Effect of Screening on Ovarian Cancer Mortality
The Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Randomized Controlled Trial

Saundra S. Buys, MD
Edward Partridge, MD
Amanda Black, PhD, MPH
Christine C. Johnson, PhD, MPH
Lois Lamerato, PhD
Claudine Isaacs, MD
Douglas J. Reding, MD, MPH
Robert T. Greenlee, PhD, MPH
Lance A. Yokochi, MD, PhD
Bruce Kessel, MD
E. David Crawford, MD
Timothy R. Church, PhD, MS
Gerald L. Andriole, MD
Joel L. Weissfeld, MD, MPH
Mona N. Fouad, MD
David Chia, PhD
Barbara O'Brien, MPH
Lawrence R. Ragard, MD
Jonathan D. Clapp
Joshua M. Rathmell
Thomas L. Riley, BS
Patricia Hartge, ScD, MA
Paul F. Pinsky, PhD
Claire S. Zhu, PhD
Grant Izmirlian, PhD
Barnett S. Kramer, MD
Anthony B. Miller, MD
Jian-Lun Xu, PhD
Philip C. Prorok, PhD
John K. Gohagan, PhD
Christine D. Berg, MD
for the PLCO Project Team

Context Screening for ovarian cancer with cancer antigen 125 (CA-125) and transvaginal ultrasound has an unknown effect on mortality.

Objective To evaluate the effect of screening for ovarian cancer on mortality in the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial.

Design, Setting, and Participants Randomized controlled trial of 78 216 women aged 55 to 74 years assigned to undergo either annual screening (n=39 105) or usual care (n=39 111) at 10 screening centers across the United States between November 1993 and July 2001.

Intervention The intervention group was offered annual screening with CA-125 for 6 years and transvaginal ultrasound for 4 years. Participants and their health care practitioners received the screening test results and managed evaluation of abnormal results. The usual care group was not offered annual screening with CA-125 for 6 years or transvaginal ultrasound but received their usual medical care. Participants were followed up for a maximum of 13 years (median [range], 12.4 years [10.9-13.0 years]) for cancer diagnoses and death until February 28, 2010.

Main Outcome Measures Mortality from ovarian cancer, including primary peritoneal and fallopian tube cancers. Secondary outcomes included ovarian cancer incidence and complications associated with screening examinations and diagnostic procedures.

Results Ovarian cancer was diagnosed in 212 women (5.7 per 10 000 person-years) in the intervention group and 176 (4.7 per 10 000 person-years) in the usual care group (rate ratio [RR], 1.21; 95% confidence interval [CI], 0.99-1.48). There were 118 deaths caused by ovarian cancer (3.1 per 10 000 person-years) in the intervention group and 100 deaths (2.6 per 10 000 person-years) in the usual care group (mortality RR, 1.18; 95% CI, 0.82-1.71). Of 3285 women with false-positive results, 1080 underwent surgical follow-up; of whom, 163 women experienced at least 1 serious complication (15%). There were 2924 deaths due to other causes (excluding ovarian, colorectal, and lung cancer) (76.6 per 10 000 person-years) in the intervention group and 2914 deaths (76.2 per 10 000 person-years) in the usual care group (RR, 1.01; 95% CI, 0.96-1.06).

Conclusions Among women in the general US population, simultaneous screening with CA-125 and transvaginal ultrasound compared with usual care did not reduce ovarian cancer mortality. Diagnostic evaluation following a false-positive screening test result was associated with complications.

Trial Registration clinicaltrials.gov Identifier: NCT00002540

I N THE UNITED STATES, OVARIAN cancer is among the 5 leading causes of cancer death in women.[1] The high case-fatality ratio of ovarian cancer may be attributed in part to its vague and nonspecific symptoms, which usually appear when the disease has reached an advanced stage. Ovarian cancer confined to the ovary has a 5-year survival of 92%. However, most women with ovarian cancer are diagnosed with advanced stage disease, which has a 5-year survival of

Author Affiliations are listed at the end of this article. A complete list of the PLCO Project Team is listed in the eAppendix at http://www.jama.com.

Corresponding Author: Christine D. Berg, MD, Executive Plaza N, 6130 Executive Blvd, MSC 7346, Room 3112, Rockville, MD 20852 (bergc@mail.nih.gov).

only 30%.[1] The recognition that early detection of ovarian cancer may have the potential to improve prognosis prompted the development of randomized controlled trials, including the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial, to evaluate the efficacy of transvaginal ultrasound and serum cancer antigen 125 (CA-125) as screening tools to reduce ovarian cancer mortality.[2-4]

The PLCO trial's independent data and safety monitoring board, which periodically reviews the cumulative trial data, recently concluded that the predetermined end point had been reached for the ovarian cancer component of the trial and recommended that the ovarian cancer–specific mortality results be reported.

METHODS

The design and methods of the PLCO trial have been reported.[5] In brief, the PLCO trial was designed to determine the effect of specific cancer screening tests on cause-specific mortality. Enrollment began in November 1993 and concluded in July 2001. Planned follow-up was for up to 13 years from randomization. Women were randomized to either the intervention group or the usual care group. Women in the intervention group were offered annual screening with transvaginal ultrasound and CA-125, whereas those randomized to the usual care group were offered no interventions and only received their usual medical care. Women were considered eligible if they were aged 55 to 74 years and had no previous diagnosis of lung, colorectal, or ovarian cancer. The 2 initial exclusion criteria of previous oophorectomy and current tamoxifen use were dropped in 1996 and 1999, respectively.[6] Women who had undergone previous bilateral oophorectomy were screened for lung and colorectal cancer but not for ovarian cancer. These women are not included in this analysis.

