

Design of the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial

Philip C. Prorok, PhD, Gerald L. Andriole, MD,
Robert S. Bresalier, MD, Saundra S. Buys, MD,
David Chia, PhD, E. David Crawford, MD,
Ronald Fogel, MD, Edward P. Gelmann, MD,
Fred Gilbert, MD*, Marsha A. Hasson, MS,
Richard B. Hayes, PhD, Christine Cole Johnson, PhD, MPH,
Jack S. Mandel, PhD, MPH, Albert Oberman, MD, MPH,
Barbara O'Brien, MPH, Martin M. Oken, MD,
Sameer Rafla, MD, PhD, Douglas Reding, MD, MPH,
Wilmer Rutt, MD**, Joel L. Weissfeld, MD, MPH,
Lance Yokochi, MD, MPH, and John K. Gohagan, PhD,
FACE for the PLCO Project Team***

Division of Cancer Prevention (P.C.P., J.K.G.) and Division of Cancer Epidemiology and Genetics (R.B.H.), National Cancer Institute, Bethesda, Maryland; Washington University School of Medicine, St. Louis, Missouri (G.L.A.); Henry Ford Health System, Detroit, Michigan (R.S.B., R.F., C.C.J., W.R.); University of Utah Health Sciences Center, Salt Lake City, Utah (S.S.B.); UCLA Tissue Typing Laboratory, Los Angeles, California (D.C.); University of Colorado Health Sciences Center, Denver, Colorado (E.D.C.); Georgetown University, Washington, DC (E.P.G.); Pacific Health Research Institute, Honolulu, Hawaii (F.G., L.Y.); Westat, Inc., Rockville, Maryland (M.A.H., B.O.); University of Minnesota, Minneapolis, Minnesota (J.S.M.); University of Alabama at Birmingham, Birmingham, Alabama (A.O.); Virginia Piper Cancer Center, Minneapolis, Minnesota (M.M.O.); Cancer Institute of Brooklyn, Brooklyn, New York (S.R.); Marshfield Medical Research and Education Foundation, Marshfield, Wisconsin (D.R.); and University of Pittsburgh Cancer Institute, Pittsburgh, Pennsylvania (J.L.W.)

*Deceased.

**Retired.

***A roster of the project team appears in the lead paper of this supplement. Other participants in early design considerations for this trial included former National Cancer Institute investigators David P. Byar (deceased) and Charles R. Smart (retired).

Address reprint requests to: Dorothy Sullivan, Early Detection Research Group, Division of Cancer Prevention, National Cancer Institute, EPN 330, 6130 Executive Blvd., Bethesda, MD 20892-7346 (E-mail: ds255j@nih.gov).

Received March 27, 2000; accepted May 31, 2000.

ABSTRACT: The objectives of the Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial are to determine in screenees ages 55–74 at entry whether screening with flexible sigmoidoscopy (60-cm sigmoidoscope) can reduce mortality from colorectal cancer, whether screening with chest X-ray can reduce mortality from lung cancer, whether screening men with digital rectal examination (DRE) plus serum prostate-specific antigen (PSA) can reduce mortality from prostate cancer, and whether screening women with CA125 and transvaginal ultrasound (TVU) can reduce mortality from ovarian cancer. Secondary objectives are to assess screening variables other than mortality for each of the interventions including sensitivity, specificity, and positive predictive value; to assess incidence, stage, and survival of cancer cases; and to investigate biologic and/or prognostic characterizations of tumor tissue and biochemical products as intermediate endpoints. The design is a multicenter, two-armed, randomized trial with 37,000 females and 37,000 males in each of the two arms. In the intervention arm, the PSA and CA125 tests are performed at entry, then annually for 5 years. The DRE, TVU, and chest X-ray exams are performed at entry and then annually for 3 years. Sigmoidoscopy is performed at entry and then at the 5-year point. Participants in the control arm follow their usual medical care practices. Participants will be followed for at least 13 years from randomization to ascertain all cancers of the prostate, lung, colorectum, and ovary, as well as deaths from all causes. A pilot phase was undertaken to assess the randomization, screening, and data collection procedures of the trial and to estimate design parameters such as compliance and contamination levels. This paper describes eligibility, consent, and other design features of the trial, randomization and screening procedures, and an outline of the follow-up procedures. Sample-size calculations are reported, and a data analysis plan is presented. *Control Clin Trials* 2000;21:273S–309S © Elsevier Science Inc. 2000

KEY WORDS: *Randomized trial, screening, design, prostate cancer, colorectal cancer, lung cancer, ovarian cancer*

INTRODUCTION

Lung and colorectal cancers, the most common cancers in Americans, accounted for 46% of cancer deaths in males and 34% of cancer deaths in females in 1989 when this trial was being considered [1]. In males, prostate cancer was the third leading cause of cancer mortality and accounted for 11% of cancer deaths. In females, ovarian cancer accounted for 5% of cancer deaths. Mortality statistics for these cancers are similar today. In 2000, there will be an estimated 28,500 deaths among women and 27,800 deaths among men from colorectal cancer and, respectively, 67,600 and 89,300 deaths from lung cancer. About 14,000 women will die from ovarian cancer and 31,900 men from prostate cancer [2].

The death rate for prostate cancer has increased somewhat over time, while the rate for colorectal cancer has dropped, especially for females. The death rate for lung cancer has risen rapidly in both sexes, with a recent downturn for males [2]. Successful screening programs for these three cancers could have a major impact on overall cancer mortality. The death rate for ovarian cancer has remained relatively stable. Nearly 70% of ovarian cancers present as advanced disease with a poor prognosis, while localized disease has a 90% survival rate [3]. Successful screening might substantially reduce ovarian cancer mortality.

Uncertainty regarding the value of screening for these cancers has resulted in conflicting positions in the medical community and confusion in populations at risk. A randomized, controlled trial is necessary to determine the effects of screening on disease-specific mortality. The Prostate, Lung, Colorectal and

Ovarian (PLCO) Cancer Screening Trial is a 23-year randomized trial in which 37,000 men will be screened for prostate, lung, and colorectal cancers and 37,000 women will be screened for lung, colorectal, and ovarian cancers. Prostate-specific antigen (PSA) and digital rectal examination (DRE) (for prostate), chest X-ray (for lung), 60-cm flexible sigmoidoscopy (for colorectal), and CA125 blood test and transvaginal ultrasound (TVU) (for ovary) are being investigated as screening modalities. An equal number of men and women will be followed with routine medical care as controls. There will be a follow-up period of at least 13 years from randomization for both intervention and control participants to determine the effects of screening on cause-specific mortality.

This paper describes the design of this trial at the completion of protocol development (just prior to the initiation of the pilot-phase recruitment) and protocol modifications that have occurred since. Included are the specific rationale for each cancer site, overall design features, screening and follow-up procedures, sample-size considerations, and data analysis plans. Recruitment into the pilot phase began November 16, 1993, with main-phase recruitment commencing September 30, 1994.

TRIAL RATIONALE

Prostate Cancer Screening

The DRE, the most common screening test for prostate cancer screening prior to 1990, has never been completely evaluated. Observational studies have examined sensitivity and case survival data, but without appropriate controls and with no adjustment for lead-time and length biases [4, 5].

In 1984, Chodak and Schoenberg [6] reported on 811 patients from 50–80 years of age who underwent rectal examination and follow-up. Thirty-eight of 43 patients with a palpable abnormality in the prostate agreed to undergo biopsy. The positive predictive value for prostate cancer was 29%. Forty-five percent of the cases were stage B, 6% stage C, and 18% stage D. More recent results from the same investigators revealed a 25% positive predictive value with 68% of the detected tumors clinically localized [7]. Others also reported a high proportion of localized disease when prostate cancer is detected by routine rectal examination [8–11]. In contrast, Wajsman and Chu [12] among others have reported that even with annual rectal examination, only 20% of cases are localized at diagnosis. Thompson and Zeidman [13] reported that 25% of men presenting with metastatic disease had a normal prostate exam.

A summary of the data on DRE for detection of prostate cancer concluded the following: sensitivity is 55–69%, specificity is 89–97%, positive predictive value is 11–26%, and negative predictive value is 85–96% [14]. Further, the rectal examination depends on the skill and experience of the examiner and the presence of a cancer in the posterior prostate. However, DRE is inexpensive, relatively noninvasive, nonmorbidity, and can be taught to nonprofessional health workers. What remains to be determined is whether routine annual screening by rectal examination reduces prostate cancer mortality. A case-control study involving 139 men with metastatic prostate cancer and matched controls found the relative risk of metastatic prostate cancer to be 0.9 for men with one or more rectal examinations compared with men with none. The 95% confidence

interval was 0.5–1.7, suggesting that screening by routine DRE appears to have little effect in detecting and treating prostate cancer before it becomes metastatic [15].

Prostatic imaging by ultrasound, computerized tomography, and magnetic resonance imaging have also been suggested for prostate cancer screening. Each modality has relative advantages and disadvantages. Transrectal ultrasound has received the most attention [8, 16–22]. In a summary, Waterhouse and Resnick [23] reported that the sensitivity and specificity of ultrasound are too low for the procedure to be a valuable screening tool. Sensitivity ranged from 71–92% for prostate cancer and 60–85% for subclinical disease. Specificity ranged from 41–79%, and positive predictive values in the 30% range have been reported. The sensitivity and positive predictive value of ultrasound may be better than those of DRE when each is used as a single test. However, the relatively low specificity along with the invasiveness and cost of the procedure preclude routine screening for prostate cancer by transrectal ultrasound.

Serum PSA has been examined in several observational settings, both for initial diagnosis of disease and as a tool to detect recurrence after initial therapy [8, 20, 24–27]. Parameter estimates for this test include sensitivity near 70% and positive predictive values of 17–28%, although these estimates of predictive value are strongly dependent upon the disease prevalence in the populations studied [28]. The potential value of PSA lies in its simplicity, objectivity, reproducibility, lack of invasiveness, and lower cost relative to ultrasound. The test has increased the detection rate of early stage cancers, many of which may be curable by local therapy [9, 29, 30]. However, the test must be carefully evaluated because false positives in the form of benign prostatic lesions are common, requiring biopsies and added expense, and PSA testing cannot distinguish between latent or biologically irrelevant versus aggressive tumors.

The use of serial tests to assess the rate of change of PSA has been evaluated as a method to improve the specificity of the test [31]. The combination of PSA and ultrasound has been used to determine PSA density indexed to prostate size [32–34]. In one study, volume-adjusted PSA identified a population at higher risk of carcinoma [35], but another study of intermediate levels of PSA found no advantage to volume-adjusted PSA levels for screening [36]. Ratios of free to complexed PSA can amplify the differences in PSA levels for individuals with prostate cancer versus prostatic hyperplasia [37, 38]. No statistical advantage has been established for using the ratio of free to total PSA compared to total PSA alone in a screened population [39]; however, the free to total PSA ratio did improve specificity in other studies [40].

In a study by Cooner et al. [41] to resolve questions surrounding the relative merits of the three tests, all subjects had a rectal examination, PSA determination (Hybritech assay), and a 7-mHz ultrasound examination. Most of the participants with positive results on ultrasound plus a few other individuals were biopsied. The pertinent findings of this study and a similar study by Lee et al. [20] are given in Table 1. Both studies demonstrate that the rate of cancer among subjects with positive results on ultrasonography in whom the rectal and PSA exams are normal is extremely low. Hence, ultrasound was not included as one of the screening tests in this trial.

Careful evaluation of prostate cancer screening is mandatory because the natural history of the disease is variable and appropriate treatment is not clearly

Table 1 Effect of Rectal and Prostate-Specific Antigen Examinations on Cancer Rate in Patients with Abnormal Rectal Ultrasound

	Cooner Study			Lee Study		
	Biopsies	Cancer	Rate	Biopsies	Cancer	Rate
Rectal +, PSA +	235	151	0.64	89	63	0.71
Rectal +, PSA -	166	23	0.14	23	6	0.26
Rectal -, PSA +	134	41	0.31	92	31	0.34
Rectal -, PSA -	177	12	0.07	44	2	0.05

PSA = prostate-specific antigen.

defined [28, 42, 43]. The incidence of prostate cancer found at autopsy steadily increases for each decade after age 50, and most of these lesions are clinically latent. Some progress has been made in predicting the biologic behavior of these tumors, but despite improved understanding of the relationship among histologic grade, tumor volume, and biologic behavior, it is difficult to determine appropriate therapy for any given tumor [44]. A meta-analysis indicated that patients with low-grade prostate cancer can experience long-term survival with deferred therapy [45]. Decision analyses produce indeterminate results because of uncertainty regarding treatment efficacy and metastatic rates for prostate cancer [46–48]. On the other hand, a review of 60,000 cases of prostate cancer diagnosed between 1983 and 1992 showed that men with poorly or moderately differentiated cancer had improved survival if treated rather than followed [49].

