



Published in final edited form as:

*Cancer Epidemiol.* 2012 December ; 36(6): e401–e406. doi:10.1016/j.canep.2012.08.008.

## Prostate cancer specific survival in the Prostate, Lung, Colorectal, and Ovarian (PLCO) Cancer Screening Trial

Paul F. Pinsky<sup>a,\*</sup>, Amanda Black<sup>b</sup>, Howard L. Parnes<sup>a</sup>, Robert Grubb<sup>c</sup>, E. David Crawford<sup>d</sup>, Anthony Miller<sup>e</sup>, Douglas Reding<sup>f</sup>, and Gerald Andriole<sup>c</sup>

<sup>a</sup>Division of Cancer Prevention, National Cancer Institute, National Institutes of Health, Bethesda, MD, United States

<sup>b</sup>Division of Cancer Epidemiology and Genetics, National Cancer Institute, National Institutes of Health, Bethesda, MD, United States

<sup>c</sup>Division of Urologic Surgery, Washington University School of Medicine, St. Louis, MO, United States

<sup>d</sup>University of Colorado, Denver, Denver, CO, United States

<sup>e</sup>Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada

<sup>f</sup>Marshfield Clinic, Marshfield, WI, United States

### Abstract

**Background**—The prostate component of the Prostate, Lung, Colorectal, and Ovarian (PLCO) randomized screening trial demonstrated no mortality effect of screening. Here we analyze prostate cancer specific survival in PLCO and its relation to screening.

**Methods**—76,693 men aged 55–74 were randomized to usual care ( $n = 38,350$ ) or intervention ( $n = 38,343$ ). Intervention arm men received annual prostate-specific antigen (6 years) and digital rectal exam (4 years). Men were followed for cancer diagnosis and mortality through 13 years. Medical record abstractors confirmed prostate cancer diagnoses, stage and grade. Prostate-specific survival in PLCO cases was analyzed using Kaplan–Meier analysis and proportional hazards modeling. We utilized data from the Surveillance, Epidemiology and End Results (SEER) program to compute expected survival in PLCO and compared this to observed.

**Results**—There was no significant difference in prostate-specific survival rates between arms; 10 year survival rates were 94.7% (intervention,  $n = 4250$  cases) versus 93.5% (usual care,  $n = 3815$  cases). Within the intervention arm, cases never screened in PLCO had lower 10 year survival rates (82%) than screen detected or interval (following a negative screen) cases, both around 95.5%. The ratio of observed to expected 10 year prostate-specific death (1-survival) rates was 0.59 (95% CI: 0.51–0.68) for all PLCO cases, 0.66 (95% CI: 0.51–0.81) for Gleason 5–7 cases and 1.07 (95% CI: 0.87–1.3) for Gleason 8–10 cases.

**Conclusion**—Prostate cancer specific survival in PLCO was comparable across arms and significantly better than expected based on nationwide population data. How much of the better survival is due to a healthy volunteer effect and to lead-time and overdiagnosis biases is not readily determinable.

\*Corresponding author. Tel.: +1 301 402 6480; fax: +1 301 480 0465. pp4f@nih.gov (P.F. Pinsky).

#### Conflict of interest statement

The authors have no conflicts of interest with regards to this manuscript.

## Keywords

Prostate cancer; Screening; PSA; Survival; Gleason grade

---

## 1. Introduction

In 2009, two randomized controlled trials reported the results of the effect of screening with prostate-specific antigen (PSA) on mortality from prostate cancer. The European Randomized Trial of Screening for Prostate Cancer (ERSPC) showed a 20% reduction in prostate cancer mortality; however, the Prostate, Lung, Colorectal, and Ovarian (PLCO) Cancer Screening Trial, a multi-center U.S. based study, showed no mortality benefit through 10 years of follow-up [1,2]. A recently published updated analysis of PLCO through 13 years of follow-up continued to demonstrate no benefit, with a mortality relative risk (screened versus control arm) of 1.09 [3].

As is standard in randomized trials of cancer screening, PLCO and ERSPC both used disease-specific mortality as their primary endpoint. Disease-specific survival can be a misleading endpoint, as it may be confounded by lead-time and over-diagnosis, leading to, for example, artificially inflated 5 year survival rates for cases detected through screening [4]. Note that in this context, mortality refers to the death rate (events over person years) for all randomized subjects whereas survival refers to the (cumulative) death rate over time in those subjects diagnosed with the cancer of interest. However, for several reasons, it is still of interest to examine survival in PLCO. First, this cohort provides an opportunity to examine survival not only by trial arm, but also by prior (to the PLCO trial) PSA testing and method of detection (e.g., screen detected, interval). Second, since the PLCO showed high contamination rates, i.e., high use of PSA screening in the control arm, alternative analyses, including comparison of the survival, and mortality, experience of both arms to that of the general population, may shed further light on the effects of screening and help elucidate how contamination affected the trial results. Finally, PLCO is a unique source of prostate cancer survival rates in the U.S. among a large cohort of men undergoing intensive PSA screening. As such, these rates can be applied for planning, projection, or other purposes, with the appropriate caveats, to other cohorts or populations subjected to similar screening.

