Cancer Surveillance Series: Interpreting Trends in Prostate Cancer—Part III: Quantifying the Link Between Population Prostate-Specific Antigen Testing and Recent Declines in Prostate Cancer Mortality

Ruth Etzioni, Julie M. Legler, Eric J. Feuer, Ray M. Merrill, Kathleen A. Cronin, Benjamin F. Hankey

Background: **The objective of this study was to investigate the circumstances under which dissemination of prostate-specific antigen (PSA) testing, beginning in 1988, could plausibly explain the declines in prostate cancer mortality observed from 1992 through 1994.** ***Methods:*** **We developed a computer simulation model by use of information on population-based PSA testing patterns, cancer detection rates, average lead time (the time by which diagnosis is advanced by screening), and projected decreased risk of death associated with early diagnosis of prostate cancer through PSA testing. The model provides estimates of the number of deaths prevented by PSA testing for the 7-year period from 1988 through 1994 and projects what prostate cancer mortality for these years would have been in the absence of PSA testing.** ***Results:*** **Results were generated by assuming a level of screening efficacy similar to that hypothesized for the Prostate, Lung, Colorectal, and Ovarian (PLCO) Cancer Screening Trial. Under this assumption, the projected mortality in the absence of PSA testing continued the increasing trend observed before 1991 only when it was assumed that the mean lead time was 3 years or less. Projected mortality trends in the absence of PSA screening were not consistent with pre-1991 increasing trends for lead times of 5 years and 7 years.** ***Conclusions:*** **When screening is assumed to be at least as efficacious as hypothesized in the PLCO trial, it is unlikely that the entire decline in prostate cancer mortality can be explained by PSA testing based on current beliefs concerning lead time. Only very short lead times would produce a decline in mortality of the magnitude that has been observed. [J Natl Cancer Inst 1999;91:1033–9]**

This is the third in a series of three articles on recent trends in prostate cancer incidence and mortality in the United States. The first article *(1)* reviewed trends through 1995 for evidence consistent with an effect of prostate-specific antigen (PSA) screening and enumerated several alternative explanations for the observed patterns. One of these, cause-of-death misclassification (or attribution bias), was analyzed in detail in the second article *(2)* for its role in the recent rise and fall in prostate cancer mortality.

The present article extends the scope of the previous articles by explicitly considering rates of utilization of the PSA test from the late 1980s through 1994 *(3)*. Our purpose is to quantify the link between PSA use at the population level and population declines in prostate cancer mortality. Specifically, we address whether PSA testing conducted in the late 1980s and early 1990s could plausibly explain the decline in prostate cancer mortality observed from 1992 through 1994. Naturally, the magnitude and timing of any mortality decline will depend on the frequency of testing in the population and the extent to which the test will diagnose latent disease, about which data are available *(3–5)*. However, features of disease natural history and the benefits of early diagnosis are also important factors to consider. If disease progression is relatively slow or early diagnosis confers only minimal benefit, then, even if everyone in the population is tested, it may be many years before any changes in disease mortality are observed. Conversely, an aggressive disease for which early treatment is effective could yield mortality declines relatively soon after the introduction of screening, even if only a fraction of the population is screened. For our purposes, these factors can be summarized by the following two parameters: 1) the time by which diagnosis is advanced because of screening (lead time) and 2) the amount by which screening reduces the risk of dying of prostate cancer (decreased risk of death from prostate cancer). With the aid of a computer simulation model, we show that only under short lead times and certain levels of benefit can use of the test at levels similar to those recorded in the population explain the decline in mortality observed by 1994. Alternatively, if we assume that the test works in the sense that it confers a specific outcome benefit on an individual level, then the model allows us to assess how long it would take for this to translate into a noticeable decline in prostate cancer mortality in the population.

The "Methods" section describes our computer model in some detail. The model pertains to white men, among whom mortality began declining in 1992. Declines among African-American men were not as early or as sizable as those among white men in the data considered herein (through 1994), although more recent data show declines among African-American men as well. Certainly, the basic framework presented could also be applied to data on African-American men.

METHODS

Our computer simulation model is programmed in GAUSS *(6)* on an IBM RS6000 workstation. The model generates PSA testing histories and mortality events for a cohort of men who were 65–84 years old in 1988. Given information on the dissemination of PSA testing in this population and under specific as-

Affiliations of authors: R. Etzioni, Fred Hutchinson Cancer Research Center, Seattle, WA; J. M. Legler, E. J. Feuer, K. A. Cronin, B. F. Hankey, Division of Cancer Control and Population Sciences, National Cancer Institute, Bethesda, MD; R. M. Merrill, Health Sciences Department, Brigham Young University, Provo, UT.