Screening and treatment of a large population of males could entail substantial risks and morbidity, which include urinary incontinence, urethral strictures, sexual impotence, rectal injury, and a small probability of treatment-related mortality [44, 50]. Given these circumstances, careful evaluation of prostate cancer screening is needed. Currently, there is insufficient evidence with which to decide the efficacy or effectiveness of screening asymptomatic men [44, 47]. In addition to the PLCO trial, randomized trials are underway in other countries to address these issues [51, 52].

Lung Cancer Screening

Evaluations of chest X-ray and sputum cytology, the most common screening tests for lung cancer, were first reported nearly 30 years ago. The early studies include the Philadelphia Pulmonary Neoplasm Research Project [53], a nonrandomized, uncontrolled study begun in 1951; the Veterans Administration study [54], a nonrandomized, uncontrolled study performed from 1958 to 1961; the South London Lung Cancer Study [55], a nonrandomized, uncontrolled study done in 1955 to 1963; the North London Cancer Study [56, 57], a randomized study with industrial firms randomized between screening and no screening done in the early 1960s; and the Kaiser Foundation Health Plan multiphasic screening trial [58, 59], a controlled trial with annual chest X-ray, spirometry, and medical questionnaire as part of the multiphasic screening begun in 1964. None of these studies demonstrated a significant impact of screening on lung cancer mortality. The South London study, for example, showed an increase

in the survival of screen-detected cases compared with other cases found in the same geographical region, but without adjustment for self-selection bias, lead-time bias, overdiagnosis bias, or length bias [60, 61]. These studies typically were small, and for most, follow-up was short, so that any small to moderate size effect or any long-term effect was not likely to be demonstrated.

More recent studies include a randomized trial in Czechoslovakia [62, 63], a case-control study in the former German Democratic Republic [64], and a case-control study in Japan [65]. As with some earlier studies, the randomized groups in the Czechoslovakian study were screened with cytology and X-ray at two frequencies, semiannual versus every 3 years, so that there was no unscreened control group. There was no difference in mortality between the two groups. The German case-control study evaluated chest X-rays originally used for control of tuberculosis. The Japanese case-control study considered X-ray histories among deceased lung cancer cases and matched controls. In contrast to the German study, the odds ratio of dying from lung cancer for those screened within 12 months versus those not screened was 0.72, suggesting some benefit from the screening.

Three other randomized controlled trials have been conducted. One trial, the Mayo Lung Project, was initiated in 1971 for males 45 years or older who were heavy smokers [66–68]. Participants free of lung cancer on initial screening were randomized either to a group offered screening with sputum cytology and chest X-ray every 4 months or to a group not offered screening but advised to seek it annually. In the studies at the Johns Hopkins University Hospital [69–72] and at Memorial-Sloan Kettering Cancer Center [73, 74], intervention and control groups were offered annual chest X-ray, while the intervention group was also offered sputum cytology every 4 months. In the Mayo Clinic study, cases found in the screened arm were diagnosed in earlier stages than those in the control arm. However, there was no significant reduction in lung cancer mortality between the screened group and the control group in any of these trials.

Therefore, at this point there is no solid evidence that screening for lung cancer can reduce lung cancer mortality. Sputum cytology has not been shown to be effective as an adjunct to annual chest X-ray. There is evidence that screening with chest X-ray plus sputum cytology does improve stage at diagnosis and case survival rate relative to cases diagnosed through usual care, but despite this there was no reduction in lung cancer mortality. However, modeling using data from these trials suggests that there may have been as much as an 18% mortality reduction in these trials [75–77].

The Mayo study is the only one of the three which is pertinent to studying annual X-ray in the present trial because the use of screening X-rays differed in the two arms. However, several reservations can be noted about the Mayo study finding. First, the study was designed to detect a 50% reduction in lung cancer mortality and was too small to demonstrate a lesser but important reduction of 10–15%. Second, at the time the study was terminated there were still 40 excess cases of lung cancer in the screened group. Whether these cases represent overdiagnosis or a screening benefit that would only be seen with longer follow-up is not known. Third, about 50% of the men in the control group received an annual chest X-ray [68]. Thus, the level of contamination may have been sufficient to obscure any small to moderate benefit. Finally,

Table 2 Power to Detect Various Screening Effects in Previous Studies of Chest X-Ray Screening for Lung Cancer (Based on Actual Deaths Observed)

Study	Mortality Reduction (%)				
	10	20	30	40	50
Philadelphia	0.14	0.32	0.59	0.85	0.98
VA	0.16	0.38	0.69	0.92	0.99
South London	0.14	0.31	0.57	0.83	0.97
North London	0.16	0.39	0.70	0.93	0.995
Kaiser	0.12	0.27	0.50	0.76	0.94
Czechoslovakia	0.16	0.39	0.71	0.93	0.996
Mayo	0.21	0.54	0.88	0.99	0.999

when prevalence cases were detected at the first screen, they were followed separately and were not part of the randomized comparison. Hence, any effect of X-ray on reducing lung cancer mortality among these cases could not have been determined. It can also be argued that therapeutic advances may render early detection more effective today than at the time of the Mayo trial.

The concern about insufficient size of previous studies of chest X-ray screening is illustrated in Table 2. The uncertainty in interpretation of results from completed studies has led to differences of opinion regarding the value of the annual chest X-ray. Whether a small but important benefit exists can be demonstrated only by a properly designed randomized trial.

Colorectal Cancer Screening

DRE, sigmoidoscopy, and fecal occult blood testing have each been suggested for colorectal cancer screening. However, only the fecal occult blood test has been proven to be beneficial.

Several uncontrolled studies suggesting that the fecal occult blood test leads to early detection have been reported [78–80] as have two case-control studies of the effect of occult blood testing on colorectal cancer mortality. In one study, the screening histories of fatal colorectal cancer cases and matched controls were compared, resulting in an odds ratio of 0.69 for exposure to at least one occult blood test over a 5-year period. The wide confidence interval (0.52–0.91) suggested a benefit from the screening but also the need for further data [81]. In the second study, cases were less likely to have ever been screened than controls. The odds ratio was 0.7 with a 95% confidence interval of 0.5–1.0, consistent with a screening benefit [82].

Five prospective, controlled studies of fecal occult blood testing have also been conducted. The Strang Clinic of New York undertook a nonrandomized study involving some 12,000 screenees and 7000 controls designed to test the effect of combining the stool guaiac test with annual sigmoidoscopy. Individuals were allocated to the study arms by calendar periods. A reduction in colorectal cancer mortality of borderline significance was reported [83].

A randomized trial of the stool guaiac test began in 1974 at the University of Minnesota, where nearly 47,000 persons ages 50–80 were randomized into three groups: a control group, an annually screened group, and a biennially

screened group. The preponderance of test slides were rehydrated. Recent results provided the first definitive evidence that annual testing for occult blood in the stool can reduce the death rate from colorectal cancer. The 13-year cumulative mortality from colorectal cancer was reduced by 33% (mortality ratio 0.67 with 95% confidence interval 0.50–0.87) [84].

A controlled trial in Nottingham, United Kingdom randomized approximately 76,000 individuals to each of two arms using lists of family practitioners. Fecal occult blood testing every 2 years using nonrehydrated slides was offered to the screened arm for three to six rounds of screening. A 15% reduction in colorectal cancer mortality was reported after a median follow-up time of 7.8 years [85].

Two additional randomized trials of occult blood screening were initiated more recently. A trial in Sweden targeted individuals in the narrow age range of 60–64 years [86]. A Danish trial randomized about 31,000 individuals ages 45–75 into two arms. Participants in the screened arm were offered nonrehydrated fecal occult blood tests every 2 years for five rounds over a 10-year period [87, 88]. This trial demonstrated an 18% reduction in colorectal cancer mortality [89].

In summary, testing for occult blood in the stool as a colorectal cancer screening maneuver has been studied in several trials, and a mortality reduction has been demonstrated. The focus of the PLCO trial is therefore flexible sigmoidoscopy.

DRE and rigid sigmoidoscopy were both part of the multiphasic screening program carried out by the Kaiser-Permanente Foundation, and some considered the results of this study to be evidence of the effectiveness of these tests [90]. Approximately 5000 individuals were allocated to a study group urged to receive an annual multiphasic checkup, and a comparable number served as controls. After 11 years, the screened group experienced a colorectal cancer death rate of 1.0 per 1000 participants entered compared to a rate of 3.3 per 1000 in the control group [58, 59]. The observed decrease in colorectal cancer mortality in this study could be a real effect resulting from screening. However, this conclusion has been questioned for several reasons [91]. Some cancers were detected in an investigation of anemia resulting from the multiphasic examination as well as by the two tests. Further, in a reanalysis the investigators found that rates of sigmoidoscopy were low in both groups (control: 25%; screened: 30%), that there was only a slight excess of exposure to sigmoidoscopy in the study group compared to the control group, and that there was not an appreciable difference in removal of colorectal polyps between groups. They concluded that this study should not be used as evidence either for or against sigmoidoscopy screening [92]. DRE made a minor contribution. In addition, a case-control study found no statistically significant mortality reduction from distal rectal cancer using DRE [93].

Two additional observational cohort studies of sigmoidoscopy have been reported. One involved 21,000 participants in Minnesota who underwent an annual physical examination that included sigmoidoscopy [94, 95]. Polyps discovered during screening were removed, and the number of sigmoid cancers ultimately found was only 15% of the number expected. All of the 13 cancers found were localized, and none of the patients had died as of 1979. The second study followed 26,000 men and women in New York [96]. In 50 cancer patients

identified by screening and followed over 15 years, the 5-year survival rate was reported to be 90%. The interpretation that screening was of benefit in these two studies can be questioned on several grounds. Both studies are likely to be affected by self-selection bias of participants and by exclusion of certain individuals from the follow-up process. In the New York study, seven people with a history of symptoms and eight with previously diagnosed lesions were excluded, thereby lowering the observed incidence and mortality rates. In the Minnesota study, cases found at the initial examination were excluded from the observed incidence, and only individuals without gastrointestinal symptoms were allowed to participate. Thus, the data cannot be validly compared with the general population [91]. In addition, the reported survival data from both studies are affected by lead-time and length biases, but no adjustment for these biases was attempted.

Flexible sigmoidoscopy has been shown to be more acceptable to screenees than rigid endoscopy, and the test appears to be very sensitive and highly specific for cancer [97, 98]. The test can discover a high proportion of polyps, and evidence suggests that removal of adenomas decreases the risk of colorectal cancer [99]. The need to address the impact of flexible sigmoidoscopy screening on colorectal cancer mortality has been discussed by several investigators [97, 100, 101]. Encouraging reports of the potential impact of this test come from two case-control studies and from the modeling work of Eddy et al. [102, 103], which suggests a potential mortality reduction of 25–40%. Both case-control studies were conducted in prepaid health plans and used colorectal cancer deaths as cases, with matched controls. Exposure to sigmoidoscopy in cases and controls was compared [104, 105]. Rigid sigmoidoscopy was used in one study, while a majority of the screening was by flexible sigmoidoscopy in the other study. Both studies suggested a strong effect of sigmoidoscopy in reducing colorectal cancer mortality, with unadjusted odds ratios of 0.30 and 0.21. The modeling conclusions and the case-control studies are subject to the assumptions and biases in the methodologies, so that conclusive results will only be obtained from a randomized trial.

Ovarian Cancer Screening

Traditionally, the pelvic examination has been relied on to detect ovarian cancer, but it is insensitive to early disease and small tumors [106]. Thus, most ovarian cancers present as late-stage disease. Two new technologies may be useful as screening tools: CA125 and TVU.

CA125 is an antigenic determinant on a high molecular weight glycoprotein recognized by a monoclonal antibody (OC 125) using an ovarian cell line as an immunogen. The test is performed on peripheral blood. In mostly small (50–150 patients) preoperative studies of women with ovarian masses, serum CA125 levels were elevated (typically above 35 U/mL) in 68–100% of cases averaged over all stages and in 40–50% of stage I disease. Serum CA125 may also be elevated with pregnancy, endometriosis, menstruation, benign ovarian tumors, and with breast, colon, pancreatic, lung, gastric, and liver cancers [107]. CA125 was reported to have high specificity in postmenopausal women in two prospective trials. Among 1010 postmenopausal women undergoing both pelvic examination and CA125, the only malignancy diagnosed was detected

by CA125 [107]. The specificity was 94.3%. In a study in Sweden among 5550 women over 40 years of age, nine cancers were detected, six of the nine by CA125 [108]. Specificity was 98.5% using a threshold of 35 U/mL in women 50 years of age and older. The sensitivity of CA125 was estimated in two nested case-control studies using sera available from two serum banks [109, 110]. The sensitivity for a level of at least 35 U/mL ranged from 20–57% for cases occurring within the first 3 years of follow-up. These two studies also reported a specificity of 95%.

These preoperative and prospective studies together suggest early detection potential for CA125. However, no studies have been conducted to measure sensitivity and specificity in a large screened population, and no randomized trials have been initiated to assess the impact of screening with CA125 on ovarian cancer mortality.

