Published in final edited form as:

Contemp Clin Trials. 2013 March ; 34(2): 356–363. doi:10.1016/j.cct.2013.01.003.

Enhancing Cancer Screening in Primary Care: Rationale, Design, Analysis Plan, and Recruitment Results

David M. Murray¹, Mira L. Katz^{2,7,8}, Doug M. Post^{6,8}, Michael L. Pennell³, Gregory S. Young⁵, Cathy M. Tatum^{7,8}, and Electra D. Paskett^{4,7,8}

¹Biostatistics and Bioinformatics Branch, Division of Epidemiology, Statistics, and Prevention Research, Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health

²Division of Health Behavior & Health Promotion, College of Public Health, The Ohio State University, Columbus, OH

³Division of Biostatistics, College of Public Health, The Ohio State University, Columbus, OH

⁴Division of Epidemiology, College of Public Health, The Ohio State University, Columbus, OH

⁵Center for Biostatistics, The Ohio State University, Columbus, OH

⁶Department of Family Medicine, The Ohio State University, Columbus, OH

⁷Division of Cancer Prevention and Control, Department of Internal Medicine, College of Medicine, The Ohio State University, Columbus, OH

⁸Division of Population Sciences, Comprehensive Cancer Center, The Ohio State University, Columbus, OH

Abstract

Colorectal cancer (CRC) is the third leading type of cancer and the third leading cause of cancer death in the United States. National policy-making organizations recognize and support a variety of CRC screening strategies among average-risk adults aged 50 and older based on strong evidence showing that screening decreases mortality from CRC and can also reduce the incidence of the disease. The goal of this study was to test a multi-level stepped intervention to increase CRC screening rates. We used a group-randomized trial design where the units of assignment were clinics and the units of observation were eligible patients receiving care at those clinics, with stratified random assignment of clinics to study conditions. The primary analysis was planned as a mixed-model logistic regression to account for the expected positive intraclass correlation associated with clinics. Our recruitment experience reflected the difficulties of conducting research in the real world where changes in economic conditions, staff turnover/layoff, inadequate medical records, and poor acceptance of research can significantly impact study plans. It demonstrated the problems that can emerge when procedures used in the study depart from those used in the pilot work to generate parameter estimates for power analysis. It also demonstrated the

Corresponding author: David M. Murray, PhD, Adjunct Investigator, Biostatistics and Bioinformatics Branch, Division of Epidemiology, Statistics, and Prevention Research, Eunice Kennedy Shriver National Institute of Child Health and Human Development, National Institutes of Health, 6100 Executive Boulevard, Room 2B03, Rockville, MD 20892; (301) 496-1508 phone; (301) 480-7660 fax; david.murray2@nih.gov.

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Disclosures: The authors report no conflicts of interest.

importance of allowing for attrition at the group and patient levels so that if recruitment falls short, it is possible to maintain adequate power with only a slight increase in the detectable difference. This experience should assist others planning group-randomized trials, whether in cancer screening or in other areas.

Keywords

colorectal cancer; screening; group-randomized trial

INTRODUCTION

Colorectal cancer (CRC) is the third most prevalent type of cancer and the third leading cause of cancer death among men and women in the United States (U.S.) [1, 2]. Moreover, CRC incidence (males: 60.0 vs. 57.1 per 100,000; females: 44.5 vs. 42.4 per 100,000) and mortality (males: 23.6 vs. 21.2 per 100,000; females: 16.8 vs. 14.9 per 100,000) rates are higher in Ohio than in the U.S. [1]. National policy-making organizations recognize and support a variety of CRC screening strategies among average-risk adults aged 50 and older [3–5] based on strong evidence showing that screening decreases mortality from CRC and can also reduce the incidence of the disease [3–7].

CRC screening guidelines can be challenging to follow because of the various screening tests that may be recommended alone or in combination, the varying interval for screening depending on the test, the confusion about which test to recommend when, and the barriers to screening that patients may experience [8]. As a result, the use of CRC screening tests is low. For example, in 2010, 54.9% of men and 52.0% of women in the U.S. were within guidelines for CRC screening [9]. The goal established by Healthy People 2020 is 70.5% [10]. Since approximately 85% of adults visit a physician at least once every 2 years [11] and most recommended CRC screening tests require the involvement of a physician, the primary care visit can serve as the ideal time to initiate screening. Research has shown that the primary care physician's advice is the most important motivator for preventive health behavior and specifically for CRC screening [12]. Thus, effective strategies are needed in primary care settings to increase CRC screening rates.

Previous studies have tested interventions in primary care settings to increase CRC screening rates [13] using interventions aimed only at one “level” of influence, ie patients, providers, or system-level factors, or more recently, testing interventions directed at multiple levels. Each study has had limited sustained improvement in CRC screening rates suggesting that some patients may respond directly to a provider's recommendation for CRC screening while others may need more individualized approaches. Some additional strategies may include: 1) providing information about the importance of CRC screening; 2) changing any negative attitudes about CRC screening tests; 3) improving self-efficacy to complete CRC screening; 4) reducing barriers to CRC screening; and 5) providing encouragement and support for CRC screening test completion. No study, however, has combined these intervention approaches, i.e. multiple level and then increasing in intensity only for those who still have not been screened. This strategy is promising as it allows participants to receive the amount of intervention needed to change their behavior.