Correspondence to: Ruth Etzioni, Ph.D., Fred Hutchinson Cancer Research Center, 1100 Fairview Ave. N., MP-665, P. O. Box 19024, Seattle, WA 98109-1024 (e-mail: retzioni@fhcrc.org).

See "Notes" following "References."

sumptions about lead time and decreased risk of death from prostate cancer, the model provides estimates of the number of deaths prevented by PSA testing for the 7-year period from 1988 through 1994. The model's initial year was chosen to be 1988 because it was the 1st year in which PSA testing was performed at non-negligible levels in the population.

The number of prostate cancer deaths prevented by PSA testing depends directly on the difference between the numbers of prostate cancer deaths with and without the test. This difference depends, in turn, on the number of prostate cancer patients detected in association with the test. However, only those patients whose date of diagnosis is actually advanced by the test have the potential to contribute to any reduction in prostate cancer mortality. Thus, the model first identifies these patients whom we refer to as patients with an "early diagnosis." For each patient with an early diagnosis, the model projects a date and cause of death in the presence and absence of the test. This information is then used to generate estimates of the number of deaths prevented each year.

The model uses input data on the dissemination of PSA testing and prostate cancer detection in the Medicare population from 1988 through 1994 *(3)*, all-cause mortality rates from U.S. lifetables *(7)*, and prostate cancer mortality rates from the Surveillance, Epidemiology, and End Results (SEER) Program[1] *(8)*. A full list of inputs to the model is given in Table 1. Fig. 1 summarizes the main steps in the model, which we now describe in more detail.

Generation of a Model for PSA Tests and Patients With an Early Diagnosis

The model generates data for PSA tests and patients with an early diagnosis through the following steps. First, an indicator of annual test utilization is generated for each individual in the cohort. Then, at each test, prostate cancer patients are selected from among those subjects tested; these patients are designated "detected patients" to indicate that their diagnosis occurs in association with a PSA test.

The rates used to identify PSA-detected patients are derived from a PSA utilization database that does not distinguish between screening and diagnostic tests. These rates include tests on symptomatic patients whose date of diagnosis is not likely to be affected by the test. Consequently, the number of patients with an early diagnosis will tend to be lower than the expected number of detected patients based on these rates. We anticipate this phenomenon to manifest itself particularly among individuals who have had a first test. Thus, for first tests, the model adjusts the number of detected patients downward to approximate the corresponding number of patients with an early diagnosis.

Identifying Tested Individuals

To estimate annual test utilization, the model uses recorded PSA testing rates as reported by Legler et al. *(3)*. The PSA testing rates are calculated by use of claims data from a 5% sample of Medicare recipients who resided in SEER areas from 1988 through 1994 inclusive and were 65 years of age or older in 1988. It is necessary to define a cohort that is Medicare eligible from the outset because it is impossible to know the screening history of individuals who become eligible (e.g., by turning 65 years old) at a later date. Indeed, Legler et al. *(3)* justify the choice of this cohort, noting that "this cohort was defined to maximize the likelihood that the first recorded Medicare-billed PSA is the patient's first PSA."

To obtain prediagnosis PSA testing rates, we linked data on patients diagnosed with cancer from the SEER tumor registry to the Medicare data *(9)*. The linkage allowed us to ensure that tests performed after a prostate cancer diagnosis would be excluded. Men not entitled to part B Medicare services and men enrolled in a health maintenance organization were excluded because claims data are not reported for these enrollees; these patients represented less than 8% of potentially eligible patients. In estimating PSA testing rates on a yearly basis, only one test, the first, was counted for each individual in any given calendar year.

Identifying Detected Patients

To estimate the number of detected patients associated with a given test, the model uses cancer detection rates estimated from the linked SEER–Medicare data. We chose to use cancer detection rates from SEER–Medicare rather than from published studies [e.g., *see (4,5)*] because these rates depend heavily on the study population, the study protocol, and compliance to biopsy recommendations. We believed that these factors were likely to differ substantially between clinical studies and the observational setting of the SEER–Medicare data.