TVU has been proposed for ovarian cancer screening [111], but experience with this modality is limited. In a series of 1017 tumors, 0.3% of ovarian tumors unilocular on ultrasound were malignant, while 8% of those that were multilocular and 39% of those that were solid were malignant [106]. Higgins et al. and Van Nagell et al. [111, 112] have been using TVU for screening women over the age of 40 since 1987. Using 8 cm³ as the upper limit of normal ovarian volume, 31 abnormal ultrasonograms (in 1000 women) were obtained; 24 of these women underwent laparotomy. TVU identified all three of the cancers detected.

Estimates of yield and false positivity of ultrasound are available from several studies of women offered periodic screening. In a cohort of 801 women ages 40–70 who had one or more risk factors for ovarian cancer, 163 had an abnormal abdominal ultrasound. Surgery was performed in 30 cases, and one borderline ovarian tumor was found [113]. In another study of abdominal ultrasound, 5479 asymptomatic women underwent periodic screening. Of 326 participants who had a positive test and went on to surgery, five women were diagnosed with stage IA or IB ovarian cancer, and four were diagnosed with metastatic ovarian cancer [114]. TVU was also used in a study of 3220 asymptomatic, postmenopausal women. An abnormal exam led to exploratory laparotomy in 44 women. Three primary ovarian carcinomas were found, two with stage IA cancer [115]. Finally, both transvaginal and transabdominal ultrasound were used to screen 1601 women with a first- or second-degree relative who had ovarian cancer. There were 61 positive tests, leading to six ovarian cancers, five stage I. There were five additional cancers, three ovarian and two peritoneal, reported 2–44 months after the last test [116].

The available evidence is not sufficient to determine if the sensitivity and specificity of any single ovarian cancer screening test is adequate for routine application. The modalities may be complementary when used together. The cost of a test such as TVU, as well as the risks and costs associated with surgical evaluation of any positive test result, are potential impediments to general screening. Prospective screening trials to evaluate these modalities are required.

DESIGN FEATURES

Objectives and Global Design

The PLCO trial is designed to determine, in screenees ages 55–74 at entry, whether:

In females and males

- screening with flexible sigmoidoscopy (60-cm sigmoidoscope) can reduce mortality from colorectal cancer, and
- screening with chest X-ray can reduce mortality from lung cancer.

In males

- screening with DRE plus serum PSA can reduce mortality from prostate cancer.

In females

- screening with CA125 and TVU can reduce mortality from ovarian cancer.

The secondary objectives are: (1) to assess screening variables other than mortality for each of the interventions including sensitivity, specificity, and positive predictive value; (2) to assess the incidence, stage, and survival of cancer cases; (3) to investigate the mortality predictive value of biologic and/or prognostic characterizations of tumor tissue as intermediate endpoints; and (4) to conduct biomolecular and genetic research into factors associated with cancer carcinogenesis and promotion, as well as the early detection of these factors.

The design is a two-armed, randomized, controlled trial with 37,000 females and 37,000 males, ages 55–74 at entry, in each of the two arms. Ten screening centers (SCs) will each recruit approximately 5000 to 30,000 individuals to reach the total of 74,000 females and 74,000 males. Minority representation in the aggregate participant population is sought in appropriate numbers. Participants in the control arm receive their usual medical care. In the intervention arm, men are screened for prostate, lung, and colorectal cancers, and women are screened for ovarian, lung, and colorectal cancers. (See Figure 1.) The PSA and CA125 screening tests are performed at the initial visit at entry to the trial, then annually for 5 years. The DRE, TVU, and chest X-ray exams are performed at entry and then annually for 3 years, except that there are only two annual repeat x-ray exams for participants who never smoked. Sigmoidoscopy is performed at entry and then at the 5-year point. All participants will be followed for at least 13 years from randomization.

Design Options Considered

A major design issue was whether to undertake separate trials for each of the cancer sites and corresponding screening modalities or combine them in some way. After a detailed examination of the costs of separate trials and various combinations of cancer sites, it was concluded that the most efficient use of resources would be to evaluate screening for the four cancers in one trial, thereby taking advantage of the efficiency of using one common administrative structure and one coordinating center (CC). It was also decided to use combinations of the screening tests deemed ready for evaluation, as described above, rather than evaluating each test individually. So, for example, DRE and PSA are used together rather than doing a separate trial for each. There were two main reasons for this approach: cost constraints and the ability to evaluate the

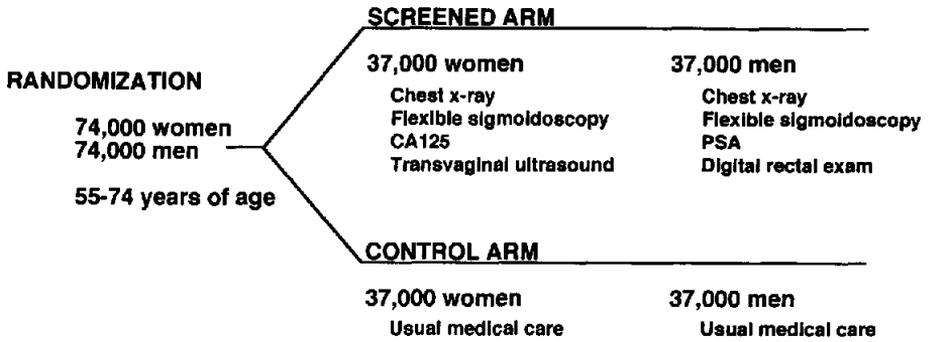


Figure 1 Schematic of the PLCO trial design.

combined, more-intensive interventions first to see if the combination works. If it does not, then testing the individual procedures is not warranted. If it does, then the individual tests can be evaluated subsequently.

Several overall designs were then considered [117]. The two primary competitors were the reciprocal control design and the all-versus-none design. The reciprocal control design would have had three arms: one devoted to screening for prostate or ovarian cancer, the second to colorectal cancer screening, and the third to lung cancer screening. Since screening would be undertaken for only one cancer site per gender in any given arm, the other arms would serve as controls. It was ultimately decided that this design would not be feasible because of the cost of bringing all participants in for screening and the anticipated substantial levels of contamination, because all participants would be coming in for screening and be aware that participants in the other arms were receiving other screening tests, which they would then request. The all-versus-none design was thus chosen in which participants would be randomized to one of two arms. One arm would serve as a control, while screening for all cancers would be done in the other arm, in the spirit of a multiphasic screening endeavor. Use of the all-versus-none design makes the reasonable assumptions for the cancers and screening tests under study that the tests for each cancer do not detect any of the other cancers, and that the endpoints—death from each of the four cancers—are not related.

Within this design it was further decided to employ the so-called “stop screen” approach in which screening is performed for a fixed number of years or screening rounds and then stopped, but follow-up continues to ascertain endpoints [118]. This approach was chosen primarily because it had been used successfully in screening trials for breast and colorectal cancers and because it is the only design that allows a direct assessment of overdiagnosis, a topic of considerable concern for prostate cancer screening in particular.

Specific Design Choices

The initial choice of four annual screens, at baseline (T0) plus three annual re-examinations (T1, T2, T3), later expanded to six screens (T0–T5) for PSA and CA125, was a trade-off between enough screens to produce an effect and

resources. Three or four screening rounds were sufficient in breast cancer screening trials [119, 120]. This also allowed sigmoidoscopy initially to be scheduled at the beginning and end of screening at a 3-year interval as suggested by many at the time the trial began. The annual interval between screens was chosen as the most frequent yet practical interval if screening is shown to be effective. Compared to less frequent screening, an annual interval also increases the likelihood of detection of a broad spectrum of the preclinical conditions in the natural history of the cancers under study. A longer interval might allow some rapidly growing lesions, which might be a source of mortality but which could be cured if found early, to escape detection. A minimum of 10 years of follow-up was initially decided upon to allow sufficient time for any mortality reduction from screening to emerge. Follow-up intervals of 7 years or more were typically required in breast cancer screening trials [119, 120], and it was assumed that the longer natural history of prostate cancer, and perhaps other cancers under study, warranted a longer follow-up period. It was recognized that these and other design parameter choices were based on the best information at the time and may be subject to change as a result of data gathered during the trial and other information.

The original age range at entry was 60–74 years. Prostate cancer screening was the driving force for the trial, and the lower age limit was based on the sharp increases in prostate cancer mortality beginning about age 60. The upper age limit was based on a combination of increasing mortality from other causes and anticipated reduction in compliance above age 80. Enrollment of individuals in the 55–59 age group began in January 1996 on the advice of the trial's Monitoring and Advisory Panel (MAP). The MAP based its recommendation on prostate cancer, considering trade-offs between younger and older ages and recognizing the natural history of untreated disease and the prognostic importance of age at diagnosis. While the importance of prostate cancer increases with age, so does the competing mortality. It is reasonable to expect that the optimal upper age to initiate screening is not above 75 years, given the relatively slow natural history of screen-detected prostate cancer, the decreasing proportion of patients over 75 who undergo surgery, and the high competing mortality. Many believe a somewhat lower age range, with a lower age limit of 50 or 55 years, is preferred for detecting progressive but curable prostate cancers in men with longer life expectancy. The impact of this change in age range on the trial's sample size is discussed below.

The time line for this trial is shown in Figure 2. This includes the 2-year pilot phase for protocol development and vanguard recruitment and screening, followed by a main phase that comprises recruitment and initial screening of participants in years 3–9, follow-up and additional screening through year 14, final follow-up through year 22, and data analysis through year 23. This time frame represents an experience-based extension of the original target to allow sufficient time to recruit the very large population required for this trial and follow this population for relevant cancer and mortality endpoints.

Pilot Phase

Protocol Development

During the first 6 months of the pilot phase, the trial investigators addressed the following components of the trial protocol:

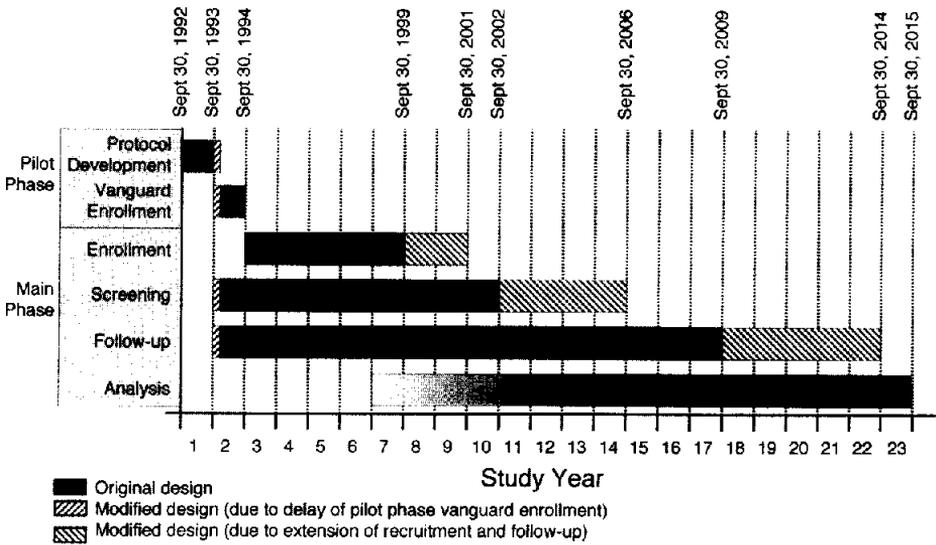


Figure 2 The PLCO time line.

1. eligibility requirements for participants entering the trial;
2. mechanism for notifying trial participants of screening results and encouraging them to seek further work-up of suspicious or positive results;
3. work-up of participants with suspicious or positive screens, including discussion of what further tests are required and in what sequence;
4. mechanism for providing appropriate therapy for cancer (or other lesions) detected by the screening program;
5. procedures for establishing and monitoring quality control of screening examinations;
6. procedures for follow-up of all randomized participants, monitoring of compliance with screening, determining cancer incidence, ascertaining cause of death, and correlating biological and tumor characteristics with mortality; and
7. educational materials for controls and for screened participants.

A number of these topics are discussed further below. The details of these and other protocol components, such as the questionnaires used in this trial, reside in the manual of operations and procedures.

Pilot Enrollment Period

Certain pilot studies were planned for the first 2 years of the trial. Decisions on the long-term commitment to the trial were based on these studies. The major activities carried out during the pilot recruitment period are summarized below:

1. Test and evaluate recruitment and randomization procedures by enrolling participants at each SC. Ten SCs enrolled 12,000 participants during the pilot phase.

2. Work out the detailed logistics by actually performing the screening examinations on the pilot-phase participants allocated to the screened arm. It was expected that all screening would take place during a single visit to an SC lasting no longer than 2 hours.
3. Assess background level of usage (contamination) of each screening modality among the participants actually randomized at each SC.
4. Assess compliance for each of the screening modalities at each SC.
5. Test in actual practice all of the data forms and procedures developed during the first year of the pilot phase.
6. Establish procedures to collect, ship, and analyze blood for PSA and CA125 assays and to collect and ship blood and tissue samples for the biorepository.