## 2. Methods

The Prostate, Lung, Colorectal, and Ovarian (PLCO) Cancer Screening Trial is a multi-center, randomized controlled trial designed to test the efficacy of screening for four types of cancer in persons aged 55–74 (ClinicalTrials.gov number NCT00002540). The methods have been described elsewhere [5]. Briefly, randomization to an intervention or control arm took place between 1993 and 2001, with almost 155,000 persons randomized at ten screening centers. Men in the screening arm received PSA and DRE at baseline (year 0) and then annually through year 3, and received PSA only (without DRE) at years 4 and 5. Exclusion criteria included history of a PLCO cancer, surgical removal of the entire prostate, having taken finasteride in the past 6 months, and starting in 1995, having had more than one PSA blood test in the past three years. At the time of randomization, subjects filled out a self-administered baseline questionnaire, which inquired about demographics, screening history, and medical history.

A PSA result of >4 ng/ml was considered positive. Prostate cancer cases were ascertained through routine follow-up of positive screens and through an annual study update (ASU) form that inquired about cancer diagnoses. Certified tumor registrars at each screening center abstracted, using a standardized protocol, clinical and pathologic (if available) T, N, and M stage characteristics, diagnostic PSA, Gleason score (biopsy, and from radical

prostatectomy if performed) and primary treatment. There was no centralized pathology review of Gleason scoring.

In the intervention arm, screen-detected cancers were defined as those diagnosed within a window extending 9 months from a positive screen or 9 months from a diagnostic evaluation that was linked to the positive screen. In subjects with at least one PLCO screen, non screen-detected cancers were classified as “interval” if they were diagnosed during the screening phase of the trial (years 0–5) and as “post-screening” otherwise. Cancers in subjects without prior PLCO screens were denoted as “never-screened.”

Deaths were ascertained primarily by means of an annual study update (ASU) form; additionally, to obtain more complete mortality data, ASU follow-up was supplemented by periodic linkage to the National Death Index (NDI). All prostate cancers and deaths from confirmed prostate cancer through 31/12/2009 or through 13 years of follow-up, whichever came first, are included in this analysis; note, these are the same cases included in the 13 year PLCO update paper [3]. Causes of death were reviewed by an end-point adjudication process [1].

## 2.1. Statistical methods

We estimated prostate-cancer specific survival rates using the Kaplan–Meier method; deaths from other causes and losses-to-follow-up were treated as censored observations [6]. The log-rank test was used to compare survival for different PLCO categories of cases. Proportional hazards modeling was used to examine the association of the following screening-related factors with survival: study arm, method of cancer detection (within the intervention arm), and prior (to PLCO) PSA testing. These factors were examined in univariate as well as multivariate analyses; for the latter, the model controlled for age, calendar year, race, education, D’Amico risk category (the standard low, intermediate, and high categories for stage I/II cancers and a 4th category for stage III/IV cancers) and primary treatment (radical prostatectomy, radiation, hormonal, none/expectant). Note that the baseline questionnaire item on PSA testing only inquired about the 3 year period prior to enrollment.

To compare the observed survival experience of a given subset of PLCO cases (e.g., intervention arm) with that expected based on national population data, specifically based on data from the Surveillance, Epidemiology and End-Results (SEER) program, we created a simulated “SEER cohort” as follows. For each case in the given PLCO subset, we generated a large number of “replicates,” with each replicate having the same age and calendar year of diagnosis (note we did not stratify by race since the proportion of black men in PLCO was small and not too different from that in the general population; 4.5% in PLCO versus 9% in the U.S. for age 55–74). We then randomly generated a “death” or censoring time for each replicate according to the SEER prostate-cancer specific survival rates for that age and diagnosis year [7]. For example, if the 1 and 2 year SEER survival rates were 99% and 97%, respectively, then roughly 1% of the replicates would have a prostate cancer death by 1 year and 3% death by 2 years. The Kaplan–Meier method was then applied to the simulated SEER cohort data to compute an expected survival curve for the given PLCO subset. We also used the SEER cohort to perform proportional hazards analysis comparing PLCO and SEER survival. The hazard ratios (HR) generated from these analyses are good summary measures of the reduction (or increase) in the hazard of prostate-specific death for the PLCO cohort as compared to the general SEER population. The other comparative measure of survival we utilize is the ratio of 10 year death rates (i.e., 1-cumulative survival rates at 10 years) for observed versus expected.

In SEER, through 2002 the cancer grades of well-differentiated, moderately differentiated, and poorly differentiated corresponded to Gleason scores of 2–4, 5–7, and 8–10, respectively; however, starting in 2003, Gleason 7 cancers were categorized as poorly instead of moderately differentiated [8]. Therefore, to compare observed versus expected survival by Gleason category, we limited PLCO cases to those diagnosed in 2002 or earlier (which comprised 56% of the cases and 77% of the survival follow-up time), and utilized the categories of Gleason 5–7 and Gleason 8–10; there were too few Gleason 2–4 cases in PLCO for meaningful analysis. According to SEER guidelines, the recorded Gleason score should be that from radical prostatectomy (RP) if that procedure was done and from the biopsy (or TURP) otherwise. Thus, to match SEER, unless otherwise specified, the Gleason categories for the PLCO data are based on the RP Gleason for subjects with RP and the biopsy Gleason otherwise.