Cancer detection rates were estimated by age and were estimated separately for first and second or later screens. To eliminate follow-up tests performed

Table 1. Computer model inputs: decreased risk of death from prostate cancer is varied as described in the text*

A) Age distribution among men 65–84 years old in 1988 *(16)*

Age, y	Probability
65–69	.39
70–74	.29
75–79	.20
80–84	.12

B) Probability of a prostate-specific antigen test

	Probability	
Year	First test[2] *(3)*	Second/later test[3] *(3)*
1988	.012	.000
1989	.036	.004
1990	.064	.018
1991	.163	.053
1992	.196	.143
1993	.146	.246
1994	.092	.314

C) Cancer detection rate: first test[4] *(5)*

Age, y	Rate, No. of cancers detected/No. of individuals tested
65–69	0.043
70–79	0.048
80–84	0.061

D) Proportion of symptomatic cases among men tested for the first time (p). Three different values were considered.

p[5]
0.0
0.016
0.032

E) Cancer detection rate: second or later test[6] *(4)*

Age, y	Rate, No. of cancers detected/No. of individuals tested
65–69	0.0174
70–79	0.0169
80–84	0.0141

F) Prostate cancer relative survival[7] *(8)*

	Relative survival			
Length of survival, y	65–69 y	70–74 y	75–79 y	80–84 y
1	0.97	0.97	0.95	0.90
2	0.93	0.92	0.90	0.83
3	0.88	0.88	0.85	0.77
4	0.84	0.84	0.81	0.72
5	0.82	0.81	0.77	0.67
6	0.79	0.78	0.74	0.64
7	0.77	0.75	0.72	0.60

G) Death rate from other cause[8] *(7)*

Age, y	Rate
65	0.025
70	0.039
75	0.060
80	0.093
85	0.155
90	0.186
95	1

H) Lead time[9]: time by which diagnosis is advanced because of screening

Mean, y	25th percentile	75th percentile
3	1.7	3.9
5	2.8	6.5
7	4.4	8.9

*Superscripted numbers are referenced in Fig. 1.

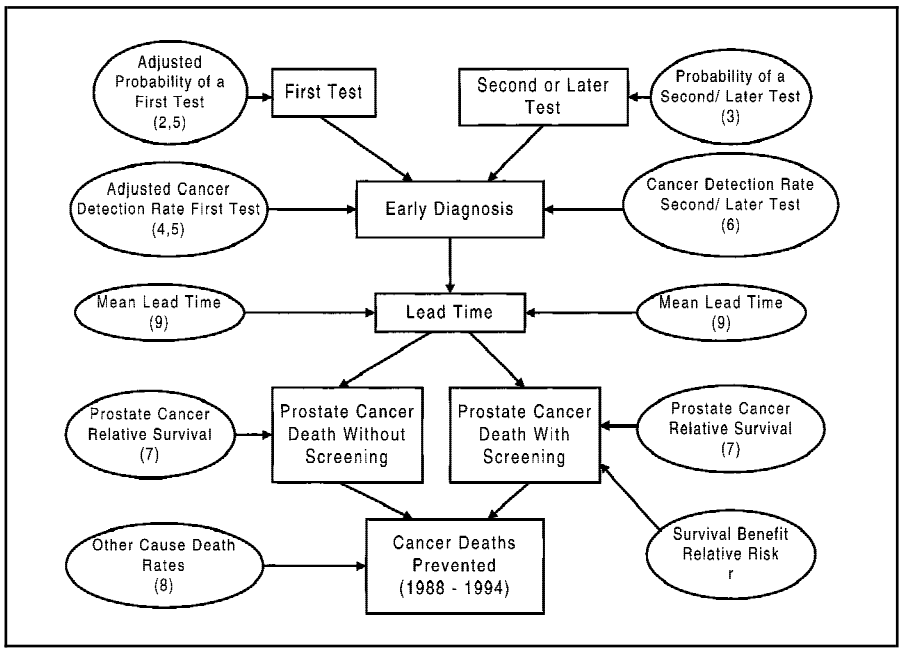

Fig. 1. Main steps in the computer model. Inputs are represented by **ovals. Numbers in parentheses** correspond to superscripts in Table 1. For each individual, indicators of first or later test utilization are generated each year from 1988 through 1994. Among those tested, indicators of early diagnosis are generated. For patients with an early diagnosis, dates of prostate cancer death are generated, given screen detection and in the absence of prostate-specific antigen screening. Finally, the number of prostate cancer deaths prevented each year is given by the difference between the number of prostate cancer deaths without screening and the number of prostate cancer deaths with screening.

shortly after a suspicious result, we defined a testing episode to include all tests performed within 6 months of an initiating test. Initiating tests were allowed to be first tests or subsequent tests performed at least 6 months after a prior test. Cancer detection rates were then estimated per episode. Specifically, the cancer detection rate for first tests consisted of all cancers diagnosed within 3 months of a test falling within a first testing episode, divided by the number of such episodes.[2] A similar rate was computed for second or later testing episodes.