Each SC identified recruitment sources and strategies appropriate to the local situation. Randomization and enrollment of participants, initially scheduled to begin in September 1993, began in November 1993 when the protocol was completed and the screening procedures set up. During the pilot phase, representatives of the CC attended all meetings of the investigators and were responsible for documenting all decisions reached and compiling the trial protocol as it developed. In addition, the CC was responsible for developing appropriate trial forms, setting up data entry and editing systems, and writing a manual of operations for all procedures to be used in the trial. During this phase, efforts were made to standardize all examinations to the extent deemed feasible in the trial settings. Efforts were made to monitor performance at each SC and to correct any deficiencies so as to maintain adequate recruitment for the main phase of the trial. Participants randomized during the pilot phase are treated as a vanguard group and are included with the participants recruited later in the main phase.

Main Phase

Pilot-phase activities were concluded satisfactorily, and full-scale recruitment to the main phase of the trial began on September 30, 1994. Each SC is randomizing approximately 5000 to 30,000 participants into the trial during years 3–9 of the trial. After recruitment is complete, further screening is required for 5 years, and annual determination of cancer incidence and of deaths among trial participants is needed for the remaining years of the trial.

Trial Population and Exclusion Criteria

Proposals were solicited from all groups nationwide capable of assembling the necessary staff and facilities to recruit participants, conduct the screening, and follow all randomized participants for at least 10 years after entry into the trial. Proposals were peer reviewed. Selection was competitive. The population under study therefore comprises volunteer participants recruited from a variety of organizations. Characteristics of the enrolled population are ascertained via a baseline questionnaire. Potential participants are excluded for the following reasons:

1. men and women who at the time of randomization are < 55 or > 74 years of age;
2. individuals currently undergoing treatment for cancer, excluding basal cell and squamous cell skin cancer;
3. individuals with known prior cancer of the prostate, lung, colon, rectum, or ovary;
4. individuals with previous surgical removal of the entire prostate, one lung, or the entire colon;
5. individuals who are participating in another cancer screening or cancer primary prevention trial;
6. males who have taken Proscar (Finasteride) in the past 6 months;
7. males who have had more than one PSA blood test in the past 3 years;
8. individuals who have had a colonoscopy, sigmoidoscopy, or barium enema in the past 3 years;
9. individuals who are unwilling or unable to sign the consent form.

The intent of exclusion criteria 7 and 8 is to exclude individuals who might be undergoing the screening tests being studied in this trial. These criteria were implemented on April 15, 1995 on the advice of the trial's MAP to minimize contamination in the trial population. During protocol deliberations, it was recognized that the level of contamination by participants already receiving one or more of the trial interventions was a very important parameter about which there was considerable uncertainty. It therefore was decided to allow a more open recruitment at the start so as to measure directly the contamination levels, with the option of making adjustments in the main phase of the trial. That adjustment was made on April 15, 1995 after data revealed that the level of PSA testing among trial participants was nearly double that anticipated prior to the initiation of recruitment.

Prior to October 1996, women with previous surgical removal of both ovaries were excluded from the trial. However, enrollment of women proved more difficult than for men, and the MAP recommended that the trial enter all women, regardless of ovary status, to increase enrollment of women and thereby enhance the trial's ability to reach a valid scientific conclusion regarding the colorectal and lung cancer endpoints. Thus, beginning in October 1996, women without ovaries were no longer excluded.

Consent

At the outset, each SC had the option of using one of two informed consent procedures. Prerandomization consent requires informed consent from participants before randomization into the trial. It was believed that this consent, covering all aspects of participation including screening and data collection, could lead to greater contamination, but could also lead to greater compliance. Dual consent randomization consisted of an initial consent from all participants for administration of the baseline questionnaire and follow-up for cancer incidence and vital status. A second consent was then obtained only from participants randomized to screening, the purpose of which was to obtain consent to conduct the screening tests. This approach could reduce contamination, but could also lead to lowered compliance. Three of the initial ten SCs chose the

dual consent approach. However, after several years of recruitment, all centers were directed to switch to a single consent because the extra complexities of the dual consent were impeding enrollment and compliance with screening.

A prototype consent form was designed using standard language across SCs describing the randomized design of the trial, the obligations of the participants, the nature of the screening tests being studied, and the potential risks and benefits of participation. Discomforts associated with the screening procedures were listed, including the possibility of discomfort from the physical examination of the ovaries, the DRE, the sigmoidoscopic examination, and the TVU. Risks mentioned included the very rare perforation of the bowel during sigmoidoscopy, the very slight chance of a vaginal tear from the ultrasound probe, and the small amount of radiation received as part of the chest X-ray. The possibility of local bruising or bleeding at the puncture site of the blood draw was also noted. The consent form indicated that diagnosis (and treatment) of cancers detected in this trial might not prolong life and that the screening tests might falsely suggest that a person has cancer. Consequently, some participants might suffer pain, anxiety, and expense that would not have occurred if the individuals had never undergone the screening tests.

The consent form was reviewed and approved by the National Cancer Institute (NCI), the National Institutes of Health (NIH) Office of Protection from Research Risks, and the U.S. Office of Management and Budget. The prototype is part of the PLCO manual of operations and procedures. Individualized versions, possibly involving minor revisions, were approved by the institutional review boards of each SC. Consent is also sought for use of biologic samples. Consent issues are discussed elsewhere in this supplement [121].

Randomization

Individuals who meet the eligibility criteria are randomized individually into intervention and control arms. The randomization scheme uses blocks of random permutations of varying lengths and is stratified by SC, gender, and age. Random assignment is implemented using compiled software and encrypted files loaded on SC microcomputers. As each person is successfully randomized into the trial, data including name, gender, date of birth, and study arm are automatically stored in encrypted data tables. At the same time, a second protected set of synchronized tables, stored on a backup device, are also updated. This device also carries all of the randomization software. In the event that the system crashes, randomization can be performed directly from the backup device linked to another computer.

Screening Procedures

Whenever possible, screening examinations are conducted in a central screening facility to promote efficiency, accelerate the examination process, eliminate confusion and anxiety among screenees, facilitate patient education, and enhance compliance. The goal is to complete screens for all cancer sites within a 2-hour period for each screenee. The trial protocol specifies the qualifications, experience, and training of the examiners; equipment requirements;

examination procedures; and definitions of positive tests. To enhance acceptability of the screening process by screenees, the protocol recommends but does not require that technologists of the same gender as the screenee perform digital rectal and sigmoidoscopic examinations, as well as palpation and TVU examinations of the ovaries.

The blood draw for the PSA or CA125 test is usually performed first (for males blood is drawn prior to the DRE) and includes the collection of up to 45 mL of blood. The specimens are centrifuged, and the serum is separated from the clot and frozen within 2–4 hours of blood collection. Samples are stored at -70°C or colder and shipped weekly overnight on dry ice to the central laboratory (LAB) at UCLA for PSA or CA125 assay. Blood not used for the PSA or CA125 test is stored in a central repository at -70°C for future research. The biorepository is discussed in another paper in this supplement [122]. The LAB transmits assay results electronically to the SCs. A PSA result > 4 ng/mL is considered positive, while a CA125 result > 35 units/mL is considered positive.

A posteroanterior chest X-ray is taken by a qualified technologist and interpreted by a radiologist. The X-ray is taken using dedicated high-kV equipment (approximately 110–140 kV) at a tube-to-film distance of 6–10 feet. A wide latitude film with a 12:1 standard grid or higher is used. The exam is considered negative if evaluation reveals midline structure and heart to be of normal size and not displaced or enlarged, and pulmonary parenchyma reveals no abnormality suspicious for cancer. The exam is positive (suspicious) for lung cancer if evaluation reveals any of the following pulmonary abnormalities: nodule, mass, hilar or mediastinal lymph node enlargement, major atelectasis/lobar collapse, infiltrate/consolidation/alveolar opacity, or pleural mass.

The DRE is performed by a qualified medical examiner. The participant may bend at the waist over the end of the examination table, kneel in the knees-to-chest position, or lie in the lateral decubitus position with knees pulled up to chest. The examiner applies a lubricated gloved index finger at the 6 o'clock position to relax the sphincter and then introduces the finger into the anal area to palpate the prostate. The examiner explores the anterior portion of the rectal vault, i.e., the base, apex, and lateral lobes of the prostate. The exam is considered negative if the prostate is symmetric, soft, and non-nodular. The result is positive (suspicious) if there is nodularity or induration of the prostate gland or if the examiner judges the prostate to be suspicious based on other criteria such as asymmetry or loss of anatomic landmarks.

Preparation for the ovarian palpation examination requires the participant to empty her bladder within 30 minutes of the examination. The participant is placed in the dorsal recumbent position, in stirrups and draped. The qualified examiner notes right and left adnexa and ovaries separately, and a rectovaginal examination is performed. The cervix is not visualized. The examiner notes if there is blood on the glove after examination. The examiner is blinded to the results of the TVU examination. Findings are considered negative if an adequate examination is completed and no ovarian masses or other abnormalities are detected. For obese participants with nonpalpable ovaries, the examination is considered negative. A positive (suspicious) result is defined as a palpable ovarian mass or cul-de-sac nodularity.

The TVU examination is performed prior to the flexible sigmoidoscopy. It is performed by a qualified sonographer using a 5–7.5-MHz transvaginal probe. The sonographer images both the left and right ovaries in the transverse and longitudinal planes. The examiner searches for no less than 5 minutes per ovary for each ovary to ensure an adequate search for the ovaries; however, if the iliac vessels are visualized and no ovaries are visualized, the examiner may conclude the search for the ovaries. The examiner is blinded to the results of the ovarian palpation examination.

Using calipers, the sonographer takes measurements on the image along the major and minor axes in both transverse and longitudinal planes. A negative examination is one that clearly defines the ovaries with no abnormalities found. If one or both ovaries cannot be visualized, the examination is also considered negative, provided no other abnormalities are found. The prolate ellipsoid formula ($\text{width} \times \text{height} \times \text{thickness} \times 0.523$) is used to calculate the volume of each ovary and/or cyst. A finding positive (suspicious) for ovarian cancer consists of one or more of the following features: any ovary or cyst greater than 10 cc in volume, any solid area or papillary projection extending into the cavity of a cystic ovarian tumor of any size, or any mixed (solid/cystic) component within a cystic ovarian tumor.

Sigmoidoscopy is performed by a qualified examiner using a 60-cm flexible sigmoidoscope. Bowel preparation includes a Fleet enema, repeated once or twice if necessary, on the morning of the exam. A brief medical history is obtained before sigmoidoscopy, and vital signs are measured before and after the procedure. After a DRE, the examiner introduces the lubricated scope and visualizes the rectum and the colon as the scope is advanced. Air for inflation or exposure is used at the discretion of the examiner. Once the scope is fully advanced, the examiner withdraws the scope and visualizes the colon and the rectum as the scope is withdrawn. The examination is considered negative if the examiner is able to advance the scope at least 50 cm, there is adequate bowel preparation, and no abnormalities are seen. A finding that is positive (suspicious) for colorectal cancer or neoplastic lesion is defined as visible (or palpable from rectal exam) evidence of a mucosal abnormality including rectal nodule(s), rectal and colon mass(es), or rectal and colon polyp(s).

The examination procedures used in this trial also identify benign abnormalities such as colon diverticuli seen on flexible sigmoidoscopy or heart enlargement seen on chest X-ray. These are reported to the participants and their physicians. Training and qualifications of the examiners and documentation, reporting, and quality assurance procedures for each screening test are described in detail in the relevant chapters of the trial's manual of operations and procedures. Details of the quality assurance procedures for the trial appear in another paper in this supplement [123].

Rescreening Protocol

For individuals with negative screens, a scheduling and tracking procedure was implemented at each SC to ensure regular attendance at repeat screens. A similar procedure was established for individuals with findings that are considered suspicious or positive by screening, but for whom subsequent diagnostic evaluations do not reveal a prostate, lung, colorectal, or ovarian cancer.

Diagnostic and Therapeutic Follow-Up

As described above, a major part of the PLCO trial protocol is an active screening intervention aimed at achieving high compliance and high-quality examinations. If participants are found to have an abnormality suspicious for cancer on any one of the screening tests, they are notified of their test results within about 3 weeks. These individuals are referred to a physician of their choice for appropriate diagnostic workup and treatment. The directing role of the SCs ceases after screening, communication of findings, and referral. The trial has no direct control over diagnostic and therapeutic procedures.

The trial is designed to have adequate statistical power for detecting decreases in mortality separately for each cancer site. Therefore, participants with positive screening tests must undergo appropriate diagnostic and therapeutic procedures in order for the screening process to produce an effect. An attempt was made to develop uniform diagnostic and treatment protocols for each organ during the protocol development phase. However, after lengthy discussion it was decided that the trial could not dictate medical practice. In addition, by design, the PLCO trial aims to measure the independent effects of screening and early cancer detection. The investigators wanted to avoid activities that might unfairly bias the type of cancer therapy received by participants in the screened arm. For this reason, the investigators decided not to promulgate treatment guidelines for participants with screen-detected cancers. Further, the funding agency could not pay for diagnostic and therapeutic procedures. Thus, there are no PLCO-prescribed protocols for diagnosis and therapy.