Because survival rates may be related to over-diagnosis, we also utilized SEER data to calculate the expected number of prostate cancer cases in PLCO, overall and by Gleason category. Specifically, age and calendar year-specific SEER incidence rates were applied to the corresponding person years (PY) in PLCO [7]. As above, expected incidence rates for Gleason 5–7 and Gleason 8–10 cancers were only computed through 2002. Unknown Gleason grades in SEER (about 7% for these years and ages) were accounted for by prorating these to the distribution of known Gleason categories (e.g., if 20% of the cases with known Gleason for a given age group and year were Gleason 8–10, then 20% of the unknown Gleason cases for that age group and year were assigned Gleason 8–10 status). We computed the standardized incidence ratio (SIR) for PLCO as the ratio of observed to expected number of cases. The SIR can be taken as a rough measure of the relative over-diagnosis rate in PLCO as compared to that of the general population; however, the SIR does not account for other factors, in addition to screening, that could affect prostate cancer incidence, such as a healthy volunteer effect. By superimposing SEER prostate cancer-specific survival rates onto the SEER incidence rates, we also estimated a standardized mortality ratio (SMR) in PLCO [9]. Note that since, by protocol, PLCO subjects could not have had a prior diagnosis of prostate cancer, one cannot simply utilize population prostate cancer mortality rates to generate the expected number of prostate cancer deaths. Confidence intervals for the SIR and SMR were computed assuming a Poisson distribution for the observed count of PLCO incident cases/deaths; the expected number of cases was assumed known without sampling error.

### 3. Results

A total of 38,343 and 38,350 men were randomized to the intervention and control arms, respectively, of PLCO. Table 1 displays characteristics of the 4250 intervention and 3815 control arm prostate cancer cases observed through 13 years of follow-up. The median (inter-quartile range) follow-up time of cases from cancer diagnosis was 6.1 (3.2–9.1) years. About three quarters were age 65 and over, and slightly over half in each arm received PSA screening prior to trial entry. Of intervention arm cases, almost half (48%) were screen detected, and about 40% post-screening. Over half (53–57%) of the cancers were Gleason 2–6, about one third (31–33%) were Gleason 7, and 10–13% were Gleason 8–10. The absolute number of Gleason 8–10 cancers was borderline significantly greater in the control arm ( $n = 496$ ) than in the intervention arm (434);  $p = 0.04$ . Primary treatments were similar across arms, with about 40% in each arm receiving radical prostatectomy and another 40% receiving radiation (with or without hormonal treatment).

Fig. 1 shows cumulative prostate-specific survival rates by study arm, overall and by Gleason category (2–6, 7, 8–10). Survival curves were not significantly different by arm overall ( $p = 0.29$ , log-rank test), or within Gleason strata ( $p = 0.82$ – $0.91$ , log-rank test).

Survival rates for all cases through 5, 10, and 12 years were 97.7%, 94.7%, and 91.9% in the intervention arm and 97.3%, 93.5%, and 92.5% in the control arm. Ten year survival by Gleason category was 97.9%, 94.9%, and 75.5% for intervention arm Gleason 2–6, 7, and 8–10 cases, respectively, compared to 97.8%, 95.0%, and 73.9% for control arm cases.

Table 2 shows the results of the univariate and multivariate proportional hazards modeling, which examined the association of survival with study arm, method of detection and prior PSA screening. On univariate analysis, prior (to PLCO) PSA screening was significantly associated with survival in both arms combined, with a hazard ratio (HR) for dying of prostate cancer of 0.76 per PSA test. The HRs were attenuated in the multivariate analysis, and were no longer statistically significant. Point estimates of univariate HRs were similar in the intervention arm alone as compared to both arms combined. Study arm was not significant in multivariate (or univariate) analysis.

Within the intervention arm, as compared to the referent group of screen detected cancers, never-screened cancers had significantly poorer survival in both univariate (HR = 3.9) and multivariate (HR = 2.4) analyses; note the latter controlled for D'Amico score and prior PSAs. Survival for post-screening cases did not significantly differ from that of screen detected cases. Although interval cases had similar survival (HR = 1.3) to screen detected cases in the univariate analysis, they had significantly worse survival in the multivariate analysis (HR = 1.8; 95% CI: 1.03–3.2). Ten year survival rates were 95.8% for screen detected cancers, 95.3% for interval cancers and 81.6% for never-screened cancers.

Fig. 2 shows observed and expected prostate-specific survival rates for all cases and by Gleason category (Gleason 5–7 and 8–10). For all cases, and for the Gleason 5–7 cases, observed survival rates were significantly increased compared to the expected rates. In contrast, observed survival rates for the 8–10 cases were similar to expected rates.