The goal of this study was to test a multi-level stepped intervention to increase CRC screening rates and to find the most effective strategies for patients who may need increasingly intense levels of encouragement to complete a screening test. Our multi-level CRC screening intervention focused on increasing CRC screening rates among average-risk adults (age 50 and older) who use primary care clinics. The novel factors of the intervention

included using increasingly intensive intervention components that focus on recalcitrant patients in a stepped approach. The first phase included a clinic and provider-level intervention, followed by a patient-level phase that included 3 steps: a letter from the physician, a CRC screening barriers counseling telephone call, and an in-person counseling visit. This paper describes the design of the study, including the intervention program, the plan for the primary analysis, and the results from the recruitment of clinics and patients.

METHODS

Research Design

Because a major component of the intervention was delivered at the clinic-level, a randomized clinical trial with randomization of individual patients was not possible. Instead, we used a group-randomized trial design [14] where the units of assignment were clinics and the units of observation were eligible patients receiving care at those clinics. In particular, we used a nested cohort design [14] with stratified random assignment of clinics to study conditions. In a cohort of patients from each clinic, medical record reviews were scheduled at baseline, 12 months, and at the end of study while patient surveys were scheduled at baseline and at the end of study to assess the effects of the intervention on completion of CRC screening.

Participating Clinics and Patients

In the grant application submitted to support the study, we identified 12 clinics from the OSU Primary Care Network that agreed to participate in the study. Ten were family practice clinics and two were general internal medicine clinics. All had participated previously in research projects and we received letters of support from the physicians who oversaw those clinics. All 12 clinics had access to CRC screening modalities and referred patients with symptoms or for screening to GI practices within the OSU network. OSU was developing an electronic medical record system at the time the grant application was submitted; we anticipated that this system would be fully operational by the time the study began and would greatly facilitate the medical record reviews planned for the study.

During the second year of the study, potentially eligible patients were randomly selected from lists of all patients age 49 and older at each clinic. A medical record review was conducted for those patients to ensure that they met the eligibility criteria. To be eligible for the cohort, patients had to: 1) be age 49 or older; 2) have no prior history of familial/hereditary cancer syndrome (e.g. hereditary non-polyposis colon cancer), polyps, or inflammatory bowel disease (e.g. Crohn's disease); 3) have a current phone number; 4) have had 2 or more visits to the clinic in the past 2 years (current patients); 5) have a current address in their medical record and no definite plans to move within the next year; 6) be at average risk for CRC; and 7) be in good health (e.g., no contraindications to CRC screening, such as dementia) as judged by their primary care doctor. Patients who appeared eligible based on the medical record review were contacted by phone to confirm eligibility and were recruited to the study. A time was arranged where the study was described in general terms; questions were answered; informed consent, HIPAA consent, and a medical release were obtained; and the baseline survey was administered by the interviewer. Based on power calculations detailed later in the paper, the goal was to recruit 94 patients per clinic. Each patient was told that study staff would contact him/her at various points during the next 3 years and that another survey would be administered in 2–3 years. Informed consent procedures and the study protocol were approved by the Institutional Review Board of The Ohio State University.

Measures

The baseline participant survey collected data on use of medical care, health beliefs, knowledge of CRC, risk factors for CRC, the most recent CRC screening test (where it was obtained and the result), reported exposure to information about CRC, and demographic characteristics (e.g. marital status, education, race, age, and income). Additional questions provided process evaluation measures for the clinic and patient-directed intervention components and assessed the degree of contamination among participants in the usual care clinics. In addition, medical record reviews were conducted at baseline and were planned at 12-months and at the end of the study to assess the effect of each step of the intervention on CRC screening. Based on power calculations, the goal was to recruit 94 participants per clinic at baseline and to assess 75 participants per clinic at the end of the study.

The primary outcome in this study, CRC screening within guidelines, was defined as having any one of the following based on medical record review (MRR) since the time the participant was recruited into the cohort: 1) FOBT; 2) flexible sigmoidoscopy; 3) double-contrast barium enema; or 4) colonoscopy. Potential participants who were within CRC guidelines could not enter the cohort. If they had any of the four procedures after entry to the study, they would be considered as within CRC screening guidelines at 24-months follow-up.

Intervention

The multi-level intervention to improve CRC screening rates was directed at primary care practices and targeted the clinic, providers and average-risk patients age 50 and older. Formative data collection using focus groups and intervention development took place during the first year of the study. The intervention components were based on previous investigator experience and evidence from the literature [15–17]. The provider CRC screening intervention was modified from an evidence-based tobacco cessation program (the 5 A's: Ask, Advise, Assess, Assist, and Arrange; and the 5 R's: Relevance, Risks, Rewards, Roadblocks, and Repetition) and focused on promoting CRC screening rather than tobacco cessation [18].