Identifying Patients With an Early Diagnosis

The identification of patients with an early diagnosis requires a further addition to the model that represents an estimate of how often the test is used for diagnostic versus screening purposes. Specifically, let p denote the ratio of the number of patients whose date of diagnosis is not affected by the test to the number of tested individuals. The patients whose data are in the numerator of p are typically those with palpable lesions and will henceforth be referred to as "symptomatic patients." A higher value for p implies greater use of the test in definitive diagnosis of the disease and less in screening for clinically occult disease.

For any given value of p, adjusted testing and cancer detection rates can be derived that effectively eliminate data for symptomatic patients from the corresponding testing and detection rates provided as input (*see* Table 1). The adjusted rates more closely approximate true screening rates. Details of the adjustment are provided in the "Appendix" section. For data from first screens, the adjusted rates are used as indicated in Fig. 1.

Information about the parameter p is not easily obtained from the available data. Therefore, in our analysis, we consider a range of values for p as a sensitivity analysis. We use these values to examine how different relative frequencies of symptomatic patients and patients with an early diagnosis impact on the results of the model. Table 2 gives some adjusted rates that correspond to several values of p. In the model, we use a value for p of 0.016 to represent a moderate level of diagnostic use of the PSA test and a value for p of 0.032 to represent a relatively high level of diagnostic use. These values can be interpreted by considering the implied proportion of symptomatic patients among all patients diagnosed at a first test (Table 2). Thus, for example, among those 70–79 years old, $p = 0.016$ implies that two thirds of detected patients are patients with an early diagnosis, and $p = 0.032$ implies that one third of detected patients are patients with an early diagnosis. Note that, by definition, p cannot exceed the recorded cancer detection rate associated with the test.

Lead Time and Survival

For each patient with an early diagnosis, the model generates a date of death if testing had not occurred and if cancer was detected by a PSA test. Under both scenarios, a date of prostate cancer death is generated; thus, two such dates are produced for each patient with an early diagnosis. In addition, a single date of other-cause death is produced for each individual *(7)* in the original cohort. The date of death is defined as the earlier of the date of death due to other causes and the date of death due to prostate cancer; prostate cancer is assigned as being the cause of death if the date of death due to prostate cancer precedes that due to other causes.

The dates of prostate cancer death with and without screening depend on the lead time and the decreased risk of death from prostate cancer due to screening as follows: Let T_1 be the time of death from prostate cancer if PSA testing was not done and T_2 be the time of death from prostate cancer if PSA testing was done. Then, $T_1 = t_d + lt + s_1$ and $T_2 = t_d + lt + s_2$, where t_d is the time of PSA detection, lt is the lead time, $t_d + lt$ is the original date of diagnosis without PSA testing, and s_1 and s_2 are the survival times from this point to prostate cancer death in the absence and presence of PSA testing, respectively. Thus, $s_2 - s_1$ represents the decreased time to death from prostate cancer due to screening. Note that the model implies that no one can die of prostate cancer during their lead time. We now describe in detail how lt, s_1, and s_2 are generated.

Some information on the average lead time can be obtained from studies of longitudinal changes in PSA before prostate cancer diagnosis. For example, Pearson et al. *(10)* used a mathematical model to estimate that the rate of change in PSA accelerates in patients with prostate cancer beginning at a median time of 7.3 years before diagnosis for local or regional cancers and 9.2 years before diagnosis for advanced or metastatic cancers. The model was applied to retrospectively assayed serum from a cohort of men who were later diagnosed with clinical prostate cancer; serum was stored at roughly 2-year intervals. An empirical estimate of mean lead time from this study, based on the elapsed time from first PSA test that measured greater than 4.0 ng/mL to diagnosis, is only 2.9 years *(11)*. In an analysis of men who developed prostate cancer within 10 years of the start of the Physicians' Health Study, Gann et al. *(12)* estimated the mean lead time for all such cancers as 5.5 years.