Participants were recruited at 10 screening centers across the United States. Recruitment targeted individuals from the general population resid-ing within the catchment area of each of the screening centers. The various localities of the screening centers are reflected in the ethnic diversity of the PLCO participants. The recruitment strategies have been described.[7,8]

Eligible individuals were randomized with a scheme that used blocks of random permutations of varying lengths and was stratified by screening center, age, and sex. Random allocation was achieved using compiled software and encrypted files loaded to the screening center computers by the PLCO coordinating center. Due to the invasive nature of the screening procedures, participants were not blinded to their randomization allocation. Each institution obtained annual approval from its institutional review board to perform the study, and all participants provided written informed consent. At study entry, participants completed a self-administered baseline questionnaire, which included demographics (such as race/ethnicity classified as white, black, Asian/Pacific Islander, American Indian/Alaskan Native, or Hispanic origin), general risk factors, and screening and medical histories.

Screening Examinations

Women in the intervention group underwent screening with a CA-125 blood test and transvaginal ultrasound at baseline, an annual transvaginal ultrasound for an additional 3 years, and an annual CA-125 for an additional 5 years. The original protocol specified that women undergo a CA-125 annually for only 4 years. However, women were eligible by 1999 to have CA-125 screening for the fifth and sixth study years. Therefore, depending on when women were randomized, they may have had the opportunity to attend 4, 5, or 6 screening rounds. Bimanual examination of the ovaries was originally part of the screening procedures but was discontinued in December 1998 because no cancers were detected solely by ovarian palpation; in the usual care group, a high proportion of women underwent bimanual examination with ovarian palpation. The CA-125 screen-ing assays were analyzed centrally at the Immunogenetics Laboratory (University of California, Los Angeles). The CA-125 assay results that were 35 U/mL or greater were classified as abnormal.[6] In addition, blood and tissue samples were collected and stored in a central biorepository.[9]

Transvaginal ultrasound was conducted by trained examiners using a 5- to 7.5-MHz transvaginal probe.[6] At least 5 minutes were spent looking for each ovary before categorizing the ovaries as being nonvisualized. The following transvaginal ultrasound results were classified as abnormal: (1) ovarian volume greater than 10 cm^3, (2) cyst volume greater than 10 cm^3, (3) any solid area or papillary projection extending into the cavity of a cystic ovarian tumor of any size; and (4) any mixed (solid and cystic) component within a cystic ovarian tumor. Quality assurance included duplicate screening examinations on a sample of participants, independent observation of the examination, independent review of transvaginal ultrasound films, or both.[10]

Participants and their physicians were notified in writing about suspicious abnormalities. In accordance with standard US medical practice, it was the responsibility of the participant's primary care physician to manage the diagnostic process to assess abnormalities. The PLCO coordinating center obtained the medical records for all diagnostic and therapeutic follow-up procedures.

For the purposes of this study, cancers detected by screening were defined as those diagnosed as a result of investigations initiated after a screening test with a positive result and without a lapse in the diagnostic evaluation exceeding 9 months. A false-positive result was defined as a positive screening examination result that did not result in cancers detected by screening. Interval cancers were defined as cancers not detected by screening and diagnosed within 12 months of the woman's last expected screening examination. Other cancers not detected by screening included those diagnosed among women who never un-

derwent a screening examination (non-compliant) and those diagnosed after the screening phase was completed.

Ascertainment of Study End Points

The primary study end point was ovarian cancer–specific mortality. Secondary end points included ovarian cancer incidence, cancer stage, survival, potential harms of screening, and all-cause mortality. Ovarian, primary peritoneal, and fallopian tube cancers were considered malignant ovarian neoplasms for this report (*International Classification of Diseases for Oncology, Second Revision,* codes C569, C481, C482, and C570). Tumors of low malignant potential (borderline tumors) were not included in the definition of malignant ovarian neoplasms, but are included as false-positive results (n=21 in the intervention group; n=6 in the usual care group).

All incident cancers (both PLCO trial and other cancers) and deaths were ascertained primarily by an annual study update questionnaire mailed to participants. Population-based cancer registries also were used when possible. Additionally, to obtain more complete mortality data, annual study update follow-up was supplemented by periodic linkage to the National Death Index. All ovarian cancers and deaths from confirmed ovarian cancer known as of February 28, 2010, are included in this analysis. Medical records pertaining to diagnosed cancers were obtained by the PLCO screening centers; data on the stage, histology, and grade of PLCO cancers were abstracted by certified tumor registrars. In addition, treatment information during the first year post-diagnosis was abstracted.

The underlying cause of death was determined in a manner that was uniform and unbiased and was based on the death certificate and relevant medical records; determination of the underlying cause of death also has been described in detail elsewhere.[11] Briefly, deaths potentially related to a PLCO trial cancer and those of unknown or uncertain cause were reviewed by at least 1 individual from a panel with appropriate expertise

(epidemiology, surgery, medicine, radiation oncology); panel members were not otherwise affiliated with the trial and reviewers were blinded to the randomization group of the deceased participant.

Early on during the death review process, each death had 2 independent reviews and discrepancies were resolved by consensus after a third reviewer had examined the records. After 2 years, the process was streamlined. As a result, a primary reviewer considered the record without having access to the death certificate. If he or she recorded an underlying cause of death different from that of the death certificate, a second reviewer independently reviewed the record together with the death certificate. Disagreement between the reviewers resulted in