The first phase of the intervention focused on the clinic and the healthcare providers. It began after the baseline participant survey and medical record review and lasted one year. This phase of the intervention was based on the Chronic Care Model [19] which postulates that the community and the healthcare system each contribute to a productive provider-patient interaction. Intervention components in this phase included: 1) self-management support by having waiting room and examination room posters and brochures for the patients; and 2) decision support by conducting a CME session for physicians that addresses evidence-based screening guidelines and focused on two communication strategies: the 5 A's approach to assist providers with promoting CRC screening among patients and the 5 R's approach to assist patients to complete screening.

After the first year of intervention, a second medical record review identified participants who were still not within CRC screening guidelines. The second phase of the intervention targeted those participants and was conducted in three increasingly intensive steps. This phase of the intervention was based on well-established health behavior theories, including the Health Belief Model (perceived severity, susceptibility, benefits, and barriers) [20], the Stages of Change from the Transtheoretical Model (barriers counseling) [21], and social support (patient navigators) [22].

The first step was a letter sent to the participant from their primary care doctor with CRC screening brochures. One brochure was a commercially available (American Cancer Society) CRC screening brochure and the second was a study-specific brochure targeting the

patient's gender and race. Targeted messages included quotes from formative work with non-study, clinic patients, conducted prior to beginning the study. If the participant did not arrange or complete a screening test in three months, the second step of the intervention took place.

The second step of the patient-level intervention was a barriers-counseling telephone call that addressed the individual participant's specific CRC screening barriers. The call was made by a patient navigator trained to address CRC screening barriers and one who had experience assisting patients with arranging screening tests within the medical system. If a participant did not arrange or complete a CRC screening test within three months after the barriers call, they were contacted for an in-person visit with the patient navigator, the third step of the intervention. The in-person visit used a structured presentation that included: 1) information about the importance of completing a CRC screening test, 2) screening test options, 3) how to discuss screening with providers, 4) CRC screening barriers, and 5) tips to prevent colon cancer. Additionally, the goal of this in-person session was to assess the participant's knowledge about CRC screening, answer the participant's questions, address identified screening barriers, and provide encouragement and support for CRC screening.

Process evaluation was conducted during each step of the intervention. Evaluation of the clinic and provider components of the intervention included regularly confirming that CRC screening posters and brochures were visible in the waiting and examination rooms of the clinics. In addition, educational sessions were conducted with pre-post tests measuring change in knowledge and attitudes for all providers at the intervention clinics at the start of the study and for any new providers joining the intervention clinics during the study.

Process evaluation of the patient-level components of the intervention included: 1) tracking outgoing physician letters containing brochures to identify any changes in the participant's residence; 2) monitoring the types (patient, system) of CRC screening barriers reported and monitoring patient navigator calls to make certain that specific barriers were being appropriately addressed; and 3) observing patient navigator in-person visits to determine the CRC screening barriers reported by recalcitrant participants and to ensure that the navigator was providing encouragement and support for CRC screening.

Analysis Plan

Observations taken from members (here patients) of the same identifiable social group (here clinic) are expected to be correlated due to commonalities in selection, exposure, mutual interaction, or some combination of those factors [23]. This intraclass correlation (ICC) reflects a component of variance attributable to the group above and beyond variation attributable to its members. The presence of ICC causes outcome data from group-randomized trials to violate the assumption of independent errors associated with the most familiar analytic methods, including those based on the General Linear Model [24]. Cornfield identified this problem in 1978 [25] and simulation studies have confirmed that an analysis that ignores the extra variation carries an inflated Type I error rate [26–29]. From the outset, we planned to avoid this problem by using analytic methods appropriate to the structure of the design [14, 30]. In particular, we planned analyses based on the General Linear Mixed Model [31–36] and on the Generalized Linear Mixed Model [37, 38]. The General Linear Mixed Model is appropriate when there are multiple sources of random variation and all sources are distributed Gaussian. The Generalized Linear Mixed Model is appropriate when there are multiple sources of random variation and the observation-level data have a non-Gaussian distribution. We planned to implement the General Linear Mixed Model using SAS PROC MIXED and the Generalized Linear Mixed Model using SAS PROC GLIMMIX, Version 9.2 [39]. MIXED allows the user to model a variety of random-effects covariance structures with all sources of random variation assumed Gaussian.

GLIMMIX has most of the same features as MIXED but allows the user to specify link and variance functions appropriate to the distribution of the outcome variable.

The primary analysis to evaluate intervention effects was proposed based on a mixed-effects logistic regression model, which is a special case of the Generalized Linear Mixed Model. In this analysis, the follow-up data are regressed on condition, modeled as a fixed effect, with clinic included as a nested random effect, and with regression adjustment for covariates identified in advance. The intervention effect is defined as the odds ratio comparing the intervention and control conditions for screening up to date. The expected df are calculated as $2(g-1)$ where there are g groups per condition. In this study, we planned for 6 clinics per condition so the expected $df=2(6-1)=10$. This formula is appropriate given balanced data; we expect to use the Kenward-Roger approach for df [40], as the data are not likely to be balanced. This analysis assumes that there are two sources of random variation (clinics and residual error). The observations are assumed independent conditional upon these random effects.