Table 3 displays hazard ratios and 1-survival ratios (at 10 years) for observed versus expected prostate-specific mortality for various categories of PLCO cases. The HRs for the intervention and usual care arms were significantly below 1, 0.50, and 0.57, respectively (0.54 for both arms combined). The never-screened (intervention arm) cases had an elevated HR of 1.44, although this did not reach the level of statistical significance; note that there were only 21 deaths in this group of 194 cancers, so the power to detect a significant difference was low. With respect to Gleason categories, the HR for Gleason 5–7 cases (both arms) was substantially below one, 0.62 (95% CI: 0.52–0.75), whereas the HR for all Gleason 8–10 cases was slightly above 1.0 (HR = 1.07; 95% CI: 0.88–1.30). The HRs for the Gleason 5–7 and Gleason 8–10 cases in the intervention arm were similar to the HRs for both arms combined. Across all comparisons, the ratios of observed versus expected 1-survival rates were generally similar to the estimated HRs.

The HRs for the different Gleason categories may be related to the relative over-diagnosis rates for these categories. Table 4 shows standardized incidence ratios (SIR) for all cases and by Gleason category. For the period through 2002, SIRs for Gleason 5–7 disease were 174 (95% CI: 166–181) for the intervention and 135 (95% CI: 127–140) for the control arm. In contrast, SIRs for Gleason 8–10 disease during this period were 0.64 (95% CI: 0.55–0.73) and 0.72 (95% CI: 0.63–0.81) for the intervention and control arms, respectively. Due to these differences, the percent of cases (diagnosed through 2002) in PLCO that were Gleason 8–10, 9.0%, was considerably lower than the expected 20%. For all prostate cancers through 2002 (which was study year 6.2 on average), SIRs were 1.52 (95% CI: 1.46–1.58) for the intervention and 1.21 (95% CI: 1.15–1.26) for the control arm. For the entire period of follow-up the SIRs were 1.23 (95% CI: 1.19–1.27) for intervention and 1.11 (95% CI: 1.07–1.15) for control. SMRs for the entire period of follow-up (for all prostate cancers) were



0.60 (95% CI: 0.50–0.69) in the intervention and 0.55 (95% CI: 0.46–0.64) in the control arm.

We also examined HRs (observed versus expected) for the outcome of all-cause survival. HRs were 0.69 (95% CI: 0.65–0.73) for all cases, 0.76 (95% CI: 0.70–0.82) for Gleason 5–7 cases and 0.90 (95% CI: 0.77–1.03) for Gleason 8–10 cases. For all cases, 23% of deaths were due to prostate cancer; these percentages were 14% and 57% for Gleason 5–7 and Gleason 8–10 cancers, respectively.

#### 4. Discussion

Prostate-specific survival did not significantly differ by study arm in the PLCO trial, either overall or within Gleason strata. Further, within the intervention arm, survival rates, in univariate analyses, did not differ between screen detected cases, interval cases (generally, those following a negative screen) and post-screening period cases. However, never-screened cancers had significantly worse survival compared to the screen detected cancers, an association that persisted, although with a somewhat attenuated HR, in a multivariate analysis that controlled for D'Amico risk category and primary treatment.

That interval and screen detected cancers had similar survival might be explained by the fact that only a minority of the men with interval cancers (around 30%) presented with symptoms consistent with prostate cancer, with roughly half being referred due to high or “rising” PSA (these findings are derived from medical record review; the specific types of symptoms were not recorded). A similar profile was seen for the post-screening cases, with only about 20% presenting with symptoms. On the other hand, the never screened cases had a similarly low percentage (26%) presenting with symptoms (some of these cases may have been screened outside of the trial). Note also that given the high survival rates, many of these symptoms were not likely resultant from the prostate cancer itself but from benign prostatic hyperplasia (BPH); if so, the cancer could be considered to be detected early, as with PSA-based detection. Interestingly, on multivariate, but not univariate analysis, the interval cases had significantly worse survival than the screen detected cases. This is explained, in part, by the fact that the D'Amico risk category profile was actually more favorable in the interval (62% low risk) than in the screen detected cases (53% low risk) and the interval cases were more likely to have had prior PSA testing than the screen detected cases (68% versus 56%); thus, controlling for D'Amico risk and prior PSA, the interval cases had modestly worse survival. With respect to the never-screened cases, behavioral, health care access and/or health status factors, over and above the choice of primary treatment, which was controlled for in the model, might have affected prostate-specific survival.

We found that having had fewer PSA tests prior to PLCO enrollment was associated with worse survival on univariate analysis. This may be related to a “selection bias” effect. Men with prior PSA tests are less likely to have detectable pre-existing, preclinical prostate cancer than men without a prior PSA since such men (with prior tests) would likely have had their cancers detected by the PSA (and later diagnosed), making them ineligible for PLCO. This implies that the cancers in men with no prior PSA tests would have, on average, longer pre-clinical sojourn times at diagnosis than those in men with prior PSAs. If these cancers were further along in their progression at diagnosis, this could result in reduced observed survival. Men with prior PSA tests had slightly more favorable D'Amico risk profiles (48% low risk) than men with no prior PSA tests (43% low risk). After controlling for D'Amico risk category, and other covariates, prior PSA testing was no longer significantly associated with survival.