Figure 1. Flow of Patients Through the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial
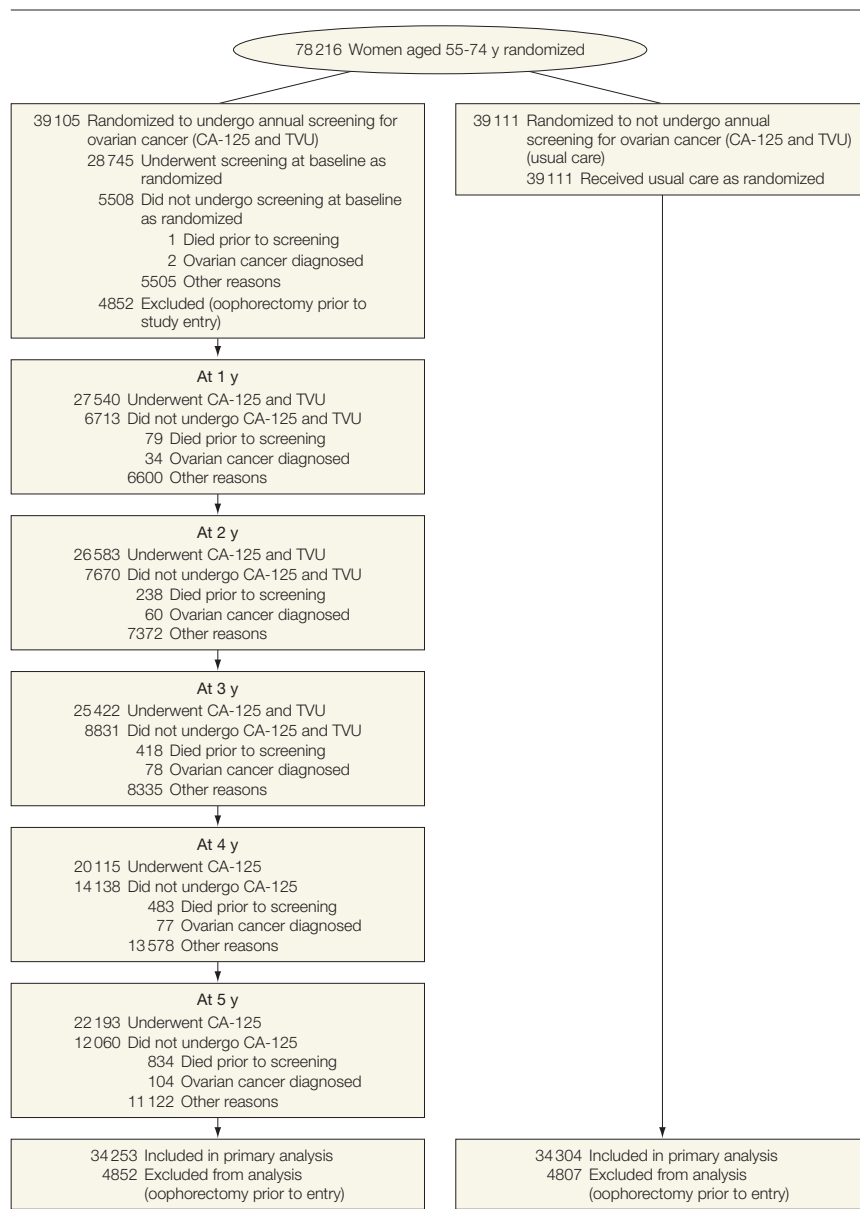
CA-125 indicates cancer antigen 125; TVU, transvaginal ultrasound.

Table 1. Participant Characteristics in the PLCO Cancer Trial

Characteristic	No. (%) of Women[a]	
	Intervention Group (n = 34 253)	Usual Care Group (n = 34 304)
Age, y		
55-59	11 698 (34.2)	11 728 (34.2)
60-64	10 405 (30.4)	10 398 (30.3)
65-69	7497 (21.9)	7506 (21.9)
70-74	4653 (13.6)	4672 (13.6)
Race/ethnicity		
White (non-Hispanic)	29 502 (88.6)	29 292 (88.4)
Black (non-Hispanic)	1896 (5.7)	1897 (5.7)
Hispanic origin	512 (1.5)	512 (1.5)
Asian/Pacific Islander	1138 (3.4)	1190 (3.6)
American Indian/Alaskan Native	250 (0.8)	253 (0.8)
Education		
<High school	2175 (6.6)	2142 (6.5)
High school graduate	13 308 (40.0)	13 308 (40.5)
Some college	7688 (23.1)	7483 (22.7)
College graduate	5156 (15.5)	4999 (15.1)
Postgraduate study	4926 (14.8)	5020 (15.2)
Prior hysterectomy	9083 (27.3)	8979 (27.2)
Prior oral contraceptive use	17 822 (53.6)	17 883 (54.1)
Prior hormone therapy use	20 992 (63.4)	20 744 (63.0)
Nulliparous	3092 (9.3)	3040 (9.2)
Personal history of breast cancer	1200 (3.6)	1186 (3.6)
Family history of breast or ovarian cancer	5592 (17.6)	5437 (17.3)

Abbreviation: PLCO, Prostate, Lung, Colorectal and Ovarian.
[a] For all variables except age, percentages exclude missing or unknown values. On average, approximately 3% of responses were missing for the variables presented.

another independent review, which was subsequently resolved by a meeting or teleconference.

Screening-Related Harms

Complications directly associated with screening by transvaginal ultrasound and CA-125 (eg, bleeding, fainting, nausea, bruising) were recorded at the time of the examinations. More severe complications such as infection, cardiovascular events, and bowel injury associated with cancer diagnoses and false-positive test results were assessed from the medical records. Oophorectomy rates in each study group were assessed through the use of a supplemental questionnaire, which was sent in 2006 or 2007 to participants after all PLCO screening had been completed. Similar to the baseline questionnaire, the supplemental questionnaire inquired about general risk factors and medical history, including oophorectomy. Oophorectomy rates among those women with at least 1 ovary

at baseline were then calculated in each group as the number of oophorectomies reported on the supplemental questionnaire divided by the total person-time from enrollment to completion of the supplemental questionnaire.

Compliance and Contamination

Women were expected to undergo their annual screening if they had not been diagnosed with ovarian cancer, had not died, and had not undergone oophorectomy prior to the screening date. The rate of compliance with screening was calculated as the number of women actually screened divided by the number expected for the test. The use of screening outside of the trial protocol (contamination) was measured in 2 ways. Baseline contamination (defined as testing with either CA-125 or transvaginal ultrasound at least twice during the 3 years prior to trial entry) was determined from the baseline questionnaire. After trial enrollment, annual or biennial surveys were

conducted in a 1% random sample of women in the usual care group to estimate the subsequent level of contamination. Contamination during the trial was defined as having reported a screening by CA-125 or transvaginal ultrasound in the prior year.