Clinics were randomized from within pairs matched on the proportion of African American patients to ensure balance on factors closely related to race-ethnicity. Even so, we planned to ignore the matching in the outcome analysis in order to preserve power. Expected degrees of freedom are halved in a matched analysis, compared to an unmatched analysis, and when they are very limited, as in this study, the cost in terms of power from giving up half the degrees of freedom is substantial. Fortunately, ignoring the matching has no impact on the Type I error rate in a group-randomized trial [41, 42], so that the decision to reflect matching in the analysis can be made based on the expected effect on power [14].

We have made a concerted effort to track, locate and measure all participants at each survey. However, even in the best of studies, there is almost always some missing data. Based on our previous experience, we estimate attrition of no more than 20% of the cohort over 2–3 years. Multiple imputation is now widely regarded as an effective method for replacing missing data and we planned our analysis using this approach [43]. For the primary outcome and other 0/1 outcomes, we plan to fit a logistic regression equation using baseline data from patients to predict status on the outcome at follow-up; we plan to include fixed effects for condition and baseline covariates and a nested random effect for clinic [44]. We plan to use that equation to generate an imputed value to replace each missing outcome value, adding a random error component based on the distribution of those with complete data. By repeating that procedure several times, we can obtain several complete data sets that differ only in the values imputed to replace the missing outcome data. We can then analyze those data sets and combine their results to provide a summary estimate and test for the intervention effect. We plan to employ R to implement the multiple imputation procedures [45].

Secondary analyses are planned explore the assumption of homogeneity of the intervention effect across subgroups of participants. The primary subgroup variables are race (African American vs. Other) and age group (> age 65 vs. other).

Power

We hypothesized that at the end of the study, 35% of the cohort participants would be screened up-to-date in the intervention condition, while only 20% would be screened up-to-date in the control condition. We estimated the total variance using the expected average rate of CRC screening-within-guidelines at the end of the study, or

$\sigma_y^2 = \overline{pq} = \left(\frac{0.20+0.35}{2} \right) \left(1 - \frac{0.20+0.35}{2} \right) = 0.20$. We estimated the clinic-level ICC using data on CRC screening-within-guidelines from the clinics we proposed for the study; that estimate was 0.00455. To be conservative, we projected no benefit from regression adjustment for

covariates, so that the two thetas were set at 1.0. We anticipated an intervention effect of $35\% - 20\% = 15\%$; but because we also anticipate 20% attrition over 2–3 years, with multiple imputation of missing data, we performed the sample size calculation as though those missing had no intervention effect and reduced the effect accordingly, to $15\% * (100\% - 20\%) = 12\%$. With these considerations, 89% power was available for the primary outcome with 6 clinics per condition and 94 cohort participants per clinic at baseline allowing for 20% attrition [14].

Timeline

Participant recruitment began in January 2009 and finished in August 2010. The clinic intervention began in January 2011 and finished in December 2011. The patient directed intervention began in January 2012 and finished in December 2012. The end of study participant survey and medical record review is planned for early 2013.

RESULTS

Recruitment of Clinics and Patients

Recruitment of clinics and patients proved much more difficult than expected. As noted above, we proposed 12 clinics from the OSU Primary Care Network at the time the grant application was submitted in March, 2006; all were expected to have fully operational electronic medical records by the time the project began. That set of 12 included 4 clinics from the Columbus Neighborhood Health Centers, Federally-Qualified Health Centers (FQHC's). The director of the Centers provided a letter of support for the grant application. When the study began in May 2007, a new director was in place and staff were being cut due to budget reductions. Moreover, there were no electronic medical records at these FQHC's nor were the medical charts reliable. Thus, there was no easy way to identify potentially eligible patients. As a result, these health centers were not included in the study.

Within the OSU Primary Care Network, we had originally identified 14 primary care clinics to approach about participation, with a goal of recruiting eight. When the grant was funded, 4 of these clinics had been dissolved due to low volume and cost concerns for the practice and their patients were included in the remaining 10 clinics of the OSU Primary Care Network. We were able to recruit all 10 of the remaining clinics for the study, but that left us two short of the 12 clinics that had been proposed in the grant application. These ten clinics had participated in previous research projects and all had access to CRC screening modalities and referred patients to gastrointestinal (GI) practices within the OSU network for screening. The seven family practice clinics provided services to 14,842 adult patients who were 50 or older, of which 33% were aged 65+, 59% were female, 26% were racial or ethnic minorities, and 56% were married. The three general internal medicine clinics served 2067 adult patients who were 50 or older, of which 34% were aged 65+, 63% were female, 11% were of racial or ethnic minorities, 58% were married, and 59% had private insurance. Clinics were matched on the proportion of African American patients and randomized from within pairs to the intervention and control conditions with five clinics randomized to each condition.