As has been described in detail, there was considerable use of PSA screening in the control arm of PLCO [10]. The estimated mean number of PSA tests during the screening period of PLCO (year 0–5) was 2.7 in the control arm, compared to 5.0 in the intervention arm. Further, 74% of control arm men were estimated to have had at least one screening PSA during this period, versus 95% in the intervention arm. This led to the relatively low observed over-diagnosis rate of 11% between arms in PLCO and also to a relatively small estimated increase in lead time in the intervention arm; among cases diagnosed in the first 7 study years (screening period plus 2 years for extended diagnostic follow-up), intervention arm cases were diagnosed on average 0.7 years earlier than usual care cases.

In contrast to the findings here, the 11 year follow-up of the ERSPC did demonstrate significantly increased prostate cancer-specific survival for intervention versus control arm men within Gleason strata (2–6 and 7–10) [11]. These discrepant findings between PLCO and ERSPC are likely due in part to the substantially higher contamination rate in PLCO, which attenuated the differential (by arm) effects on observed survival of lead-time and over-diagnosis. Note that compared to the 1.11 rate ratio for prostate cancer incidence in the intervention compared to control arm in PLCO, in ERSPC the corresponding rate ratio was 1.63, indicating much less relative overdiagnosis in PLCO [11]. The discrepancy also reflects the differential mortality relative risks in the two trials.

Screening PSA usage in control arm men was estimated to be 30% higher than in the general population (for a similarly aged group), and as described above, intervention arm men had about twice as many PSA tests as control arm men [10]. The increased lead time and over-diagnosis resultant from this extra screening helps explain the finding that the survival of control arm, as well as that of intervention arm, cases was considerably higher than expected based on overall population survival rates, with HRs for observed versus expected of 0.57 (control) and 0.50 (intervention).

The SMRs for the cohort also demonstrated significantly decreased risks of prostate cancer death (0.60 and 0.55 for intervention and control arm, respectively). Note that SMRs, which are derived from the entire population, not just cases, avoid the biases of lead time and overdiagnosis that are present with survival comparisons. However, the SMRs, possibly to a greater extent than the survival comparisons, are influenced by selection bias, in the form of a healthy volunteer effect. Such an effect has been previously demonstrated in PLCO [12].

To help gauge the magnitude of a possible healthy volunteer effect we used two genitourinary “control” cancers, bladder cancer and renal cancer. Again, as for prostate cancer, a simulated SEER cohort was generated and a proportional hazards analysis was performed; SMRs were also calculated. The hazard ratios for bladder and renal cancer survival were 0.89 (95% CI: 0.75–1.06) and 0.87 (95% CI: 0.75–1.04), respectively, indicating a modest possible healthy volunteer effect for survival from these cancers. If a similar healthy volunteer effect held for prostate cancer, then roughly one third to one quarter of the observed difference in survival could be due to such an effect. For the remainder, resolving whether lead-time and overdiagnosis could explain most or all of it will probably require complex modeling.

SMRs were 0.73 (95% CI: 0.60–0.86) for bladder and 0.77 (95% CI: 0.63–0.92) for renal cancer; both were statistically significantly higher than the SMR for prostate cancer. However, the healthy volunteer effect may affect different cancers differently; it is possible that this effect is simply stronger for prostate cancer than for renal or bladder cancer.

It is of interest that the HR for observed versus expected survival was decreased for the Gleason 5–7 cases (HR = 0.62) but not for the Gleason 8–10 cases (HR = 1.08). This clearly demonstrates a differential impact of screening by Gleason category. This finding may be

related to the difference in overdiagnosis patterns between the two groups of cases; in contrast to the Gleason 5–7 cases, which showed significantly elevated SIRs in both arms, the intervention and control arm Gleason 8–10 cases had SIRs significantly below 1 (0.64 and 0.72, respectively). In addition to the SIRs under 1, there was also a 12% decrease in the number of Gleason 8–10 cases in the intervention compared to the control arm (434 versus 496;  $p = 0.04$ ). This observed reduced rate of Gleason 8–10 disease in a highly screened population is potentially important in that it provides indirect evidence that Gleason grade progresses over time, since if Gleason grade progresses one would expect earlier diagnosis to reduce the rate of Gleason 8–10 disease. Based on data presented in their supplementary appendix, the 11 year follow-up of ERSPC demonstrated a statistically significant 18% reduction in the population rate of Gleason 8–10 cancers in the screened versus the control arm [11]. There is some evidence from serial biopsy studies that Gleason score 6 cancers can progress; Sheridan et al. [13] found a 19% rate of progression, although some (or all) of this could be due to sampling variability in the biopsies. Choo et al. [14] reported no consistent change in Gleason score on repeat biopsy at a median of 22 months follow-up. A reduced rate of Gleason 8–10 disease caused by PSA screening might result in an eventual beneficial effect of such screening on prostate cancer mortality. In PLCO, further follow-up (past 13 years) is ongoing so the trial will be able to assess the longer term consequences of this reduction on mortality.