Individuals with positive findings or with symptoms not arising as a result of a PLCO-initiated screening test undergo diagnostic procedures determined within their own medical care environment. If requested, referral physicians are provided with standard-of-practice guidelines for diagnostic procedures by the local PLCO SC. All sequelae to screening tests and/or subsequent diagnostic procedures are identified and recorded. These include any morbid or medical events potentially associated with a positive or negative screen or a diagnostic procedure subsequent to a positive screen. Individuals diagnosed with prostate, lung, colorectal, or ovarian cancer as a result of a trial-initiated screening test are referred to qualified medical personnel for appropriate therapy. Treatment is expected to be in accordance with current accepted practice for appropriate stage of disease, age, and medical condition of the participant. Physicians are referred to published references such as those listed in the NIH's Physician's Data Query (PDQ) system for treatment guidelines, if requested.

Data pertaining to diagnosis and (at least initial) treatment of all PLCO cancers are collected in both the screened and control arms of the trial to enable uniform staging and other prognostic criteria to be applied. In addition, the SCs actively track participants with screening test abnormalities and actively seek, collect, assemble, organize, and abstract medical record information related to diagnostic follow-up and treatment. These data collection activities may, and probably do, stimulate contact with physicians and diagnostic follow-up among intervention arm participants who have screening test abnormalities.

Endpoints

Cause-specific mortality for each of the PLCO cancers is the primary endpoint. In addition, cancer incidence, stage shift, and case survival data will be

collected and monitored as secondary endpoints to help understand and explain the results. Biologic and/or basic prognostic characteristics of the cancers will be measured and correlated with mortality to determine the mortality predictive value of these intermediate endpoints. Collection of endpoint data will involve several processes to ensure as far as possible complete and unbiased ascertainment of endpoints. These include both active and passive follow-up activities, as well as cause-of-death review procedures.

All prevalent and incident PLCO cancers and all deaths that occur among participants during the trial will be ascertained, primarily by means of an active follow-up process involving a mailed annual study update (ASU) questionnaire. The ASU asks about type and date of cancer diagnosed in the previous year. This is accompanied by a follow-up locator form approximately every other year that updates the participant's mailing address and personal contacts. Participants who do not return the ASU are contacted by repeat mailing and/or telephone. Compliance with the ASU is expected to be about 95%. SCs that are located in regions having population-based cancer registries also make use of these resources for follow-up.

The ASU activity will be supplemented by linkage to the National Death Index (NDI) to enhance completeness of endpoint ascertainment. Each SC collects and retains the information necessary to search the NDI files: name, social security number, and date of birth. Other useful information is state of birth, last known state of residence, race, and marital status. Procedures for and timing of NDI searches will be determined as the trial progresses. Although one could expect substantially complete endpoint identification using the NDI, the PLCO trial uses an active approach as the primary follow-up process to obtain more timely information and to promote contact with participants so as to enhance acquiring consent and clinical follow-up information should a participant develop cancer or die.

Underlying cause of death is determined for all participants who die during the trial. Since the true underlying cause may not always be evident or accurately portrayed on the death certificate, the trial will employ a death review process to assess cause of death in a uniform and unbiased manner. Each SC collects and makes available for the death review process all documents needed for ascertaining the underlying cause of death, including death certificates, pathology and other medical forms, available autopsy reports, and pathology slides as necessary. These documents will be reviewed by a panel of individuals with appropriate expertise who are not otherwise affiliated with the trial and who will be blinded as to the randomized arm of the deceased participants [124].

Etiologic and Early Marker Studies

To support these objectives, the PLCO collects and archives (1) baseline and supplemental demographic and risk factor information on all participants, and (2) specifically processed blood products and other tissue samples for molecular/genetic research. Serial, prospective collection of biologic samples (serum, plasma, red blood cells, and buffy coat) from screened participants will make future studies to evaluate new early detection markers of PLCO cancers inexpensive and rapid. It will also make possible molecular epidemiologic and etiologic risk assessment studies of the highest scientific quality.

A full discussion of the biorepository appears in a companion paper in this supplement [122].

Sample-Size Calculations

Sample size was calculated using the method suggested by Taylor and Fontana [125], modified to allow for arbitrary magnitude of screening impact, arbitrary sample-size ratio between screened and control arms, and arbitrary levels of compliance in the screened and control arms. Let N_C be the number of individuals randomized to the control arm and N_S be the number randomized to the screened arm, with $N_S = f N_C$, where f is a proportionality constant. For $0 < r < 1$, assume the trial is designed to detect a $(1 - r) \times 100\%$ reduction in the cumulative disease-specific death rate over the duration of the trial. Also let P_C be the proportion of individuals in the control arm who comply with the usual-care protocol and P_S be the proportion of individuals in the screened group who comply with the screening protocol. The total number of disease-specific deaths needed for a one-sided α -level significance test with power $1 - \beta$ is then:

$$D = \frac{[(Q_C + f Q_S) Z_{1-\alpha} - \sqrt{Q_C Q_S} (1 + f) Z_\beta]^2}{f (Q_C - Q_S)^2}$$

where $Q_C = r + (1 - r)P_C$ and $Q_S = 1 - (1 - r)P_S$. The number of participants required in the control arm is:

$$N_C = \frac{D}{(Q_C + f Q_S) R_C Y}$$

where Y is the duration of the trial from entry to end of follow-up in years and R_C is the average annual disease-specific death rate in the control arm expressed in deaths per person per year.

A one-sided hypothesis testing approach to sample-size calculation was employed based on the nature of the question being addressed. The PLCO trial is intended to provide definitive evidence of the effect of screening on cause-specific mortality compared to usual medical care, analogous to phase III placebo-controlled trials in the therapeutic setting. The focal question for each of the four cancers is whether screening reduces mortality. This is inherently a one-sided research question, implying a one-sided design and analysis approach. The question is not whether screening reduces *or* increases mortality. Determining whether screening increases mortality is not an objective of this trial. Furthermore, if the screening intervention has no effect or if it is harmful, the consequences in terms of a public health decision are the same—screening is not recommended. This further dictates a one-sided approach [126].

The estimation procedure is illustrated for prostate cancer for white males. Prostate cancer screening was the impetus for the trial and is the primary focus for sample-size calculations. Similar calculations can be done for the other sites using the data in Table 3. This illustration is based on calculations done for the original design prior to the pilot phase, when the eligible age range was 60–74 years and the trial duration was 10 years from randomization for each participant.

Table 3 Cancer Mortality Rates per Person per Year ($\times 10^{-5}$), Estimated Using 1983–1987 Data

Age (years)	Prostate		Ovarian		Colorectal	Lung
	White	Black	White	Black	White ^a	White ^b
Males						
50–54	3.5	11.1	—	—	20.8	88.4
55–59	11.5	31.9	—	—	39.6	165.4
60–64	30.4	80.4	—	—	64.6	252.6
65–69	71.1	174.1	—	—	104.4	367.6
70–74	137.8	332.3	—	—	156.1	470.2
75–79	244.8	515.7	—	—	216.0	543.9
80–84	402.8	838.7	—	—	296.0	555.0
85+	606.3	937.1	—	—	378.6	441.3
Females						
50–54	—	—	14.2	10.4	16.5	46.5
55–59	—	—	20.3	15.3	28.1	75.3
60–64	—	—	27.5	23.3	43.9	104.9
65–69	—	—	35.3	27.4	67.9	138.0
70–74	—	—	41.5	33.8	100.1	152.9
75–79	—	—	45.2	34.5	141.9	143.8
80–84	—	—	49.8	41.1	200.5	127.2
85+	—	—	44.7	35.0	289.2	103.5

^a Rates for black males are very similar. Rates for black females in age group 65–79 years are about 15% higher.

^b Average rate for black males in age group 65–79 years is about 13% higher. Average rate for black females in age group 65–79 years is about 20% lower.

Calculation of N_C requires an estimate of R_C . It was assumed that the trial would enroll an equal number of participants in each of three age strata: 60–64, 65–69, and 70–74 years. Because individuals recruited for screening trials are expected to be healthier than the general population, the usual cancer mortality rate obtained from national or registry data will overestimate the mortality rate of the participants, at least for the early part of the trial. Therefore, for a 10-year prostate cancer screening trial with men entered between the ages of 60–74, it was assumed that for the first 2 years the mortality rate in the control arm is 25% of the usual rate, for the next 3 years it is 50% of the usual rate, and for the last 5 years it equals the usual rate. The usual mortality rate was estimated by the unweighted average prostate cancer mortality rate for men ages 65–79 years. This age range was used to adjust for aging over the 10 years of the trial. The usual mortality rates from national data are shown in Table 3 [127]. The estimated rate for this example is $R_C = 103.763 \times 10^{-5}$.

Results of sample-size calculations for the trial are given in Table 4. These calculations assume a 10-year trial using a one-sided, 0.05-level test, $P_C = P_S = 1$, and possible mortality reductions as shown in a screened group compared to an equal-sized, usual care group ($f = 1$). The sample sizes are based on mortality rates for whites. Including blacks in the trial does not substantially alter sample size. A sample size of 37,000 (rounded up from 36,221 in Table 4) screened and 37,000 controls of each gender was chosen on the following basis. A high power of at least 90% is mandatory to yield a meaningful negative result, should that happen, and to achieve a high level of scientific validity

Table 4 Number of Participants Ages 60–74 Years at Entry Needed in Each Arm of the Trial

Site	Power	Mortality Reduction (%)			
		10	20	30	35
Prostate	0.9	153,577	36,221	15,078	
(males)	0.8	110,906	26,182	10,920	
Lung	0.9	76,721	18,095		
(males and females)	0.8	55,404	13,080		
Colorectum	0.9	177,208	41,794	17,397	
(males and females)	0.8	127,971	30,211	12,600	
Ovary	0.9		134,697	56,069	39,733
(females)	0.8		97,365	40,606	28,817

because a trial of this magnitude addressing these questions is not likely to be repeated. In addition, it was felt that for an effect of prostate cancer or colorectal cancer screening to be of public health importance, it must be at least 20% or greater. Given the magnitude of the lung cancer problem, it was felt that a screening effect of 10% or greater would be very important. To estimate whether a 20% effect for prostate cancer screening was realistic, two calculations were performed. The first used plausible stage shifts due to screening and survival by stage to project possible improved outcome for screen-detected cancers. The second used projections from a computer model [128]. Both gave mortality reduction estimates in the range of 25% with perfect compliance.

Power calculations are displayed in Table 5. With 37,000 men and women in each arm, the trial has 91% power to detect a 20% mortality reduction in prostate cancer mortality and 89% power to detect a 10% lung cancer mortality reduction. The power is nearly 90% to detect a 15% colorectal cancer mortality reduction and 99% for a 20% effect. For ovarian cancer, the power is nearly 90% to detect a 35% mortality reduction.

It was recognized that compliance will not be perfect in either randomized group. Contamination or drop-in will occur in the control arm ($P_C < 1$) and noncompliance or dropout is to be anticipated in the screened arm ($P_S < 1$). The target mortality reductions of 20% for prostate and colorectal cancers and 10% for lung cancer therefore are to be interpreted as effects that the trial seeks

Table 5 Power by Percent Reduction in Mortality with 37,000 Men and 37,000 Women in Each Arm

Site	Gender	Mortality Reduction (%)						
		5	10	15	20	25	30	35
Prostate	male	—	—	0.71	0.91	0.98	—	—
Lung	both genders	0.41	0.89	0.997	—	—	—	—
	female	0.17	0.41	0.69	—	—	—	—
	male	0.34	0.81	0.985	—	—	—	—
Colorectum	both genders	—	—	0.89	0.99	0.999	—	—
	female	—	—	0.56	0.79	0.93	—	—
	male	—	—	0.72	0.92	0.99	—	—
Ovary	female	—	—	—	0.45	0.62	0.77	0.88

Table 6 Percent Mortality Reduction Required When Compliance Is 100% in Both Groups, Based on a Mortality Reduction of 20% in the Presence of Noncompliance, as a Function of P_S and P_C

Compliance in the Control Group (P_C)	Compliance in the Screened Group (P_S)					
	0.5	0.6	0.7	0.8	0.9	1.0
0.5	—	100	67	50	40	33
0.6	90	71	53	42	34	29
0.7	77	56	43	36	30	26
0.8	59	45	37	31	27	24
0.9	48	39	32	28	24	22
1.0	40	33	29	25	22	20

to detect in the presence of whatever noncompliance and contamination exist in the populations. This implies that if there were perfect compliance, the mortality reductions would be greater since they would not be diminished by noncompliance.

One can assess the relationship between true effect size and level of noncompliance during the screening period by examining Table 6, which shows what the mortality reductions with perfect compliance would have to be to realize a 20% mortality reduction for various levels of noncompliance in the screened and control groups. For example, if 90% of participants in the screened group undergo a PSA test ($P_S = 0.9$) while 20% of controls are so screened ($P_C = 0.8$), then the prostate cancer mortality reduction from such screening would have to be 27% with perfect compliance for there to be a 20% effect in the presence of noncompliance. The 27% figure corresponds very closely to the modeling estimate. Thus, compliance of at least 90% and contamination of no greater than 20% for prostate cancer screening, particularly with PSA, were chosen as the target values for these parameters.