The recruitment of patients from these 10 clinics was difficult due to higher than expected refusal and ineligibility rates. Figure 1 summarizes the screening and recruitment of patients. A total of 8405 patients age 50 and older were screened across the 10 clinics. A total of 3233 appeared eligible after medical record review, but 1758 were ruled ineligible during the telephone contact for recruitment. Another 916 declined to participate in the study, leaving 559 who were eligible and agreed to participate. Eight were judged ineligible after consent, leaving only 551 patients who were fully eligible, consented, and began the study.

Individuals who refused were on average almost 5 years older (60.8 yrs vs. 56.1 yrs), more likely to be males (46.5% vs. 41.9%), more likely to be white (69.7% vs. 63.7%), and less likely to be African American (18.9% vs 29.2%). Individuals who refused were more likely to be married (54.0% vs. 47.0%) and less likely to be divorced (12.2% vs 16.5%), widowed (8.6% vs. 4.9%), or single (23.8% vs. 29.4%).

Characteristics of Participating Patients

Table 1 summarizes the distribution observed for demographic and basic health variables for the control and intervention conditions, separately and combined. There was little difference in the distribution of gender, ethnicity, race, marital status, employment status, insurance status, or smoking status. Compared to the intervention condition, there were more college graduates, more higher income patients, and more patients with a medical condition in the control condition. Compared to the control condition, there were more individuals with some college and better perceived relative health in the intervention condition. There was no difference in the average age at the time of the medical record review -- 56.1 years in both conditions.

DISCUSSION

The goal of this study is to test the effect of a stepped intervention on increasing CRC screening rates among average-risk patients age 50 and older from primary care practices. A group-randomized trial design was used to avoid contamination of patients within clinics. The original grant application planned for six clinics per condition and 94 patients per clinic to be enrolled at baseline, providing 89% power for an intervention effect of 15% in the percentage screened within guidelines, allowing for a 20% attrition rate among participants.

Recruitment of clinics and patients proved much more difficult than expected and we were able to recruit only five clinics per condition and 55 cohort patients per clinic at baseline. Reasons for lower than expected accrual of clinics centered around health system factors. A change in medical directors, reductions in staff due to budget cuts, and delays in implementation of the electronic medical record system in the FQHCs prevented us from working with 4 health centers. Four other OSU clinics closed due to low patient volume and cost concerns. This left us with only 10 clinics, which we were able to recruit into the study.

The recruitment of patients from these 10 clinics was more difficult than expected due to higher than expected refusal and ineligibility rates. Reasons for refusals included not being interested, no time, and not wanting to participate in a research study. The main reasons for ineligibility were already had a CRC screening test within guidelines, too old, or higher-than-average risk for CRC. We underestimated the percentage of patients at higher-than-average risk *a priori* because history of a prior adenomatous polyp was not always listed in the primary medical chart and was identified only during the screening interview. In addition, CRC screening rates were higher than determined in our pilot study, possibly due to patients obtaining their CRC screening tests outside of the OSU system and these results not being included in the OSU medical chart; those tests were also uncovered only during the screening interview. Anecdotally, a few providers mentioned that they often referred patients to other health care systems providing colonoscopies because there was a 6-month wait for colonoscopy appointments within the OSU health care system.

When recruitment falls short of expectations, power is necessarily reduced. The original grant application estimated 89% power based on 6 clinics per condition and 94 patients per clinic at baseline, with a 20% attrition rate for patients, providing 75 patients per clinic at follow-up. As implemented, the study had 5 clinics per condition and 55 patients per clinic at baseline, providing only 70% power for the original intervention effect of 15% given

attrition of 20%. Even so, 80% power was available for an intervention effect that was only slightly larger at 16.8%.

Randomization was from within pairs of clinics matched on the proportion of African American patients. As a result, it is not surprising that we observed little difference between the two conditions in terms of the distribution of race-ethnicity groups. There were, however, apparent differences on other variables, including education level, total household income, medical conditions, and perceived health. This illustrates a common challenge in group-randomized trials, where the limited number of units of assignment makes it difficult for randomization to achieve balance on all possible confounders. To avoid any potential for confounding, these variables should be included as covariates in the primary analysis.

Implications for Future Studies

This study demonstrates the difficulties of conducting research in the real world where changes in economic conditions, staff turnover/layoff, inadequate medical records, inaccurate estimates of key parameters used in the design of the study, and poor acceptance of research can significantly impact study plans. Some of these challenges are difficult to anticipate, but each bears some consideration here.

It was not possible in March 2006 for us to anticipate the economic difficulties the country would face as the economy began to deteriorate later that year and through 2007, resulting in the deepest and longest recession since the Great Depression. Even so, it underscores the need for investigators to plan studies so as to allow for attrition. In a group-randomized trial, investigators should plan for attrition at the level of the unit of assignment, or group, as well as at the level of the unit of observation, or member. In our case, we allowed for attrition at the clinic level indirectly, as we powered at 89%. That allowed us to maintain 80% power even after losing two clinics simply by increasing the target intervention effect slightly. Had we powered originally at the nominal level of 80%, we would have been forced to generate an even larger intervention effect for 80% power.