Statistical Methods

The primary analysis was a comparison of ovarian cancer mortality rates between the 2 study groups by intention to perform screening. Secondary aims included comparison of ovarian cancer incidence, cancer stage, survival, potential harms of screening, and all-cause mortality between the 2 groups.

The trial had 88% power to detect a 35% reduction in ovarian cancer mortality at a 1-sided α level of .05.[5] These calculations included adjustment for expected contamination in the usual care group and noncompliance in the intervention group. The projected compliance rate in the screening group was greater than 90% for CA-125 and greater than 85% for transvaginal ultrasound; the projected contamination in the usual care group was less than 10% for both screening methods.

An interim analysis plan was used to monitor the primary end point for efficacy and futility. The plan used a weighted log-rank statistic with the weights increasing in proportion to the pooled ovarian cancer mortality. The weighted statistic was chosen because of the presumed delay in effect of screening on ovarian cancer mortality. The efficacy boundary was constructed via the Lan-DeMets approach using an O'Brien-Fleming spending function, and a nonbinding futility boundary was constructed via stochastic curtailment.[12]

After taking into consideration the fact that the futility boundary for the monitoring statistic had been crossed, the PLCO trial's data and safety monitoring board concluded in November 2009 that the primary aim of the ovarian component had been achieved and recommended that the results be reported. Further details about the interim analyses are provided in the

eTable at http://www.jama.com. Statistical analyses were performed using SAS software version 9 (SAS Institute Inc, Cary, North Carolina) and R statistical software.[13]

Event rates were defined as the number of events (diagnoses of ovarian cancer or deaths) divided by the person-time at risk for the event. Person-time was measured from randomization to the earliest date of diagnosis, death, or censoring for incidence rates and to the earliest date of the death or censoring for mortality rates. The censoring date was the earliest of the date of last contact, the 13-year anniversary of randomization, or February 28, 2010.

Rate ratios (RRs) were derived as the unadjusted ratio of event rates. The 95% confidence interval (CI) for the ovarian mortality RR was derived based on the sequential design according to the weighted method used to monitor the trial, which allows for a varying RR.[14] The 95% CIs for the other RRs and the nominal 95% CI for the ovarian mortality RR were calculated using a single test normal critical value, which assumed a Poisson distribution for the number of events and a normal distribution for the logarithm of the ratio according to asymptotic methods.[15] Additional details and references are provided in the eMethods

and eReferences at http://www.jama.com)

RESULTS

The current analysis excludes women with a pretrial history of bilateral oophorectomy. Women with a history of bilateral oophorectomy were included in the trial after a protocol change in 1996 permitted enrollment of such women (12% of the participants). From the original enrollment allocation of 39 105 women in the intervention group and 39 111 women in the usual care group, 4852 and 4807, respectively, were removed, leaving a total of 34 253 in the intervention group and 34 304 in the usual care group who were included in this analysis (**FIGURE 1**). The baseline characteristics of the participants in the 2 groups were well balanced (**TABLE 1**). Because randomization was performed between 1993 and 2001, some participants did not reach 13 years of follow-up by the cutoff date (February 28, 2010). However, all participants, regardless of follow-up time, were included in the analysis. The median follow-up was 12.4 years (25th-75th percentile, 10.9-13.0 years).

Contamination and Compliance

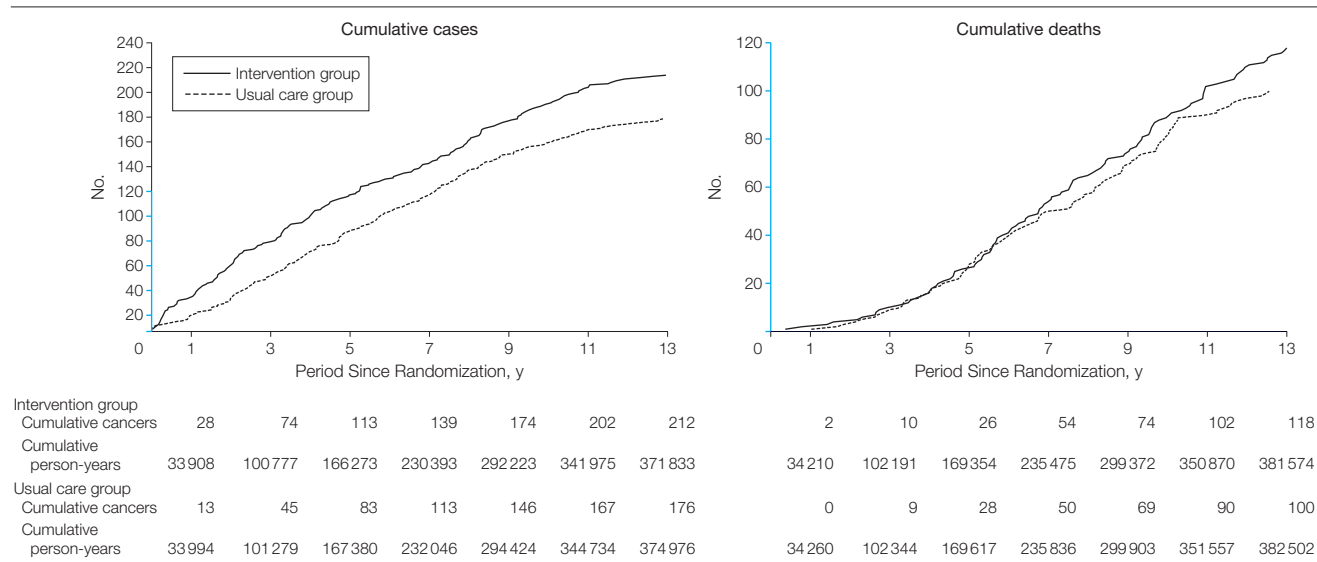Screening outside the trial protocol (contamination) was minimal. In each study

group, baseline contamination rates (ie, use of screening in the 3 years prior to trial entry) were 2.8% for CA-125 and 9.4% for transvaginal ultrasound. During the screening phase of the trial, contamination in the usual care group ranged from 2.3% to 3.2% per year for CA-125 and from 2.7% to 4.6% for transvaginal ultrasound.