Inquiries into potential screening compliance and screening contamination for the four cancer sites being studied in this trial indicated that the ranges of reasonable target values at the time of initiation of recruitment were as shown in Table 7. In addition to direct contact with health maintenance organizations and existing SCs, published data from the 1987 National Health Interview Survey were used to gauge these effects [129, 130]. These numbers were necessarily somewhat subjective. Additional estimates were obtained directly from the trial population during the pilot phase, and further assessment will occur as the trial progresses, possibly leading to sample-size adjustment.

In the context of these levels of contamination and compliance, the required true levels of mortality effect (effect size) with perfect compliance are, approximately, lung 20%, colon 25%, and prostate 25%. These requirements are consistent with expected effects based on modeling efforts [74, 75, 103].

Regarding the ovarian cancer objectives of this trial, if the mortality reduction from screening for ovarian cancer were 35%, this design would have almost a 90% power to demonstrate this effect. However, if the mortality effect were only 25%, 84,000 screened women and an equal number of controls would be required to achieve 90% power. Thus, the ovarian component of this trial is to be viewed as a two-step process. Near the end of the screening phase of the trial, sufficient cases of ovarian cancer should accrue to provide good estimates

Table 7 Design Contamination and Compliance Ranges Projected by Modality

	Compliance (%)	Contamination (%)
Digital rectal exam	>90	<20
Prostate-specific antigen	>90	<20
CA125	>90	<10
Ovarian palpation	>85	<10
Transvaginal ultrasound	>85	<10
Sigmoidoscopy	>85	<15
Chest X-ray	>85	<40

of sensitivity for each screening modality. Specificity and predictive value can also be estimated. If as a result any one or combination of the tests appears sufficiently promising to justify a full mortality study, the female population base of this trial could be supplemented or a meta-analysis of data from this trial and other relevant studies could be done to increase power.

As noted above, in January 1996 the lower age limit for trial participation was reduced from 60 to 55 years. Given the lower mortality rates in the 55–59 age stratum, this would ordinarily imply the need for an increase in the sample size. However, this protocol change took place after the April 1995 eligibility criterion change, also noted above, to exclude men who had prior repeat PSA screening, thereby reducing the contamination level. Sample-size estimates for prostate cancer screening for the age range 55–74 years are shown in Table 8. For compliance of 90% and a revised estimate of contamination of 10–15%, a sample of 37,000 men (and therefore 37,000 women) in each trial arm is still appropriate. A similar conclusion holds for the other cancer sites as well. As mentioned, this estimate is monitored regularly during the enrollment phase of the trial to determine if adjustment is required.

Based on the monitoring of design parameters, further protocol modifications were adopted in December 1998. These were to change from a 3-year to a 5-year interval for flexible sigmoidoscopy for individuals who had not yet had their second exam, and at the same time to add year 4 and 5 PSA and CA125 tests. Also, the remaining third annual chest X-ray exams are offered only to current or former smokers, and follow-up is extended 3 years, so that all participants will be followed at least 13 years from randomization. A final change was that the ovarian palpation exam, which had been part of the original protocol, was eliminated.

Table 8 Number of Males Required in Each Arm to Achieve 90% Power with Age at Entry Range 55–74 Years, as a Function of P_s and P_c

P_c	P_s		
	0.85	0.90	0.95
0.80	53,057	45,338	39,134
0.85	46,087	39,787	34,650
0.90	40,440	35,225	30,918

The interval between flexible sigmoidoscopy was lengthened to coincide with recommendations in the community and was based on preliminary information suggesting that sigmoidoscopy at 3 years finds polyps, but very few are likely to be of any significance. A delay of 2 years was expected to yield more polyps and cancers, leading to a greater potential for mortality reduction. The addition of 2 extra years of PSA and CA125 blood tests and at least 3 additional years of follow-up were adopted to provide assurance of sufficient screening effect and statistical power in the event that initial design assumptions were incorrect. The final round of chest X-ray testing for individuals who never smoked was eliminated because of the very low yield of this exam. Finally, the ovarian palpation exam was deleted because of very low yield and the fact that a very high proportion of women participating in the trial regularly underwent pelvic examination, thereby diluting any possible effect of the palpation exam.

Data Reporting

The data management system for the trial has the following operational requirements: ability for the NCI and the CC to access SCs remotely, synchronization of databases on multiple platforms, preparation of high-quality analysis datasets, secure backup and archiving of data, and robust configuration management. Data are exchanged via a distributed data entry system and are transmitted among collaborators via common carrier service using modems, with transmission to the NIH mainframe on a regular basis. A detailed description of the system is provided in a companion manuscript in this supplement [131].

Various forms were developed for collection of information in this trial. Included are eligibility and consent forms, male and female versions of the baseline questionnaire, a dietary questionnaire, examination forms for each screening procedure, diagnostic evaluation and treatment forms, and a questionnaire for regular follow-up of participants. Additional forms are developed as needed as the trial progresses. Most forms are scanned into the data system. All trial forms are catalogued in the trial's manual of operations and procedures.

Pertinent data items include but are not necessarily restricted to the following:

1. participant trial identification number;
2. participant demographic and risk factor information;
3. participant randomized group, date of birth, and date of entry into the trial;
4. date and result of each screening test for each screened group participant;
5. sufficient information regarding diagnostic procedures performed as a result of a positive or suspicious screening test to allow determination of whether a cancer was or was not diagnosed as a result of screening;
6. for all screening tests, detailed physical findings and any complications or morbid events possibly associated with the test, and description of any diagnostic procedures subsequent to a positive test;
7. for every PLCO cancer diagnosed during the trial in both randomized groups, date of diagnosis, histology and stage at diagnosis, and initial therapy;

8. for every death that occurs during the trial in both randomized groups, date of death and cause of death as noted on the death certificate, coded using the ICD-9-CM classification.

Data Analysis Plan

Analytical methods used will include standard descriptive statistics and techniques such as regression, analysis of variance and covariance, analysis of rates and proportions, contingency table methods, and analysis of survival data. New methods of analysis or modeling will be developed and applied as needed.

Topics addressed in the pilot phase included recruitment progress and problems, evaluation of randomization and the delivery of screening, participant follow-up, compliance, contamination, assessment of quality assurance practices, evaluation of the distributed data entry system, evaluation of data forms processing and information flow, and operation of the biorepository. Intra- and intercenter comparisons in the areas mentioned above are conducted using descriptive statistics to monitor progress and practices. Compliance and contamination are examined, and quality assurance is monitored using summary statistics on screening results and comparing screening exams with quality assurance exams.

Topics to be addressed in the main phase include the following:

1. quality assurance, recruitment, delivery of screening, follow-up, compliance, contamination, and information system evaluation—all continued from the pilot phase;
2. characteristics of the enrolled population;
3. determination of screening test operating characteristics including sensitivity, specificity, and predictive value;
4. prevalence and incidence;
5. characteristics of the four PLCO cancers including stage, histology, survival, prognostic variables, and interval versus screen-detected cases;
6. lead-time estimation and modeling;
7. incidence rates of interval cancers and advanced stage disease;
8. cause-specific and all-cause mortality;
9. therapy of cancer cases;
10. complications of interventions;
11. surrogate end points;
12. cost variables;
13. quality of life.

Each of these areas will be examined for the trial overall, and variability among SCs will be investigated. Cancer site-specific mortality rates for each of the four PLCO cancers will be calculated and compared between the screened arm and the control arm on an intent-to-treat basis as the primary analyses in this trial. These mortality rates are calculated as the number of cause-specific deaths per 1000 person-years at risk among all individuals randomized to a given arm of the trial. In addition, the death rates from other causes and total mortality will be scrutinized to assess the comparability of the randomized

populations. The rates will be calculated from time of entry, yearly and cumulatively, by age group at entry and for all ages combined, and by gender and for both genders combined where appropriate. These rates will be compared using Poisson tests, Poisson regression analysis, and the log-rank test. Stratification will be taken into account. Since it is likely that the mortality or hazard rates in the screened and control arms of this trial will not be proportional throughout the trial, statistical methods that apply to this situation will also be considered.

Sensitivity, specificity, and predictive value will be estimated for each screening test and test combination for each cancer site for each screen. After the completion of screening, overall estimation of these parameters will be undertaken. In addition to the standard calculation of sensitivity as the ratio of screen-detected cases to all cases found through some interval after a screen, usually 1 year, a second method will be employed that uses the proportion of expected incidence in time periods after a negative screen [132]. Prevalence will be calculated as the number of cancers detected per 1000 individuals screened on the first screen for each cancer site and SC, and will be pooled to indicate overall prevalence. Incidence will be similarly calculated as the number of cancers per 1000 person-years at risk. Incidence rates will be calculated yearly and cumulatively over the course of the trial. The ratio of prevalence to incidence will be used as an estimate of the mean duration of preclinical disease.

For cancer case characteristics such as stage and histology that carry prognostic implications, the distribution of each characteristic will be calculated for each cancer site among usual care group cases, all screened group cases, screen-detected cases, and interval cases. The distributions can be compared using chi-square tests. Survival distributions will also be calculated for the same subsets of cancer cases using the Kaplan-Meier method and will be compared using the log-rank test and Cox proportional hazards regression methods. These distributions will be calculated cumulatively through each successive year of the trial to assess whether there is some suggestion of screening benefit. However, these intermediate endpoints cannot be relied upon for definitive evaluation because they are subject to lead-time and length biases.

Estimation of lead time is important as an intermediate indicator of the early detection capability of the screening procedures in the trial. Average lead time will initially be estimated using the prevalence to incidence ratio under the assumption of an exponential distribution of preclinical duration. After screening has been completed, other modeling approaches will be employed. These include the Day-Walter model [133, 134], which allows estimation of the lead-time distribution, and newer approaches that examine differences in long-term case survival rates to estimate mean lead time [135, 136].

The specific therapy used for each PLCO cancer will be collected in the screened and control populations within each disease stage. These data will be examined to assess whether, within each stage of disease, therapy is comparable between randomized groups to eliminate any confounding effect of therapy in assessing the impact of the screening protocol. Treatment distributions can be compared using chi-square procedures. In addition, there will be consideration of alternative endpoints that can act as a surrogate for mortality, but which can be ascertained sooner, perhaps at reduced cost. Possible surrogates

include advanced stage rate [137], functions of stage or other cancer case characteristics, or functions of lead time.

Complications of the screening, diagnostic, and treatment procedures administered to trial participants will be monitored. These include any medical complications or risks and any mortality potentially related to trial procedures, particularly the more invasive procedures such as colonoscopy or laparotomy that might follow a positive colorectal screen or ovarian screen, respectively. These will be examined for each cancer site and each SC at least for each month up to 1 year after a screening episode. Cancer incidence will also be tracked carefully to alert the investigators to the possibility of substantial overdiagnosis of one of the cancers being studied. This is thought to be a problem particularly for prostate cancer. Guidelines for termination in the event of adverse effects of the screening process will also be developed by the MAP.

Sequential monitoring will be an integral part of this trial and will be conducted separately for each PLCO cancer site. Statistical monitoring guidelines have been established by the trial investigators and the MAP to use in periodic examinations of the emerging data from the trial to decide on continuation or termination. Beginning in study year 9, the accruing mortality data and secondary endpoints will be examined annually by the MAP, which will determine if and when a protocol change is warranted that would result in an early decision about screening for one or more of the cancer sites. The data will be analyzed in two ways: one addressing the prospect of termination due to a significantly large effect, and one focusing on early termination due to a negligible effect. For early termination with a large effect, a variant of the O'Brien-Fleming boundary arising from Method 1 of Lan and DeMets [138] will be used. The test statistic will be a weighted log-rank statistic with weights linear in cumulative mortality. Choice of this combination of boundary and weights was based on power computation conducted using simulation methods. For early termination with a negligible effect, the stochastic curtailment procedure in Lin et al. [139] will be implemented. This procedure makes allowance for nonproportional hazards, frequently encountered in a screening trial.

DISCUSSION

The PLCO trial represents a major commitment of resources and personnel designed to evaluate early detection procedures that have great potential to reduce the burden of cancer in the population of the United States as well as other countries. The four cancers targeted for intervention are among the major sources of cancer incidence and mortality in the United States in the targeted age group. Within the constraints of the design assumptions, the sample size is sufficient to ensure scientifically valid assessment of the impact of the screening tests on cause-specific mortality. In addition, the trial will investigate secondary endpoints and disease natural history questions, and will provide data that can be used to address relationships among costs, risks, and benefits.

The screening tests being examined in this trial have been in use for many years, if not decades, but until now none has been rigorously evaluated in a proper scientific trial. The prospective randomized design was chosen for this undertaking because of the uncertainty in interpretation of observational studies of cancer screening due to selection bias and the difficulty in understanding

the reasons for changes over time in population cancer rates. Furthermore, it was decided to assess intervention for four cancers in one trial to take advantage of the cost savings achieved by having one administrative structure and one CC, as well as to mimic to some extent the multiphasic nature of screening in the community.