Even more important for our study was our reliance on parameter estimates from pilot studies that did not accurately reflect the values for those parameters in our target population. Statisticians will be the first to point out that the accuracy of a power analysis depends substantially on the accuracy of the parameter estimates used in that power analysis. In our case, we underestimated the proportion of the sample that would be ineligible because they were already screened within guidelines or because they were at higher than average risk. We also underestimated the proportion that would refuse to participate in the study. These errors were often the result of not using the same procedures in the pilot work that we planned to use in the funded study. For example, we relied on estimates based on medical record reviews, rather than refining those estimates by calling the patients to get additional information about their screening and health history. Had we used the same procedures in the pilot studies that we used in the funded study, we likely would have had more accurate estimates of eligibility and refusal rates, and been able to plan accordingly.

Another error on our part was the assumption that electronic medical records would be fully operational in all of the clinics proposed for the study by the time the project was launched. Implementation of that complex electronic medical record system took longer than expected, and the paper records in some clinics proved an inadequate substitute. We might have avoided this problem by confirming, in advance, that there was an adequate backup system if it turned out that the electronic medical records were not available at the time the project launched.

Research conducted in real-world environments is often more difficult than laboratory experiments conducted under carefully controlled conditions. Even so, such research is often the only way to answer important research questions. As illustrated by our experience, it is prudent to test as many assumptions as possible during a pilot or preliminary stage, whether those assumptions involve eligibility rates, refusal rates, adequacy of medical records, screening rates, or other parameters important to the design of the study. Failure to do so can make it even more difficult to adequately test the study hypothesis, and to move the science forward.

Acknowledgments

Financial support: Supported by Research Grants R01 CA116487 and P30 CA016058 (to E.D. Paskett), and K07 CA107079 (to M.L. Katz) from the National Cancer Institute. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Cancer Institute or the National Institutes of Health.

Clinical Trials registration no. NCT01568151.

REFERENCES

1. Cancer Facts and Figures, 2010. American Cancer Society; Atlanta: 2010.
2. Colorectal Cancer Facts and Figures, 2008–2010. American Cancer Society; Atlanta: 2008.
3. Screening for colorectal cancer: U.S. Preventive Services Task Force recommendation statement. *Ann Intern Med.* 2008; 149(9):627–637. [PubMed: 18838716]
4. Smith RA, Cokkinides V, Brawley OW. Cancer screening in the United States, 2009: a review of current American Cancer Society guidelines and issues in cancer screening. *CA Cancer J Clin.* 2009; 59(1):27–41. [PubMed: 19147867]
5. Winawer S, Fletcher R, Rex D, et al. Colorectal cancer screening and surveillance: clinical guidelines and rationale–Update based on new evidence. *Gastroenterology.* 2003; 124(2):544–560. [PubMed: 12557158]
6. Mandel JS, Church TR, Bond JH, et al. The effect of fecal occult-blood screening on the incidence of colorectal cancer. *N Engl J Med.* 2000; 343(22):1603–1607. [PubMed: 11096167]
7. Selby JV, Friedman GD, Quesenberry CP Jr. Weiss NS. A case-control study of screening sigmoidoscopy and mortality from colorectal cancer. *N Engl J Med.* 1992; 326(10):653–657. [PubMed: 1736103]
8. Vernon SW, Meissner H, Klabunde C, et al. Measures for ascertaining use of colorectal cancer screening in behavioral, health services, and epidemiologic research. *Cancer Epidemiol Biomarkers Prev.* 2004; 13(6):898–905. [PubMed: 15184243]
9. Cancer Prevention and Early Detection Facts and Figures, 2010. American Cancer Society; Atlanta: 2010.
10. Pederson LL, Sidney K, Lefcoe NM. An objective measure of the validity of children's responses to a questionnaire on health and smoking. *Canadian Journal of Public Health.* 1977; 68(6):497–498.
11. Fletcher SW. The periodic health examination and internal medicine: 1984. *Ann Intern Med.* 1984; 101(6):866–868. [PubMed: 6497198]
12. Vernon SW. Participation in colorectal cancer screening: a review. *J Natl Cancer Inst.* 1997; 89(19):1406–1422. [PubMed: 9326910]
13. Holden, D.; Harris, R.; Porterfield, D., et al. Enhancing the use and quality of colorectal cancer screening. Agency for Healthcare Research and Quality; 2010.
14. Murray, DM. Design and Analysis of Group-Randomized Trials. Oxford University Press; New York, NY: 1998.
15. Katz ML, Tatum C, Dickinson SL, et al. Improving colorectal cancer screening by using community volunteers: results of the Carolinas cancer education and screening (CARES) project. *Cancer.* 2007; 110(7):1602–1610. [PubMed: 17665496]