Compliance with screening in the intervention group was 85% for CA-125 and 84% for transvaginal ultrasound at baseline, with compliance declining modestly to 79% and 78%, respectively, by the fourth screening. At the fifth and sixth screenings (at which transvaginal ultrasound was not administered), the proportions screened by CA-125 were 75% and 73%, respectively. Positive screening rates ranged from 1.4% to 1.8% for CA-125 during the 6 screening rounds and from 2.9% to 4.6% for transvaginal ultrasound during the 4 screening rounds.

Ovarian Cancer Incidence

Through the follow-up period, 212 ovarian cancer cases (5.7 per 10 000 person-years) were diagnosed in the intervention group and 176 cases (4.7 per 10 000 person-years) in the usual care group (RR, 1.21; 95% CI, 0.99-1.48). The accumulation of ovarian cancer cases over

Figure 2. Ovarian Cancer Cumulative Cases and Deaths

Intervention group														
Cumulative cancers	28	74	113	139	174	202	212	2	10	26	54	74	102	118
Cumulative person-years	33 908	100 777	166 273	230 393	292 223	341 975	371 833	34 210	102 191	169 354	235 475	299 372	350 870	381 574
Usual care group														
Cumulative cancers	13	45	83	113	146	167	176	0	9	28	50	69	90	100
Cumulative person-years	33 994	101 279	167 380	232 046	294 424	344 734	374 976	34 260	102 344	169 617	235 836	299 903	351 557	382 502

Y-axis shown in blue indicates range of 0 to 120 cumulative events.

Table 2. Detection and Tumor Characteristics of Ovarian Cancers

	No. (%) of Women[a]	
	Intervention Group (n = 212)	Usual Care Group (n = 176)
Cancer site		
Primary invasive neoplasm of ovary	169 (80)	151 (86)
Primary peritoneal cancer	29 (14)	18 (10)
Primary invasive neoplasm of fallopian tube	14 (7)	7 (4)
Histology		
Serous	116 (55)	103 (59)
Mucinous	5 (2)	3 (2)
Endometrioid	19 (9)	8 (5)
Clear cell	6 (3)	6 (3)
Not specified and other	66 (31)	55 (31)
Unknown	0	1 (1)
Grade		
1	12 (6)	7 (4)
2	29 (14)	21 (12)
3	132 (62)	109 (62)
Could not be assessed	7 (3)	6 (3)
Unknown	32 (15)	33 (19)
Detection		
Screening at baseline	20 (9)	NA
Screening at 1-5 y	53 (25)	NA
Interval cases[b]	37 (17)	NA
After screening phase (compliant)	78 (37)	NA
Never screened (noncompliant)		
During screening phase	16 (8)	NA
After screening phase	8 (4)	NA

Abbreviation: NA, data not applicable.
[a]Subsections do not equal 100% because percentages were rounded to the nearest whole number.
[b]Defined in the "Methods" section.

time since randomization appears in **FIGURE 2**. The excess of cases in the intervention group compared with the usual care group increased during the first 2 years and then remained approximately constant after year 3.

Ovarian Cancer Mortality

Examination of the ovarian cancer deaths in the final interim analysis demonstrated that the boundary for futility had been reached. There were 118 deaths caused by ovarian cancer (3.1 per 10 000 person-years) in the intervention group and 100 deaths (2.6 per 10 000 person-years) in the usual care group (mortality RR, 1.18; 95% CI, 0.91-1.54 [unadjusted] and mortality RR, 1.18; 95% CI, 0.82-1.71 [sequentially adjusted]). The cumulative numbers of ovarian cancer deaths over time appears in Figure 2.

Tumor Characteristics

TABLE 2 presents characteristics of the ovarian cancers by study group. Primary peritoneal and fallopian tube cancers, which were considered ovarian cases for this analysis, accounted for 20% of all ovarian cancers diagnosed in the intervention group (n=43) and 14% in the usual care group (n=25). Eliminating these cases did not alter the statistical significance of any of the findings (mortality RR, 1.04; 95% CI, 0.79-1.38). The histological subtypes were similar across groups and 55% were serous cystadenocarcinomas in the intervention group and 59% in the usual care group. The majority of cancers in each study group were high grade.

Table 2 also shows the cases in the intervention group by method and time of diagnosis. Of the 212 intervention group cases, 126 were diagnosed during the screening phase of the trial (59%), of which 58% were screen-detected cases, 29% were interval cases, and 13% were detected in those without prior screening (noncompliant).

Overall, the stage distributions were similar by study group with stage III and IV cancers comprising the majority of cases in both the intervention group (77%) and usual care group (78%) (**TABLE 3**). The absolute number of stage IV cancers was slightly higher in the usual care group (n=54) than in the intervention group (n=43), although this difference was not statistically significant. There was no significant association between study period and stage distribution; 76% of intervention group cases diagnosed during the screening phase of the study (years 0-5)

were stage III or IV compared with 79% of those diagnosed during the post-screening phase. Among actual screen-detected cases (n=73), 69% were stage III or IV (data not shown).

Treatments for ovarian cancer were similar across study groups overall and within each stage (**TABLE 4**). In the intervention group, 81% underwent surgery and received systemic therapy compared with 80% in the usual care group. Among women with stage III cancer, 93% in the intervention group and 87% in the usual care group underwent surgery and received systemic therapy.