The PLCO timetable calls for recruitment to be completed in the year 2001 with follow-up continuing for another 13 years. The accruing data will be monitored on a regular basis, so that any findings on the main mortality outcome endpoints that occur sooner can be reported promptly. For any of the four cancer sites being studied, if the findings are positive, the trial will provide a quantitative estimate of the effect that will be used in cost-effectiveness planning for public health purposes. Negative findings will provide scientific evidence for abandoning a test and shifting resources elsewhere.

REFERENCES

1. Silverberg E, Lubera JA. Cancer statistics, 1989. *CA Cancer J Clin* 1989;39:3-20.
2. Cancer Facts & Figures-2000. Pub. 00-300M-No. 5008.00. Atlanta: American Cancer Society; 2000.
3. Gallion HH, van Nagell JR, Donaldson ES, et al. Adjuvant oral alkylating chemotherapy in patients with stage I epithelial ovarian cancer. *Cancer* 1989;63:1070-1073.
4. Gilbertsen VA. Cancer of the prostate gland. Results of early diagnosis and therapy undertaken for cure of the disease. *JAMA* 1971;215:81-84.
5. Jenson CB, Shahon DB, Wangenstein OH. Evaluation of annual examinations in the detection of cancer. *JAMA* 1960;174:1783-1788.
6. Chodak GW, Schoenberg HW. Early detection of prostate cancer by routine screening. *JAMA* 1984;252:3261-3264.
7. Chodak GW, Keller P, Schoenberg HW. Assessment of screening for prostate cancer using the digital rectal examination. *J Urol* 1989;141:1136-1138.
8. Cooner WH, Mosley BR, Rutherford CL Jr, et al. Clinical application of transrectal ultrasonography and prostate specific antigen in the search for prostate cancer. *J Urol* 1988;139:758-761.
9. Babaian RJ, Mettlin C, Kane R, et al. The relationship of prostate-specific antigen to digital rectal examination and transrectal ultrasonography. Findings of the American Cancer Society National Prostate Cancer Detection Project. *Cancer* 1992;69:1195-1200.
10. Thompson IM, Ernst JJ, Gangai MP, et al. Adenocarcinoma of the prostate: Results of routine urological screening. *J Urol* 1984;132:690-692.
11. Thompson IM, Rounder JB, Teague JL, et al. Impact of routine screening for adenocarcinoma of the prostate on stage distribution. *J Urol* 1987;137:424-426.
12. Wajsman Z, Chu TM. Detection and diagnosis of prostatic cancer. In: Murphy GP, ed. *Prostatic Cancer*. Littleton, Massachusetts: PSG Publishing; 1979.
13. Thompson IM, Zeidman EJ. Presentation and clinical course of patients ultimately succumbing to carcinoma of the prostate. *Scand J Urol Nephrol* 1991;25:111-114.
14. Resnick MI. Editorial comment. In: Rattiff TL, Catalona WJ, eds. *Genitourinary Cancer*. Boston: Martinus Nijhoff Publishers; 1987:94-99.
15. Friedman GD, Hiatt RA, Quesenberry CP Jr, et al. Case-control study of screening for prostatic cancer by digital rectal examinations. *Lancet* 1991;337:1526-1529.
16. Chodak GW, Schoenberg HW. Progress and problems in screening for carcinoma of the prostate. *World J Surg* 1989;13:60-64.

17. Chodak GW, Wald V, Parmer E, et al. Comparison of digital examination and transrectal ultrasonography for the diagnosis of prostatic cancer. *J Urol* 1986;135:951-954.
18. Clements R, Griffiths GJ, Peeling WB, et al. How accurate is the index finger? A comparison of digital and ultrasound examination of the prostatic nodule. *Clin Radiol* 1988;39:87-89.
19. Lee F, Littrup PJ, Torp-Pedersen ST, et al. Prostate cancer: Comparison of transrectal US and digital rectal examination for screening. *Radiology* 1988;168:389-394.
20. Lee F, Torp-Pedersen S, Littrup PJ, et al. Hypoechoic lesions of the prostate: Clinical relevance of tumor size, digital rectal examination, and prostate-specific antigen. *Radiology* 1989;170:29-32.
21. McClennan BL. Transrectal US of the prostate: Is the technology leading the science? *Radiology* 1988;168:571-575.
22. Torp-Pedersen ST, Littrup PJ, Lee F, et al. Early prostate cancer: Diagnostic costs of screening transrectal US and digital rectal examination. *Radiology* 1988;169:351-354.
23. Waterhouse RL, Resnick MI. The use of transrectal prostatic ultrasonography in the evaluation of patients with prostatic carcinoma. *J Urol* 1989;141:233-239.
24. Lange PH, Ercole CJ, Lightner DJ, et al. The value of serum prostate specific antigen determinations before and after radical prostatectomy. *J Urol* 1989;141:873-879.
25. Oesterling JE, Chan DW, Epstein JI, et al. Prostate specific antigen in the preoperative and postoperative evaluation of localized prostatic cancer treated with radical prostatectomy. *J Urol* 1988;139:766-772.
26. Stamey TA, Yang N, Hay AR, et al. Prostate-specific antigen as a serum marker for adenocarcinoma of the prostate. *N Engl J Med* 1987;317:909-916.
27. Partin AW, Oesterling JE. The clinical usefulness of prostate specific antigen: Update 1994. *J Urol* 1994;152:1358-1368.
28. Coley CM, Barry MJ, Fleming C, et al. Early detection of prostate cancer. Part I: Prior probability and effectiveness of tests. The American College of Physicians. *Ann Intern Med* 1997;126:394-406.
29. Brawer MK, Chetner MP, Beatie J, et al. Screening for prostatic carcinoma with prostate specific antigen. *J Urol* 1992;147:841-845.
30. Catalona WJ, Smith DS, Ratliff TL, et al. Measurement of prostate-specific antigen in serum as a screening test for prostate cancer. *New Engl J Med* 1991;324:1156-1161.
31. Carter HB, Pearson JD, Metter EJ, et al. Longitudinal evaluation of prostate-specific antigen levels in men with and without prostate disease. *JAMA* 1992;267:2215-2220.
32. Benson MC, Whang IS, Pantuck A, et al. Prostate specific antigen density: A means of distinguishing benign prostatic hypertrophy and prostate cancer. *J Urol* 1992;147:815-816.
33. Kane RA, Littrup PJ, Babaian R, et al. Prostate-specific antigen levels in 1695 men without evidence of prostate cancer. Findings of the American Cancer Society National Prostate Cancer Detection Project. *Cancer* 1992;69:1201-1207.
34. Oesterling JE. Prostate-specific antigen: Improving its ability to diagnose early prostate cancer. *JAMA* 1992;267:2236-2238.
35. Lee F, Littrup PJ, Loft-Christensen L, et al. Predicted prostate specific antigen results using transrectal ultrasound gland volume. Differentiation of benign prostatic hyperplasia and prostate cancer. *Cancer* 1992;70(Suppl 1):211-220.
36. Bangma CH, Grobbee DE, Schroder FH. Volume adjustment for intermediate prostate-specific antigen values in a screening population. *Eur J Cancer* 1995;31A:12-14.
37. Demura T, Watarai Y, Togashi M, et al. Measurement of prostate specific antigen and gamma-seminoprotein ratio: A new means of distinguishing benign prostatic hyperplasia and prostate cancer. *J Urol* 1993;150:1740-1745.

38. Lilja H. Significance of different molecular forms of serum PSA. The free, non-complexed form of PSA versus that complexed to alpha 1-antichymotrypsin. *Urol Clin North Am* 1993;20:681-686.
39. Bangma CH, Kranse R, Blijenberg BG, et al. The value of screening tests in the detection of prostate cancer. Part I: Results of a retrospective evaluation of 1726 men. *Urology* 1995;46:773-778.
40. Vashi AR, Wojno KJ, Henricks W, et al. Determination of the "reflex range" and appropriate cutpoints for percent free prostate-specific antigen in 413 men referred for prostatic evaluation using the AxSYM system. *Urology* 1997;49:19-27.
41. Cooner WH, Mosley BR, Rutherford CL Jr, et al. Prostate cancer detection in a clinical urological practice by ultrasonography, digital rectal examination and prostate specific antigen. *J Urol* 1990;143:1146-1154.
42. Johansson JE, Adami HO, Andersson SO, et al. Natural history of localised prostatic cancer. A population-based study in 223 untreated patients. *Lancet* 1989;1:799-803.
43. Miller GJ. Histopathology of prostate cancer: Prediction of malignant behavior and correlation with ultrasonography. *Urology* 1989;33(Suppl 6):18-26.
44. Kramer BS, Brown ML, Prorok PC, et al. Prostate cancer screening: What we know and what we need to know. *Ann Intern Med* 1993;119:914-923.
45. Chodak GW, Thisted RA, Gerber GS, et al. Results of conservative management of clinically localized prostate cancer. *New Engl J Med* 1994;330:242-248.
46. Fleming C, Wasson JH, Albertsen PC, et al. A decision analysis of alternative treatment strategies for clinically localized prostate cancer. Prostate Patient Outcomes Research Team. *JAMA* 1993;269:2650-2658.
47. Barry MJ, Fleming C, Coley CM, et al. Should Medicare provide reimbursement for prostate-specific antigen testing for early detection of prostate cancer? Part IV: Estimating the risks and benefits of an early detection program. *Urology* 1995;46:445-461.
48. Beck JR, Kattan MW, Miles BJ. A critique of the decision analysis for clinically localized prostate cancer. *J Urol* 1994;152:1894-1899.
49. Lu-Yao GL, Yao S. Population-based study of long-term survival in patients with clinically localised prostate cancer. *Lancet* 1997;349:906-910.
50. Prestigiacomo AF, Stamey TA. Physiological variation of serum prostate specific antigen in the 4.0 to 10.0 ng./ml. range in male volunteers. *J Urol* 1996;155:1977-1980.
51. Schroder FH, Bangma CH. The European randomized study of screening for prostate cancer (ERSPC). *Br J Urol* 1997;79(Suppl 1):68-71.
52. Auvinen A, Rietbergen JB, Denis LJ, et al. Prospective evaluation plan for randomized trials of prostate cancer screening. The International Prostate Cancer Screening Trial Evaluation Group. *J Med Screen* 1996;3:97-104.
53. Boucot KR, Weiss W. Is curable lung cancer detected by semiannual screening? *JAMA* 1973;224:1361-1365.
54. Lilienfeld A, Archer PG, Burnett CH, et al. An evaluation of radiologic and cytologic screening for the early detection of lung cancer: A cooperative pilot study of the American Cancer Society and the Veterans Administration. *Cancer Res* 1966;26:2083-2121.
55. Nash FA, Morgan JM, Tomkins JG. South London lung cancer study. *Br Med J* 1968;2:715-721.
56. Brett GZ. The value of lung cancer detection by six-monthly chest radiographs. *Thorax* 1968;23:414-420.
57. Brett GZ. Earlier diagnosis and survival in lung cancer. *Br Med J* 1969;4:260-262.
58. Dales LG, Friedman GD, Collen MF. Evaluating periodic multiphasic health check-ups: A controlled trial. *J Chronic Dis* 1979;32:385-404.