The mean lead time among screen-detected cancers is of interest in the current study; this is likely to exceed the estimates from both previous studies because these studies included only men clinically diagnosed within a specific time horizon. In our model, we use three estimates of mean lead time, i.e., 3, 5, and 7 years. The lead time is generated according to a statistical distribution with the specified mean.

Prostate Cancer Death Without PSA Screening

As an example of how the model generates prostate cancer mortality had screening not occurred, consider a patient who is diagnosed with early prostate cancer at age 73 years with an estimated lead time of 3 years. If we assume that PSA testing did not take place and if this patient survived until age 76, he would have been diagnosed with prostate cancer at age 76 years and been subject to disease-specific mortality similar to that of unscreened patients aged 76 years at diagnosis. Thus, in the model, if this patient survives until age 76 years, the time from this point to prostate cancer death is generated by use of population survival curves from the SEER database for prostate cancer patients diagnosed between ages 75 and 79 years inclusive *(8)*. We use data on diagnoses from 1980 through 1987 to represent a recent calendar period that precedes the introduction of PSA testing. We use data from SEER to represent the population survival experience, i.e., the expected survival corresponding to the mix of patients and treatments at the population level. If the date of death due to other causes precedes the date of death due to prostate cancer, then this patient's death is not counted as a prostate cancer death.

Table 2. Adjustment of prostate-specific antigen (PSA) testing and cancer detection rates (CDRs) for the presence of symptomatic prostate cancer patients among patients tested*

Testing rate	CDR	*p*	Adjusted testing rate	ACDR	Proportion of patients with an early diagnosis among detected patients
0.196	0.043	0.0	0.196	0.043	0.0
		0.016	0.193	0.027	0.63
		0.032	0.190	0.011	0.26
0.196	0.048	0.0	0.196	0.048	0.0
		0.016	0.193	0.033	0.67
		0.032	0.190	0.017	0.33
0.196	0.061	0.0	0.196	0.061	0.0
		0.016	0.193	0.046	0.73
		0.032	0.190	0.030	0.48

*Shown is the PSA testing rate from Surveillance, Epidemiology, and End Results–Medicare for 1992 in a cohort of men 65–84 years old in 1988. The CDRs are the rates corresponding to first screens for the age groups 65–69 years, 70–79 years, and 80–84 years. The values *p* represent the proportion of symptomatic patients among men tested. The adjusted cancer detection rates (ACDRs) eliminate the symptomatic patients from the original detection rates (*see* "Appendix" section for the calculation). The proportion of patients with an early diagnosis among detected patients = (1 − *p*/CDR).

Prostate Cancer Death With PSA Screening

In the presence of PSA testing, the time to death is similarly generated; however, once the lead time has elapsed, the annual prostate cancer mortality among prostate cancer patients is a factor *r* (<1) times the prostate cancer mortality rate based on relative survival without screening. Thus, the time from age 76 years to prostate cancer death would be generated from the aforementioned relative survival curve, modified by the factor *r* (mathematically raised to the power *r*). The factor *r* quantifies the decreased risk of death from prostate cancer and is equivalent to the relative risk in a Cox proportional hazards model *(13)*.

Little is known about the decreased risk of death from prostate cancer as a result of screening. In designing the Prostate, Lung, Colorectal, and Ovarian (PLCO) Cancer Screening Trial, the reduction in prostate cancer mortality posited for the sample-size determination was 20% *(14)*. This percentage represents a clinically significant mortality reduction and is the projected relative difference in mortality between the intervention and control groups at exactly 10 years after the start of the trial. If we take into account the expected enrollment schedule, the compliance rate, and the proportion of screen-detected patients in the trial, the "Appendix" section shows mathematically that this translates into a relative risk *r* as defined above of approximately 50%.

Cancer Deaths Prevented Because of PSA Testing

Decreased risk of death from prostate cancer as a result of PSA testing is computed as the number of deaths prevented by screening each year. The number of deaths prevented can be converted to a rate by dividing by the (modeled) population alive in the middle of the year. This "rate of deaths prevented" can be added to the observed prostate cancer mortality rate; the resulting inflated mortality curve represents our best guess as to what prostate cancer mortality trends would have been in the absence of screening. If the inflated curve also shows a downturn in mortality after 1991, then we can conclude that PSA testing does not appear to be responsible for the observed trend reversal from 1991 through 1994. Rather, declines in prostate cancer mortality would have occurred regardless of the introduction of the test. On the other hand, if the inflated curve shows an increasing trend, then we can conclude that PSA testing has been sufficient to explain the observed trend reversal. If the inflated curve shows a horizontal trend after 1991, then this suggests that, without PSA testing, mortality would not have continued to increase after 1991. Consequently, PSA testing has been sufficient to transform what would have been a flat mortality curve into a declining mortality curve; equivalently, one could conclude that it has been sufficient to explain some but not all of the observed trend reversal since 1991. Naturally, any conclusions made are predicated on the validity of the model assumptions and structure.