Ovarian Cancer Survival

The lead-time bias associated with early detection of ovarian cancer is illustrated in the eFigure at http://www.jama.com. The eFigure presents ovarian cancer survival from the date of diagnosis and the date of randomization, which eliminates most of the lead-time effect. A small difference in survival was seen in the date from diagnosis (P=.18) but not in the date from randomization (P=.67) (eFigure). The difference in survival between the intervention and usual care groups was not statistically significant.

Screening-Related Harms

Minor complications such as fainting and bruising occurred at a rate of 58.3 per 10 000 women screened with CA-125 and 3.3 per 10 000 women screened with transvaginal ultrasound. Of women diagnosed with ovarian cancer, 95 in the intervention group (45%) and 91 in the usual care group (52%) experienced at least 1 major complication associated with their diagnostic procedures (ie, infection, blood loss, bowel injury, cardiovascular events).

Of 3285 women with false-positive results, 1080 underwent surgery (32.9% for oophorectomy) as part of the diagnostic workup. Of these 1080 women, 163 (15%) experienced a total of 222 distinct major complications, which yielded a rate of 20.6 complications per 100 surgical procedures (**TABLE 5**).

Oophorectomy rates were determined for the 22 955 women in the in-

tervention group and the 22 542 in the usual care group who completed the supplemental questionnaire. A total of 1771 women in the intervention group (7.7%) and 1304 in the usual care group (5.8%) reported oophorectomy, which yielded rates of 85.7 and 64.2 per 10 000 person-years, respectively (RR, 1.33; 95% CI, 1.24-1.43).

All-Cause Mortality

All-cause mortality (excluding deaths from ovarian, colorectal, and lung cancer) was similar in the 2 study groups; there were 2924 deaths (76.6 per 10 000 person-years) in the intervention group and 2914 deaths (76.2 per 10 000 person years) in the usual care group (RR, 1.01; 95% CI, 0.96-1.06). Mortality rates for the major causes of death were generally similar between the 2 study groups (TABLE 6).

COMMENT

In this randomized controlled trial, we found no statistically significant reduction in mortality from ovarian cancer in a cohort of women derived from the general population who were screened for ovarian cancer with 6 annual CA-125 tests and 4 annual transvaginal ultrasound examinations. The numbers of deaths from ovarian cancer were similar in the 2 trial groups over the entire period of follow-up, with a modestly (although not statistically significant) greater cause-specific mortality rate in the intervention group (RR, 1.18; 95% CI, 0.82-1.71).

This finding cannot be explained by differences in participant characteristics between the 2 study groups be-

cause they were virtually identical at baseline and follow-up was almost complete in both groups. Stage-specific treatments also were similar across the 2 study groups. Compliance with screening was high in the intervention group and contamination was low in the usual care group. We conclude that the screening intervention, as implemented in this trial, was not effective in reducing mortality caused by ovarian cancer.

In this trial, there was a lack of an observed stage shift. A stage shift (ie, a decrease in the absolute number of late stage [III or IV] cases in the intervention group compared with the usual care group) is thought to be necessary but not sufficient for a mortality benefit to be realized. The lack of a stage shift was apparent in the finding that the total number of advanced stage cancers was greater in the intervention group (n=163) than in the usual care group (n=137). This was true overall as well as when limited to the screening phase of the trial (years 0-5), even though the majority of intervention

group cases diagnosed during this period were detected by screening. Among cases detected by screening, 69% of the cancers were in the late stage, which is a percentage only slightly lower than the usual care group (78%).

This lack of a stage shift suggests the possibility that the 2 screening modalities used (CA-125 and transvaginal ultrasound) with the cutoffs used in the PLCO trial for screening positivity were not effective in detecting ovarian cancers early enough when the cancers were still in a nonadvanced stage. Some evidence from modeling suggests that ovarian tumors need to be found when they are relatively small, considerably smaller than the current threshold used for transvaginal ultrasound (10 cm^3 for cysts) to be in an early stage at detection.[16]

Similarly, had a CA-125 threshold lower than 35 U/mL been used, it may have been possible to detect cancers at an earlier stage; however, this would be at the expense of more false-positive results (and perhaps overdiagnosis of clinically indolent tumors). Other ap-

Table 3. Cancer Stage by Study Period

	No. (%) of Women					
	Intervention Group			Usual Care Group		
	Study Year			Study Year		
	0-5 (n = 126)	6-12 (n = 86)	Total (n = 212)	0-5 (n = 99)	6-12 (n = 77)	Total (n = 176)
Cancer stage						
I	19 (15)	13 (15)	32 (15)	13 (13)	5 (6)	18 (10)
II	11 (9)	4 (5)	15 (7)	13 (13)	7 (9)	20 (11)
III	75 (60)	45 (52)	120 (57)	46 (46)	37 (48)	83 (47)
IV	20 (16)	23 (27)	43 (20)	27 (27)	27 (35)	54 (31)
Unknown	1 (1)	1 (1)	2 (1)	0	1 (1)	1 (1)

Table 4. Treatment by Disease Stage and Trial Group

	No. (%) of Women							
	Intervention Group				Usual Care Group			
	Total	Surgery Plus Systemic Therapy	Surgery Only	No Treatment or Unknown	Total	Surgery Plus Systemic Therapy	Surgery Only	No Treatment or Unknown
Disease stage								
I	**32** (100)	18 (56)	12 (38)	2 (6)	**18** (100)	12 (67)	6 (33)	0
II	**15** (100)	13 (87)	1 (7)	1 (7)	**20** (100)	18 (90)	0	2 (10)
III	**120** (100)	111 (93)	2 (2)	7 (6)	**83** (100)	72 (87)	5 (6)	6 (7)
IV	**43** (100)	28 (65)	4 (9)	11 (26)	**54** (100)	38 (70)	0	16 (30)
Unknown	**2** (100)	1 (<1)	1 (<1)	0	**1** (100)	0	0	1 (<1)
Total	**212** (100)	**171** (81)	**20** (9)	**21** (10)	**176** (100)	**140** (80)	**11** (6)	**25** (14)