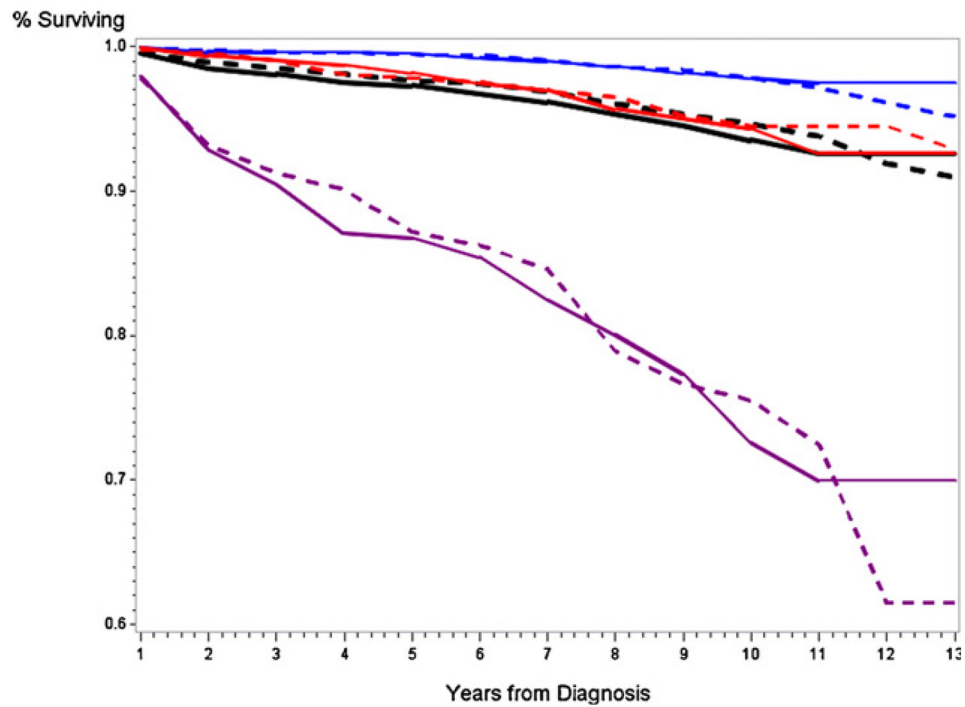
In conclusion, prostate cancer specific survival in the PLCO trial was comparable across arms and significantly better than expected based on nationwide population data. This difference was limited to Gleason 5–7 cases, however, and was not observed for Gleason 8–10 cases. SIRs, indicators of relative over-diagnosis, were significantly elevated (in both arms) for Gleason 5–7 disease, but significantly decreased for Gleason 8–10 disease. SMRs were significantly decreased in both arms. How much of the improved survival and mortality is due to a healthy volunteer effect and to lead-time and overdiagnosis biases (for survival) is not readily determinable.

## References

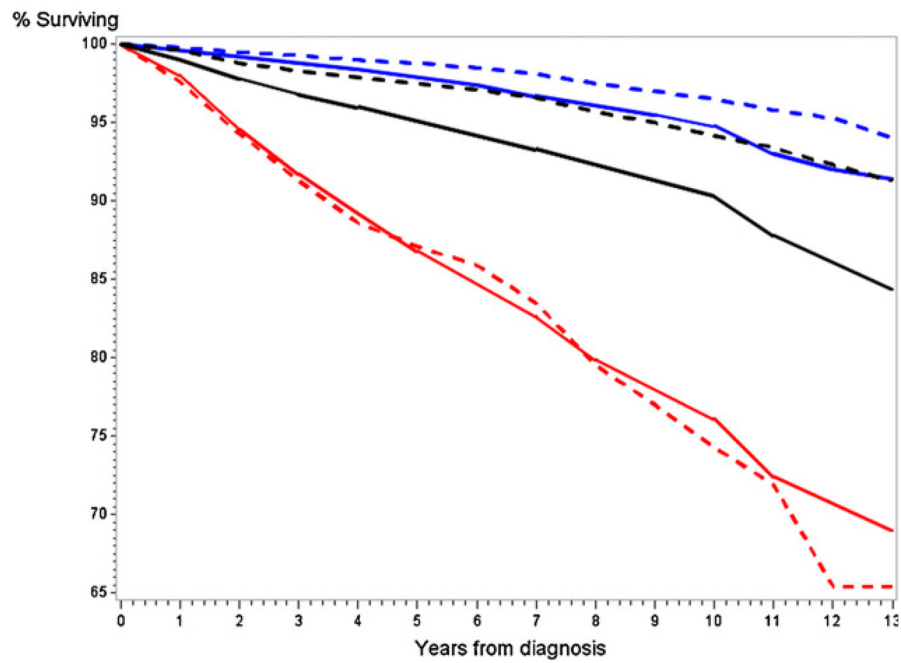
1. Andriole GA, Crawford ED, Grubb R, Buys S, Chia D, Church TR, et al. Mortality results from a randomized prostate-cancer screening trial. *N Engl J Med*. 2009; 360:1310–9. [PubMed: 19297565]
2. Schroder JH, Hugosson J, Roobol MJ, Tammela T, Ciatto S, Nelen V, et al. Screening and prostate-cancer mortality in a randomized European study. *N Engl J Med*. 2009; 360:1320–8. [PubMed: 19297566]
3. Andriole GL, Crawford ED, Grubb R, Buys S, Chia D, Church TR, et al. Screening for prostate cancer, 13-year update of the results of the prostate component of the Prostate, Lung, Colorectal and Ovarian randomized cancer screening trial. *J Natl Cancer Inst*. 2012; 104:1–8.
4. Croswell JM, Ransohoff DF, Kramer BS. Principles of cancer screening: lessons from history and study design issues. *Semin Oncol*. 2010; 37:202–15. [PubMed: 20709205]
5. Prorok P, Andriole GL, Bresalier RS, Buys S, Chia D, Crawford ED, et al. Design of the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial. *Control Clin Trials*. 2000; 21:273S–309S. [PubMed: 11189684]
6. Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. *J Am Stat Assoc*. 1958; 53:457–81.
7. Surveillance Research Program. [accessed July, 2011] National Cancer Institute SEER\*Stat software. (<http://www.seer.cancer.gov/seerstat>) version 6.5.1
8. Lin DW, Porter M, Montgomery B. Treatment and survival outcomes in young men diagnosed with prostate cancer. *Cancer*. 2009; 115:2863–71. [PubMed: 19466697]
9. Pinsky PF. Antioxidant supplementation and cancer prevention. *J Am Med Assoc*. 2009; 301:1878.