16. Paskett E, Tatum C, Rushing J, et al. Randomized trial of an intervention to improve mammography utilization among a triracial rural population of women. *Journal of the National Cancer Institute*. 2006; 98(17):1226–1237. [PubMed: 16954475]
17. Post DM, Katz ML, Tatum C, Dickinson SL, Lemeshow S, Paskett ED. Determinants of colorectal cancer screening in primary care. *Journal of cancer education : the official journal of the American Association for Cancer Education*. 2008; 23(4):241–247. [PubMed: 19058074]
18. Fiore, MC.; Jaen, CR.; Baker, TB. Treating Tobacco Use and Dependence: 2008 Update. Clinical Practice Guideline. U.S. Department of Health and Human Services, Public Health Service; Rockville, MD: 2008.
19. Wagner EH. Chronic disease management: what will it take to improve care for chronic illness? *Eff Clin Pract*. 1998; 1(1):2–4. [PubMed: 10345255]
20. Rosenstock IM, Strecher VJ, Becker MH. Social learning theory and the Health Belief Model. *Health Educ Q*. 1988; 15(2):175–183. [PubMed: 3378902]
21. Prochaska, JO.; Redding, CA.; Evers, KE. The Transtheoretical Model and Stages of Change. In: Glanz, K.; Rimer, BK.; Viswanath, K., editors. *Health Behavior and Health Education. Theory, Research, and Practice*. 4 ed. Jossey-Bass, A Wiley Imprint; San Francisco: 2008. p. 97-122.
22. Heaney, CA.; Israel, BA. Social networks and social support. In: Glanz, K.; Rimer, BK.; Lewis, FM., editors. *Health Behavior and Health Education. Theory, Research, and Practice*. #rd ed. Jossey-Bass; San Francisco: 2002. p. 185-209.
23. Kish, L. Survey Sampling. John Wiley & Sons; New York, NY: 1965.
24. Searle, SR. Linear Models. John Wiley & Sons; New York, NY: 1971.
25. Cornfield J. Randomization by group: a formal analysis. *Am J Epidemiol*. 1978; 108(2):100–102. [PubMed: 707470]
26. Murray DM, Hannan PJ, Baker WL. A Monte Carlo study of alternative responses to intraclass correlation in community trials: Is it ever possible to avoid Cornfield's penalties? *Eval Rev*. 1996; 20(3):313–337. [PubMed: 10182207]
27. Murray DM, Hannan PJ, Wolfinger RD, Baker WL, Dwyer JH. Analysis of data from group-randomized trials with repeat observations on the same groups. *Stat Med*. 1998; 17(14):1581–1600. [PubMed: 9699231]
28. Murray DM, Wolfinger RD. Analysis issues in the evaluation of community trials: progress toward solutions in SAS/STAT MIXED. *J Community Psychol*. 1994; (Special Issue):140–154. CSAP.
29. Zucker DM. An analysis of variance pitfall: The fixed effects analysis in a nested design. *Educational and Psychological Measurement*. 1990; 50:731–738.
30. Donner, A.; Klar, N. Design and Analysis of Cluster Randomization Trials in Health Research. Arnold; London: 2000.
31. Donner A. Linear regression analysis with repeated measurements. *J Chronic Dis*. 1984; 37(6):441–448. [PubMed: 6725498]
32. Donner A. A regression approach to the analysis of data arising from cluster randomization. *Int J Epidemiol*. 1985; 14(2):322–326. [PubMed: 4019000]
33. Laird NM, Lange N, Stram D. Maximum likelihood computations with repeated measures: application of the EM algorithm. *Journal of the American Statistical Association*. 1987; 82(397):97–105.
34. Laird NM, Ware JH. Random-effects models for longitudinal data. *Biometrics*. 1982; 38(4):963–974. [PubMed: 7168798]
35. Stiratelli R, Laird N, Ware JH. Random-effects models for serial observations with binary response. *Biometrics*. 1984; 40(4):961–971. [PubMed: 6534418]
36. Ware JH. Linear models for the analysis of longitudinal studies. *The American Statistician*. 1985; 39(2):95–101.
37. Breslow NE, Clayton DG. Approximate inference in generalized linear mixed models. *Journal of the American Statistical Association*. 1993; 88(421):9–25.
38. Wolfinger RD, O'Connell M. Generalized linear mixed models: a pseudo-likelihood approach. *Journal of Statistics and Computational Simulation*. 1993; 48:233–243.

