to find clinically important gender and/or race/ethnicity differences in the intervention effect on the primary endpoint, the trials were not specifically powered to detect such effects. Nevertheless, in the present study, we will explore potential difference in beliefs and related bias. In particular, we will consider males separately from females, and we will consider the major racial/ethnic minority groups separately from whites.

Targeted/Planned Enrollment Table

See link below for updated notice:
http://grants.nih.gov/grants/guide/notice-files/NOT-OD-14-086.html

Inclusion of Children

Children (individuals under 21 years of age) were excluded from the trial because the research topic in the parent study (sporadic colorectal cancer) is not relevant to children. The age criterion for enrollment in the trial was 45 to 75 years of age at the time of the intake appointment. The occurrence of colon polyps in a child or young adult is a marker of an inherited (familial) colon cancer syndrome, which was an exclusion criterion for each of these studies. Hereditary colorectal cancer syndromes occur by a different mechanism than sporadic colorectal cancer. Sporadic colorectal adenomas typically develop in the 4^{th} or 5^{th} decade of life, at the earliest.

Resource Sharing

Data from this research study will be made available in accordance with the NIH Data Sharing Policy (http://grants.nih.gov/policy/data_sharing), to researchers in both the private and public sector for free or for a nominal charge to the extent permissible under our subject consents, IRB approvals, and local, state and federal laws and regulations, including the Privacy Rule. Only anonymized data files will be provided. Pending third parties' rights, we will transfer data under appropriate documentation monitored by Dartmouth's Technology Transfer Office. Generally, we will include a requirement that new data generated by recipients become a part of the publicly available dataset. Requests for data should be directed to the study's Principal Investigator, who will evaluate the requests in a timely manner, in association with the study's Executive Steering Committee. If data sharing is not possible for any reason, the investigator will be notified with a full description of the reason. If the investigator is not satisfied, (s)he may ask that the request be submitted to the study's Safety and Data Monitoring Committee for impartial review and assessment.