10. Pinsky PF, Black A, Kramer BS. Assessing contamination and compliance in the prostate component of the Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial. *Clin Trials*. 2010; 7:303–11. [PubMed: 20571134]
11. Schroder FH, Hugosson J, Roobol MJ, Tammela T, Ciatto S, Nelen V, et al. Prostate-cancer mortality at 11 years follow-up. *N Engl J Med*. 2012; 366:981–90. [PubMed: 22417251]
12. Pinsky PF, Miller A, Kramer BS, Church T, Reding D, Prorok P, et al. Evidence of a healthy volunteer effect in the Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial. *Am J Epidemiol*. 2007; 165:874–81. [PubMed: 17244633]
13. Sheridan TB, Carter HB, Wang W, Landis PB, Epstein JI. Change in prostate cancer grade over time in men followed expectantly for Stage T1c disease. *J Urol*. 2008; 179:901–5. [PubMed: 18207195]
14. Choo R, Danjoux C, Morton G, Szumacher E, Sugar L, Gardner S, et al. How much does Gleason grade of follow-up biopsy differ from that of initial biopsy in untreated, Gleason score 4–7 clinically localized prostate cancer? *Prostate*. 2007; 67:1614–20. [PubMed: 17823923]
15. D'Amico AV, Whittington R, Malkowicz SB, Weinstein M, Tomaszewski JE, Schultz D, et al. Predicting prostate specific antigen outcome preoperatively in the prostate specific antigen era. *J Urol*. 2001; 166:2185–8. [PubMed: 11696732]



**Fig 1.** Prostate cancer specific survival rates by arm and Gleason category. Solid lines are control arm cases, dotted line intervention arm cases. Black curves are for all cases, blue curves Gleason 2–6 cases, red curves Gleason 7 cases, and purple curves Gleason 8–10 cases. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)



**Fig 2.** Observed (in PLCO) versus expected (based on SEER) prostate cancer specific survival rates. Solid lines are expected rates, dashed lines observed rates. Black curves show all cases, red curves Gleason 8–10 cases and blue curves Gleason 5–7 cases. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of the article.)

**Table 1**

Prostate cancer cases in PLCO.

	Intervention arm (n = 4250 cases)	Control arm (n = 3815 cases)
	# (%) of Cases	# (%) of Cases
Age 55–64	1121 (26.4)	905 (23.7)
Age 65–74	2503 (58.9)	2237 (58.6)
Age 75+	626 (14.7)	673 (17.6)
Education <sup>a</sup>		
Less than high school	330 (7.8)	279 (7.9)
High school	2096 (49.3)	1828 (47.9)
College graduate	1748 (41.1)	1548 (40.6)
Prior PSA screening <sup>b</sup>	2039 (53.1)	1854 (54.8%)
Method of detection		
Screen diagnosed	2050 (48.2)	N/A
Interval	374 (8.8)	N/A
Never screened	192 (4.5)	N/A
Post screening	1634 (38.5)	N/A
Gleason score <sup>a,c</sup>		
2–6	2429 (57.3)	2016 (53.0)
7	1318 (31.1)	1257 (33.0)
8–10	434 (10.2)	496 (13.0)
D'Amico risk category <sup>a,d</sup>		
Low	1952 (45.9)	1474 (38.6)
Intermediate	1457 (34.3)	1353 (35.5)
High	451 (10.6)	498 (13.1)
Clinical stage III/IV	154 (3.6)	176 (4.6)
Primary treatment <sup>a</sup>		
Radical prostatectomy	1700 (40.0)	1396 (36.6)
Radiation	1684 (39.6)	1609 (42.2)
Hormonal	302 (7.1)	329 (8.6)
None/expectant	487 (11.5)	403 (10.6)
Deaths from prostate cancer	158	145

<sup>a</sup>Does not sum up to total due to a small number of unknowns.