Table 5. Major Complications Associated With Diagnostic Evaluation for Ovarian Cancer

	No. (%)		
	Intervention Group		Cancer Cases in Usual Care Group (n = 176)[b]
	No Cancer, Surgical Follow-up (n = 1080)[a]	Cancer (n = 212)[b]	
Women with complications	163 (15)	95 (45)	91 (52)
Total complications[c]	**222** (100)	**140** (100)	**143** (100)
Infection	89 (40)	32 (23)	37 (26)
Direct surgical	63 (28)	69 (49)	61 (43)
Cardiovascular or pulmonary	31 (14)	26 (19)	27 (19)
Other	39 (18)	13 (9)	18 (12)

[a] Includes only women who had a false-positive screening result for ovarian cancer during the screening phase of the trial.
[b] Includes women diagnosed with cancer during the screening phase or follow-up.
[c] Some women had more than 1 complication.

Table 6. Cause of Death Through Year 13[a]

	No. of Women	
Cause of Death	Intervention Group (n = 34 253)	Usual Care Group (n = 34 304)
Other cancer[b]	814	808
Ischemic heart disease	367	386
Cerebrovascular accident	221	205
Other circulatory disease	395	438
Respiratory disease	315	313
Digestive disease	109	101
Infectious disease	87	60
Endocrine and metabolic diseases or immune disorders	122	92
Nervous system diseases	125	130
Accidental	130	121
Other	239	260
Total	**2924**	**2914**

[a] Cause listed on death certificate.
[b] Unrelated to lung, colorectal, or ovarian cancer.

proaches to ovarian cancer screening with these same modalities, including the risk of ovarian cancer algorithm, which incorporates longitudinal CA-125 and considers changes over time rather than a single cut point, may be able to detect ovarian cancers earlier and at a reasonable cost in terms of increased false-positive results; however, the benefit of this approach has not been demonstrated.[17,18]

The risk of ovarian cancer algorithm is currently being evaluated in the randomized UK Collaborative Trial of Ovarian Cancer Screening,[19] which is a 3-group trial comparing 2 strategies of screening (specifically, annual CA-125 testing interpreted using the risk of ovarian cancer algorithm with transvaginal ultrasound as a second-line test [multimodal group] vs annual transvaginal ultrasound only vs no intervention). To date, only the results of the prevalence screening from the UK Collaborative Trial of Ovarian Cancer Screening have been published. The multimodal group had 34 primary invasive epithelial cancers and 16 were stage I or II (47%); in the transvaginal ultrasound group, 50% of 24 such cancers were stage I or II. Whether these modalities will produce a stage shift and demonstrate a mortality benefit compared with usual care await further findings.

An additional randomized controlled trial of ovarian cancer screening is the Shizuoka Cohort Study of Ovarian Cancer Screening in Japan.[4] This trial randomized women to undergo either 5 annual screenings with concurrent transvaginal ultrasound and CA-125 or to a usual care control group. Although the number of women enrolled was similar to that of the PLCO trial, only 35 cancers in the intervention group and 32 cancers in the control group were diagnosed after a mean follow-up of 9.2 years. Mortality findings from this trial have not yet been reported.

It is possible that even an optimized program of annual screening may be insufficient to detect cancers early enough to reduce mortality. Evidence from modeling suggests that aggressive cancers progress rapidly through the early stages, limiting the ability to detect these cancers with yearly screening.[20] In contrast, more ovarian cancers were diagnosed in the screened group than in the usual care group (212 vs 176), suggesting that some of the additional cancers detected by screenings were not clinically important and, if left undetected, may never have caused any symptoms or affected the women during their lifetimes (ie, overdiagnosis).

The false-positive results rate in the PLCO trial was approximately 5% of those screened at each round, with about 60% of these resulting from transvaginal ultrasound (a modality that has trouble distinguishing benign adnexal masses from malignant entities).[21] Although this rate is comparable with or slightly lower than the false-positive results rate of mammography screening, the nature of the diagnostic follow-up, which often included invasive procedures, was a serious concern.[21,22] As evidenced by the differing rate of oophorectomies between groups in this study (33% higher rate in intervention group), false-positive results resulted in more women undergoing major surgery in the intervention group than in the usual care group. Apart from health care costs, some women who were not ultimately diagnosed with ovarian cancer experienced serious medical complications associated with their diagnostic follow-up of a false-positive screening result.

The PLCO trial had certain limitations. The trial was powered for a 35% mortality reduction based on a predicted number of mortality events (n=226) that was essentially met. However, from a public health point of view, smaller effect sizes are still potentially worthwhile to detect. The sequentially adjusted lower 95% CI for the mortality

RR was 0.82, indicating at most an 18% relative benefit within the limits of reasonable probability. Additionally, the data collected on treatment were somewhat limited. The PLCO trial neither abstracted the type of systemic therapy used, nor the type of surgeon who performed the oophorectomy (eg, gynecologic oncologist or not); both factors have been shown to be related to ovarian cancer survival. However, we have no reason to suppose that these factors differed by study group.

We conclude that annual screening for ovarian cancer as performed in the PLCO trial with simultaneous CA-125 and transvaginal ultrasound does not reduce disease-specific mortality in women at average risk for ovarian cancer but does increase invasive medical procedures and associated harms.