39. Littell, RC.; Milliken, GA.; Stroup, WW.; Wolfinger, RD.; Schabenberger, O. SAS for Mixed Models - Second Edition. Second Edition ed. SAS Institute Inc.; Cary, NC: 2006.
40. Kenward MG, Roger JH. Small sample inference for fixed effects from restricted maximum likelihood. *Biometrics*. 1997; 53(3):983–997. [PubMed: 9333350]
41. Proschan MA. On the distribution of the unpaired t-statistic with paired data. *Stat Med*. 1996; 15(10):1059–1063. [PubMed: 8783442]
42. Diehr P, Martin DC, Koepsell T, Cheadle A. Breaking the matches in a paired t-test for community interventions when the number of pairs is small. *Stat Med*. 1995; 14(13):1491–1504. [PubMed: 7481187]
43. Shadish, WR.; Cook, TD.; Campbell, DT. Experimental and Quasi-Experimental Designs for Generalized Causal Inference. Houghton Mifflin Company; Boston, MA: 2002.
44. Andridge RR. Quantifying the impact of fixed effects modeling of clusters in multiple imputation for cluster randomized trials. *Biometrical Journal*. 2011; 53(1):57–74. [PubMed: 21259309]
45. R: A language and environment for statistical computing. R Foundation for Statistical Computing; Vienna: 2011.

Figure 1.
Patient Flow Chart

Table 1

Distribution of categorical variables by study condition and total.

Variable	Level	Control (n=259)	Intervention (n=292)	Total
Gender	Male	109 (42.1%)	122 (41.8%)	231 (41.9%)
	Female	150 (57.9%)	170 (58.2%)	320 (58.1%)
Hispanic	No	254 (98.1%)	284 (97.3%)	538 (97.6%)
	Yes	3 (1.2%)	6 (2.1%)	9 (1.6%)
Race	White	168 (64.9%)	188 (64.4%)	356 (64.6%)
	Black	73 (28.2%)	87 (29.8%)	160 (29.0%)
	Asian	7 (2.7%)	6 (2.1%)	13 (2.4%)
	Native American	2 (0.8%)	1 (0.3%)	3 (0.5%)
	Other	7 (2.7%)	10 (3.4%)	17 (3.1%)
Marital status	Married	120 (46.3%)	129 (44.2%)	249 (45.2%)
	Living Together	4 (1.5%)	8 (2.7%)	12 (2.2%)
	Divorced	60 (23.2%)	67 (22.9%)	127 (23.0%)
	Separated	14 (5.4%)	11 (3.8%)	25 (4.5%)
	Single	39 (15.1%)	58 (19.9%)	97 (17.6%)
	Widowed	21 (8.1%)	19 (6.5%)	40 (7.3%)
Highest level of education completed	<8 th Grade	4 (1.5%)	6 (2.1%)	10 (1.8%)
	Some High School	18 (6.9%)	16 (5.5%)	34 (6.2%)
	High School	41 (15.8%)	37 (12.7%)	78 (14.2%)
	Some College	66 (25.5%)	97 (33.2%)	163 (29.6%)
	Associate Degree	26 (10.0%)	24 (8.2%)	50 (9.1%)
	College Degree	63 (24.3%)	56 (19.2%)	119 (21.6%)
	Graduate Degree	40 (15.4%)	55 (18.8%)	95 (17.2%)
Employment status	Full Time	120 (46.3%)	146 (50.0%)	266 (48.3%)
	Part Time	20 (7.7%)	34 (11.6%)	54 (9.8%)
	Retired	45 (17.4%)	45 (15.4%)	90 (16.3%)
	Disabled	44 (17.0%)	39 (13.4%)	83 (15.1%)
	Unemployed	28 (10.8%)	27 (9.2%)	55 (10.0%)
	Volunteer	1 (0.4%)	1 (0.3%)	2 (0.4%)
Total household income last year (\$)	<10K	36 (13.9%)	36 (12.3%)	72 (13.1%)
	10–19K	29 (11.2%)	37 (12.7%)	66 (12.0%)
	20–29K	24 (9.3%)	31 (10.6%)	55 (10.0%)
	30–39K	23 (8.9%)	28 (9.6%)	51 (9.3%)
	40–49K	19 (7.3%)	28 (9.6%)	47 (8.5%)
	50–59K	19 (7.3%)	27 (9.2%)	46 (8.3%)
	60–69 K	14 (5.4%)	15 (5.1%)	29 (5.3%)
	70K+	80 (30.9%)	75 (25.7%)	155 (28.1%)
Currently have any form of health insurance	No	36 (13.9%)	41 (14.0%)	77 (14.0%)

Variable	Level	Control (n=259)	Intervention (n=292)	Total
	Yes	220 (84.9%)	251 (86.0%)	471 (85.5%)
Health compared to others your age	Excellent	31 (12.0%)	58 (19.9%)	89 (16.2%)
	Very Good	82 (31.7%)	107 (36.6%)	189 (34.3%)
	Good	83 (32.0%)	85 (29.1%)	168 (30.5%)
	Fair	49 (18.9%)	34 (11.6%)	83 (15.1%)
	Poor	9 (3.5%)	8 (2.7%)	17 (3.1%)
Any medical conditions	No	74 (28.6%)	132 (45.2%)	206 (37.4%)
	Yes	182 (70.3%)	157 (53.8%)	339 (61.5%)
Smoked at least 100 cigarettes in your entire life?	No	109 (42.1%)	134 (45.9%)	243 (44.1%)
	Yes	149 (57.5%)	155 (53.1%)	304 (55.2%)