<sup>b</sup>In 3-year period prior to randomization into PLCO. Percents exclude unknowns (about 10%).

<sup>c</sup>Gleason score based on RP for those with RP, otherwise on biopsy.

<sup>d</sup>Low is Gleason 6, diagnostic PSA <10 ng/ml and clinical stage up to T2a; intermediate is Gleason 7, PSA 10–20 ng/ml or stage T2b; high is Gleason 8–10, PSA >20 ng/ml or stage T2c [15].

**Table 2**

Proportional hazards models of prostate specific survival in PLCO.

Factor	Both arms		Intervention arm	
	Univariate HR (95% CI)	Multivariate HR (95% CI)	Univariate HR (95% CI)	Multivariate HR (95% CI)
0 Prior PSA tests	(Referent)	(Referent)	(Referent)	(Referent)
1 Prior PSA test	0.71 (0.54–0.94)	0.98 (0.74–1.3)	0.65 (0.44–0.96)	0.87 (0.58–1.29)
2 Prior PSA tests	0.60 (0.40–0.90)	0.73 (0.48–1.1)	0.64 (0.38–1.10)	0.82 (0.46–1.48)
# of PSA tests <sup>a</sup>	0.76 (0.62–0.88)	0.89 (0.74–1.06)	0.75 (0.58–0.97)	0.90 (0.69–1.16)
Control arm	(Referent)	(Referent)		
Intervention arm	0.86 (0.68–1.09)	0.95 (0.75–1.21)		
Screen detected case			(Referent)	(Referent)
Interval case			1.3 (0.7–2.2)	1.8 (1.03–3.2)
Never screened case			3.9 (2.4–6.4)	2.4 (1.3–4.4)
Post-screening case			1.4 (0.9–2.3)	1.1 (0.7–1.5)

*Note:* Multivariate model also controlled for calendar year, black race, age, education, D'Amico risk category, and primary treatment, in addition to listed factors.

<sup>a</sup>HR per additional (prior to PLCO) PSA test.



**Table 3**

Observed versus expected prostate cancer specific survival.

<b>Group</b>	<b>Hazard ratio (95% CI)<sup>a</sup></b>	<b>Ratio of 10 year case fatality rates (95% CI)<sup>b</sup></b>
All cases	0.54 (0.47–0.60)	0.59 (0.51–0.68)
Intervention arm	0.50 (0.43–0.59)	0.54 (0.42–0.67)
Control arm	0.57 (0.48–0.67)	0.66 (0.53–0.78)
Intervention (year 0–5, 1+ screen)	0.46 (0.38–0.56)	0.47 (0.37–0.57)
Intervention, no PLCO screens	1.44 (0.89–2.3)	1.79 (0.98–2.6)
All Gleason 5–7	0.62 (0.52–0.75)	0.66 (0.51–0.81)
All Gleason 8–10	1.08 (0.88–1.30)	1.07 (0.87–1.27)

*Note:* Number of deaths from prostate cancer were as follows: 303 (all cases), 158 (intervention), 145 (control), 102 (intervention, year 0–5), 21 (intervention, no PLCO screens), 126 (all Gleason 5–7), 141 (all Gleason 8–10).

<sup>a</sup>Hazard ratio for prostate survival of PLCO cases relative to expected.

<sup>b</sup>Case fatality rate here is 1-cumulative survival rate. Ratio is observed over expected, where expected is based on SEER rates.

**Table 4**

Standardized incidence (SIR) and mortality (SMR) ratios.

	<b>Gleason 5–7 (for the period through 2002)</b>	<b>Gleason 8–10 (for the period through 2002)</b>	<b>All cases (for the period through 2002)</b>	<b>All cases (all years)</b>
	<b>Ratio (95% CI)</b>	<b>Ratio (95% CI)</b>	<b>Ratio (95% CI)</b>	<b>Ratio (95% CI)</b>
<b>SIR</b>				
Control arm	1.35 (1.28–1.41)	0.72 (0.63–0.81)	1.21 (1.15–1.26)	1.11 (1.07–1.15)
Intervention arm	1.74 (1.66–1.81)	0.64 (0.55–0.73)	1.52 (1.46–1.58)	1.23 (1.19–1.27)
All	1.54 (1.49–1.59)	0.68 (0.62–0.74)	1.36 (1.32–1.40)	1.17 (1.14–1.20)
<b>SMR</b>				
Control arm				0.55 (0.46–0.64)
Intervention arm				0.60 (0.50–0.69)
All				0.57 (0.51–0.64)

*Note:* Expected incidence and mortality derived from SEER statistics. SIR (SMR) is observed over expected prostate cancer incidence (mortality).