Author Affiliations: Huntsman Cancer Institute, University of Utah Health Sciences Center, Salt Lake City (Dr Buys); School of Medicine, University of Alabama, Birmingham (Drs Partridge and Fouad); Epidemiology and Biostatistics Program (Dr Black), Division of Cancer Epidemiology and Genetics (Dr Hartge), Early Detection Research Group, Biometry Research Group, Division of Cancer Prevention (Drs Pinsky, Zhu, Izmirlian, Prorok, and Berg), and Office of Communications and Education (Dr Kramer), National Cancer Institute, Office of Disease Prevention (Dr Gohagan), National Institutes of Health, Bethesda, Maryland; Henry Ford Health System, Detroit, Michigan (Drs Johnson and Lamerato); Lombardi Cancer Center, Georgetown University, Washington, DC (Dr Isaacs); Marshfield Clinic Research Foundation, Marshfield, Wisconsin (Drs Reding and Greenlee); Pacific Health Research and Education Institute, Honolulu, Hawaii (Drs Yokochi and Kessel); Anschutz Cancer Pavilion, University of Colorado, Denver (Dr Crawford); School of Public Health, University of Minnesota, Minneapolis (Dr Church); Division of Urologic Surgery, School of Medicine, Washington University, St Louis, Missouri (Dr Andriole); University of Pittsburgh Medical Center Cancer Pavilion, Pittsburgh, Pennsylvania (Dr Weissfeld); UCLA Immunogenetics Center, Los Angeles, California (Dr Chia); Westat Inc, Rockville, Maryland (Ms O'Brien and Dr Ragard); Information Management Services, Rockville, Maryland (Messrs Clapp, Rathmell, and Riley); and Dalla Lana School of Public Health, University of Toronto, Toronto, Ontario, Canada (Dr Miller).

REFERENCES

1. Jemal A, Siegel R, Ward E, Hao Y, Xu J, Thun MJ. Cancer statistics, 2009. *CA Cancer J Clin*. 2009; 59(4):225-249.

2. Jacobs IJ, Skates SJ, MacDonald N, et al. Screening for ovarian cancer: a pilot randomised controlled trial. *Lancet*. 1999;353(9160):1207-1210.

3. Gohagan JK, Prorok PC, Hayes RB, Kramer BS; Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial Project Team. The Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial of the National Cancer Institute: history, organization, and status. *Control Clin Trials*. 2000;21(6)(suppl):251S-272S.

4. Kobayashi H, Yamada Y, Sado T, et al. A randomized study of screening for ovarian cancer: a multicenter study in Japan. *Int J Gynecol Cancer*. 2008; 18(3):414-420.

5. Prorok PC, Andriole GL, Bresalier RS, et al; Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial Project Team. Design of the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial. *Control Clin Trials*. 2000;21(6)(suppl):273S-309S.

6. Buys SS, Partridge E, Greene MH, et al; PLCO Project Team. Ovarian cancer screening in the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial: findings from the initial screen of a randomized trial. *Am J Obstet Gynecol*. 2005;193(5):1630-1639.

7. Simpson NK, Johnson CC, Ogden SL, et al; Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial Project Team. Recruitment strategies in the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial: the first six years. *Control Clin Trials*. 2000;21(6)(suppl):356S-378S.

8. Stallings FL, Ford ME, Simpson NK, et al; Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial Project Team. Black participation in the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial. *Control Clin Trials*. 2000;21(6)(suppl):379S-389S.

9. Hayes RB, Reding D, Kopp W, et al; Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial Project Team. Etiologic and early marker studies in the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial. *Control Clin Trials*. 2000;21(6)(suppl):349S-355S.

10. Weissfeld JL, Fagerstrom RM, O'Brien B; Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial Project Team. Quality control of cancer screening examination procedures in the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial. *Control Clin Trials*. 2000;21(6)(suppl):390S-399S.

11. Miller AB, Yurgalevitch S, Weissfeld JL; Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial Project Team. Death review process in the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial. *Control Clin Trials*. 2000;21(6)(suppl):400S-406S.

12. Proschan MA, Lan KKG, Wittes JT. *Statistical Monitoring of Clinical Trials: A Unified Approach.* New York, NY: Springer-Verlag; 2006.

13. R Development Core Team. *R: A Language and Environment for Statistical Computing.* Vienna, Austria: R Foundation for Statistical Computing; 2010.

14. Izmirlian G. Estimation of the relative risk following group sequential procedure based upon the weighted log-rank statistic. http://arxiv.org/abs/1102.5088. Accessibility verified May 10, 2011.

15. Ahlbom A. *Biostatistics for Epidemiologists.* Boca Raton, FL: Lewis; 1993.

16. Brown PO, Palmer C. The preclinical natural history of serous ovarian cancer: defining the target for early detection. *PLoS Med*. 2009;6(7):e1000114.

17. Skates SJ, Menon U, MacDonald N, et al. Calculation of the risk of ovarian cancer from serial CA-125 values for preclinical detection in postmenopausal women. *J Clin Oncol*. 2003;21(10)(suppl):206S-210S.

18. Menon U, Skates SJ, Lewis S, et al. Prospective study using the risk of ovarian cancer algorithm to screen for ovarian cancer. *J Clin Oncol*. 2005; 23(31):7919-7926.

19. Menon U, Gentry-Maharaj A, Hallett R, et al. Sensitivity and specificity of multimodal and ultrasound screening for ovarian cancer, and stage distribution of detected cancers: results of the prevalence screen of the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS). *Lancet Oncol*. 2009;10(4):327-340.

20. Havrilesky LJ, Sanders GD, Kulasingham S, et al. Development of an ovarian cancer screening decision model that incorporates disease heterogeneity: implications for potential mortality reductions. *Cancer*. 2011;117(3):545-553.

21. Partridge E, Kreimer AR, Greenlee RT, et al; PLCO Project Team. Results from four rounds of ovarian cancer screening in a randomized trial. *Obstet Gynecol*. 2009;113(4):775-782.

22. Elmore JG, Barton MB, Moceri VM, Polk S, Arena PJ, Fletcher SW. Ten-year risk of false positive screening mammograms and clinical breast examinations. *N Engl J Med*. 1998;338(16):1089-1096.