59. Friedman GD, Collen MF, Fireman BH. Multiphasic health checkup evaluation: A 16-year follow-up. *J Chronic Dis* 1986;39:453-463.
60. Cole P, Morrison AS. Basic issues in population screening for cancer. *J Natl Cancer Inst* 1980;64:1263-1272.
61. Prorok PC, Connor RJ. Screening for the early detection of cancer. *Cancer Invest* 1986;4:225-238.
62. Kubik A, Polak J. Lung cancer detection. Results of a randomized prospective study in Czechoslovakia. *Cancer* 1986;57:2427-2437.
63. Kubik A, Parkin DM, Khat M, et al. Lack of benefit from semi-annual screening for cancer of the lung: Follow-up report of a randomized controlled trial on a population of high-risk males in Czechoslovakia. *Int J Cancer* 1990;45:26-33.
64. Ebeling K, Nischan P. Screening for lung cancer—Results from a case-control study. *Int J Cancer* 1987;40:141-144.
65. Sobue T, Suzuki T, Naruke T. A case-control study for evaluating lung-cancer screening in Japan. Japanese Lung-Cancer Screening Group. *Int J Cancer* 1992; 50:230-237.
66. Fontana RS. Early detection of lung cancer: The Mayo Lung Project. In: Prorok PC, Miller AB, eds. *Screening for Cancer. I—General Principles on Evaluation of Screening for Cancer and Screening for Lung, Bladder and Oral Cancer*, vol. 78. Geneva: UICC Technical Report Series; 1984:107-122.
67. Fontana RS. Screening for lung cancer. In: Miller AB, ed. *Screening for Cancer*. New York: Academic Press; 1985:377-395.
68. Fontana RS. Screening for lung cancer: Recent experience in the United States. In: Hansen HH, ed. *Lung Cancer: Basic and Clinical Aspects*. Boston: Martinus Nijhoff Publishers; 1986:91-111.
69. Levin ML, Tockman MS, Frost JK, et al. Lung cancer mortality in males screened by chest X-ray and cytologic sputum examination: A preliminary report. *Recent Results Cancer Res* 1982;82:138-146.
70. Stitik FP, Tockman MS. Radiographic screening in the early detection of lung cancer. *Radiol Clin North Am* 1978;16:347-366.
71. Stitik FP, Tockman MS, Khouri NF. Chest radiology. In: Miller AB, ed. *Screening for Cancer*. New York: Academic Press; 1985:163-191.
72. Tockman MS, Frost JK, Stitik FP, et al. Screening and detection of lung cancer. In: Aisner J, ed. *Lung Cancer*. New York: Churchill Livingstone; 1985:25-39.
73. Melamed MR, Flehinger BJ, Zaman MB, et al. Detection of true pathologic stage I lung cancer in a screening program and the effect on survival. *Cancer* 1981; 47(Suppl 5):1182-1187.
74. Melamed MR, Flehinger BJ, Zaman MB, et al. Screening for early lung cancer. Results of the Memorial Sloan-Kettering study in New York. *Chest* 1984;44-53.
75. Flehinger BJ, Kimmel M. The natural history of lung cancer in a periodically screened population. *Biometrics* 1987;43:127-144.
76. Flehinger BJ, Kimmel M, Melamed MR. Natural history of adenocarcinoma-large cell carcinoma of the lung: Conclusions from screening programs in New York and Baltimore. *J Natl Cancer Inst* 1988;80:337-344.
77. Shukla R, Deddens JA, Buncher CR. Survival benefits of x-ray screening for lung cancer after bias adjustments. *Computers and Mathematics with Applications* 1989; 18:937-948.
78. Frühmorgen P, Demling L. Early detection of colorectal cancer with a modified guaiac test—A screening examination of 6,000 humans. In: Winawer S, Schottenfeld D, Sherlock P, eds. *Progress in Cancer Research and Therapy*, vol. 13. *Colorectal Cancer: Prevention, Epidemiology, and Screening*. New York: Raven Press; 1980:311-315.

79. Winawer SJ, Andrews M, Flehinger B, et al. Progress report on controlled trial of fecal occult blood testing for the detection of colorectal neoplasia. *Cancer* 1980; 45:2959-2964.
80. Winawer SJ, Fath RB Jr, Schottenfeld D, et al. Screening for colorectal cancer. In: Miller AB, ed. *Screening for Cancer*. New York: Academic Press; 1985:347-366.
81. Selby JV, Friedman GD, Quesenberry CP Jr, et al. Effect of fecal occult blood testing on mortality from colorectal cancer. A case-control study. *Ann Intern Med* 1993; 118:1-6.
82. Lazovich D, Weiss NS, Stevens NG, et al. A case-control study to evaluate efficacy of screening for faecal occult blood. *J Med Screen* 1995;2:84-89.
83. Winawer SJ, Flehinger BJ, Schottenfeld D, et al. Screening for colorectal cancer with fecal occult blood testing and sigmoidoscopy. *J Natl Cancer Inst* 1993;85:1311-1318.
84. Mandel JS, Bond JH, Church TR, et al. Reducing mortality from colorectal cancer by screening for fecal occult blood. Minnesota Colon Cancer Control Study. *N Eng J Med* 1993;328:1365-1371.
85. Hardcastle JD, Chamberlain JO, Robinson MH, et al. Randomised controlled trial of faecal-occult-blood screening for colorectal cancer. *Lancet* 1996;348:1472-1477.
86. Kewenter J, Bjork S, Haglund E, et al. Screening and rescreening for colorectal cancer. A controlled trial of fecal occult blood testing in 27,700 subjects. *Cancer* 1988;62:645-651.
87. Kronborg O, Fenger C, Snergaard O, et al. Initial mass screening for colorectal cancer with fecal occult blood test. *Scand J Gastroenterol* 1987;22:677-686.
88. Kronborg O, Fenger C, Olsen J, et al. Repeated screening for colorectal cancer with fecal occult blood test. A prospective randomized study at Funen, Denmark. *Scand J Gastroenterol* 1989;24:599-606.
89. Kronborg O, Fenger C, Olsen J, et al. Randomized study of screening for colorectal cancer with faecal-occult-blood test. *Lancet* 1996;348:1467-1471.
90. Guidelines for the cancer-related checkup: Recommendations and rationale. *CA Cancer J Clin* 1980;30:194-240.
91. Morrison AS. *Screening in Chronic Disease*. New York: Oxford University Press; 1985.
92. Selby JV, Friedman GD, Collen MF. Sigmoidoscopy and mortality from colorectal cancer: The Kaiser Permanente Multiphasic Evaluation Study. *J Clin Epidemiol* 1988;41:427-434.
93. Herrinton LJ, Selby JV, Friedman GD, et al. Case-control study of digital-rectal screening in relation to mortality from cancer of the distal rectum. *Am J Epidemiol* 1995;142:961-964.
94. Gilbertsen VA. Proctosigmoidoscopy and polypectomy in reducing the incidence of rectal cancer. *Cancer* 1974;34(Suppl 3):936-939.
95. Gilbertsen VA, Nelms JM. The prevention of invasive cancer of the rectum. *Cancer* 1978;41:1137-1139.
96. Hertz RE. Value of periodic examinations in detecting cancer of the colon and rectum. *Postgrad Med* 1960;27:290-294.
97. Crespi M, Weissman GS, Gilbertsen VA, et al. The role of proctosigmoidoscopy in screening for colorectal neoplasia. *CA Cancer J Clin* 1984;34:158-166.
98. Winawer SJ, Miller C, Lightdale C, et al. Patient response to sigmoidoscopy. A randomized, controlled trial of rigid and flexible sigmoidoscopy. *Cancer* 1987;60: 1905-1908.
99. Winawer SJ, Zauber AG, Ho MN, et al. Prevention of colorectal cancer by colonoscopic polypectomy. The National Polyp Study Workgroup. *New Eng J Med* 1993;329: 1977-1981.
100. Chamberlain J, Day NE, Hakama M, et al. UICC workshop of the Project on Evaluation of Screening Programmes for Gastrointestinal Cancer. *Int J Cancer* 1986; 37:329-334.

101. Neugut AI, Pita S. Role of sigmoidoscopy in screening for colorectal cancer: A critical review. *Gastroenterology* 1988;95:492-499.
102. Eddy DM, Nugent FW, Eddy JF, et al. Screening for colorectal cancer in a high-risk population. Results of a mathematical model. *Gastroenterology* 1987;92:682-692.
103. Eddy DM. Screening for colorectal cancer. *Ann Intern Med* 1990;113:373-384.
104. Newcomb PA, Norfleet RG, Storer BE, et al. Screening sigmoidoscopy and colorectal cancer mortality. *J Natl Cancer Inst* 1992;84:1572-1575.
105. Selby JV, Friedman GD, Quesenberry CP Jr, et al. A case-control study of screening sigmoidoscopy and mortality from colorectal cancer. *N Engl J Med* 1992;326:653-657.
106. Granberg S, Wikland M, Jansson I. Macroscopic characterization of ovarian tumors and the relation to the histological diagnosis: Criteria to be used for ultrasound evaluation. *Gynecol Oncol* 1989;35:139-144.
107. Jacobs I, Bast RC Jr. The CA 125 tumour-associated antigen: A review of the literature. *Hum Reprod* 1989;4:1-12.
108. Einhorn N, Sjoval K, Knapp RC, et al. Prospective evaluation of serum CA 125 levels for early detection of ovarian cancer. *Obstet Gynecol* 1992;80:14-18.
109. Helzlsouer KJ, Bush TL, Alberg AJ, et al. Prospective study of serum CA-125 levels as markers of ovarian cancer. *JAMA* 1993;269:1123-1126.
110. Zurawski VR Jr, Orjaseter H, Andersen A, et al. Elevated serum CA 125 levels prior to diagnosis of ovarian neoplasia: Relevance for early detection of ovarian cancer. *Int J Cancer* 1988;42:677-680.
111. Higgins RV, van Nagell JR Jr, Woods CH, et al. Interobserver variation in ovarian measurements using transvaginal sonography. *Gynecol Oncol* 1990;39:69-71.
112. van Nagell JR Jr, Higgins RV, Donaldson ES, et al. Transvaginal sonography as a screening method for ovarian cancer. A report of the first 1000 cases screened. *Cancer* 1990;65:573-577.
113. Andolf E, Jorgensen C, Astedt B. Ultrasound examination for detection of ovarian carcinoma in risk groups. *Obstet Gynecol* 1990;75:106-109.
114. Campbell S, Bhan V, Royston P, et al. Transabdominal ultrasound screening for early ovarian cancer. *BMJ* 1989;299:1363-1367.
115. DePriest PD, van Nagell JR Jr, Gallion HH, et al. Ovarian cancer screening in asymptomatic postmenopausal women. *Gynecol Oncol* 1993;51:205-209.
116. Bourne TH, Campbell S, Reynolds KM, et al. Screening for early familial ovarian cancer with transvaginal ultrasonography and colour blood flow imaging. *BMJ* 1993;306:1025-1029.
117. Freedman LS, Green SB. Statistical designs for investigating several interventions in the same study: Methods for cancer prevention trials. *J Natl Cancer Inst* 1990;82:910-914.
118. Etzioni RD, Connor RJ, Prorok PC, et al. Design and analysis of cancer screening trials. *Stat Methods Med Res* 1995;4:3-17.
119. Shapiro S, Venet W, Strax P, et al. *Periodic Screening for Breast Cancer. The Health Insurance Plan Project and Its Sequelae, 1963-1986*. Baltimore: The Johns Hopkins University Press; 1988.
120. Tabar L, Fagerberg G, Duffy SW, et al. Update of the Swedish two-county program of mammographic screening for breast cancer. *Radiologic Clin North Am* 1992;30:187-210.
121. Simpson NK, Johnson CC, Ogden SL, et al. Recruitment strategies in the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial: The first six years. *Control Clin Trials* 2000;21:356S-378S.
122. Hayes RB, Reding D, Kopp W, et al. Etiologic and early marker studies in the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial. *Control Clin Trials* 2000;21:349S-355S.

123. Weissfeld JL, Fagerstrom RM, O'Brien B. Quality control of cancer screening examination procedures in the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial. *Control Clin Trials* 2000;21:390S–399S.
124. Miller AB, Yurgalevitch S, Weissfeld JL. Death review process in the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial. *Control Clin Trials* 2000;21:400S–406S.
125. Taylor WF, Fontana RS. Biometric design of the Mayo Lung Project for early detection and localization of bronchogenic carcinoma. *Cancer* 1972;30:1344–1347.
126. Dunnett CW, Gent M. An alternative to the use of two-sided tests in clinical trials. *Stat Med* 1996;15:1729–1738.
127. 1987 Annual Cancer Statistics Review, Including Cancer Trends: 1950–1985. NIH Publication No. 88-2789. Bethesda, MD: National Cancer Institute, Division of Cancer Prevention and Control; 1988.
128. Potosky AL, Annett DQ, Coyle L. Guide to using CAN*TROL, version 2.0. Bethesda, MD: National Cancer Institute, Division of Cancer Prevention and Control; 1995.
129. Brown ML, Potosky AL, Thompson GB, et al. The knowledge and use of screening tests for colorectal and prostate cancer: Data from the 1987 National Health Interview survey. *Prev Med* 1990;19:562–574.
130. Polednak AP. Knowledge of colorectal cancer and use of screening tests in persons 40–74 years of age. *Prev Med* 1990;19:213–226.
131. Hasson MA, Fagerstrom RM, Kahane DC, et al. Design and evolution of the data management systems in the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial. *Control Clin Trials* 2000;21:329S–348S.
132. Day NE. Estimating the sensitivity of a screening test. *J Epidemiol Community Health* 1985;39:364–366.
133. Day NE, Walter SD. Simplified models of screening for chronic disease: Estimation procedures from mass screening programs. *Biometrics* 1984;40:1–14.
134. Walter S, Day NE. Estimation of the duration of a pre-clinical state using screening data. *Am J Epidemiol* 1983;118:865–886.
135. Kafadar K, Prorok PC. A data-analytic approach for estimating lead time and screening benefit based on survival curves in randomized cancer screening trials. *Stat Med* 1994;13:569–586.
136. Kafadar K, Prorok PC. Estimating the difference in location parameters of two survival curves, with applications to cancer screening. *J Statist Planning Inference* 1997;57:165–179.
137. Day NE, Williams DRR, Khaw KT. Breast cancer screening programmes: The development of a monitoring and evaluation system. *Br J Cancer* 1989;59:954–958.
138. Lan KKG, DeMets DL. Group sequential procedures: Calendar versus information time. *Stat Med* 1989;8:1191–1198.
139. Lin DY, Yao Q, Ying Z. A general theory on stochastic curtailment for censored survival data. *J Am Stat Assoc* 1999;94:510–521.