Results

Results are presented in Figs. 2 and 3 and pertain to men 71–84 years old because data for these men are available for the entire period from 1988 through 1994. In Fig. 2, the observed prostate cancer mortality in the age group 70–84 years from 1980 through 1994 is plotted as an approximation of mortality among men 71–84 years old. The observed mortality is then inflated each year, starting in 1988, by the model-generated "rate of deaths prevented" among men 71–84 years old. As noted above, the inflated curves represent our best guess as to what mortality would have been in the absence of screening. If these inflated curves continue the increasing trends seen before 1988,

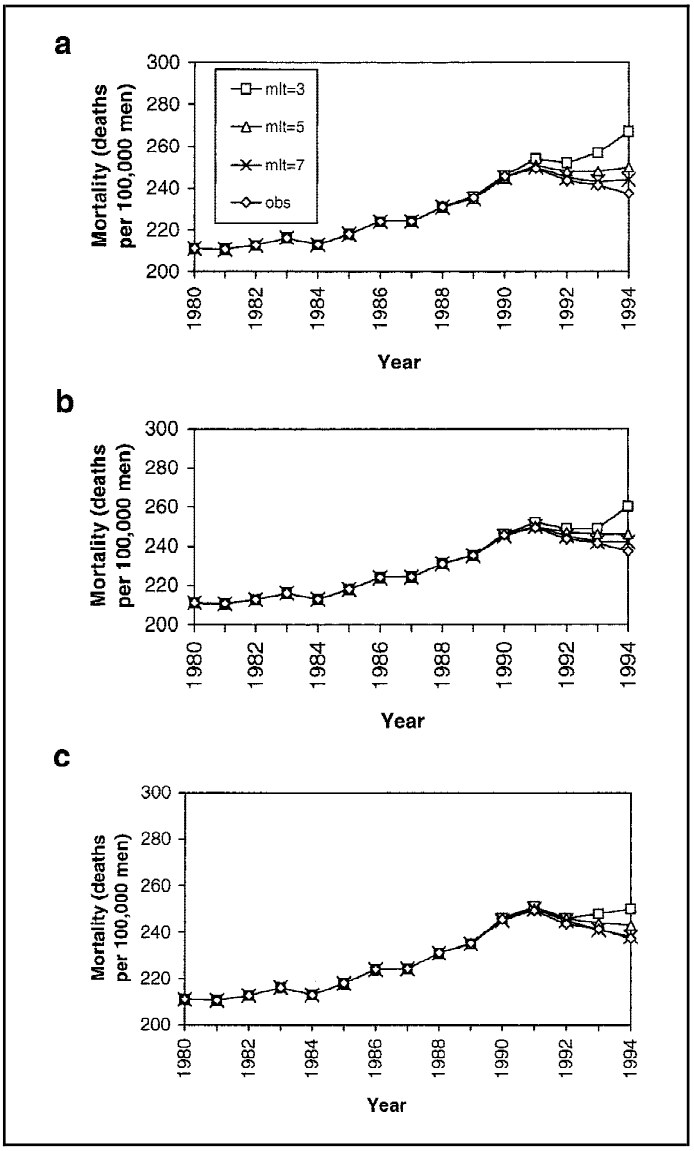

Fig. 2. U.S. prostate cancer mortality for white men 70–84 years old, observed (1980–1994) and inflated (1988–1994) by prostate-specific antigen (PSA)-prevented deaths. Annual prostate mortality rate after lead time is 50% lower among patients with screen-detected prostate cancer than among patients with clinically diagnosed prostate cancer (relative risk $r = 0.5$). **(a)** Relative frequency of symptomatic patients among those tested for the first time (p) is zero; all detected patients with prostate cancer are patients with an early diagnosis. **(b)** $p = 0.016$. **(c)** $p = 0.032$. obs = observed data; mlt = mean lead time (years).

we can conclude that screening alone is able to explain the trend reversal observed in recent years.

In Fig. 2, a, the assumed frequency of clinical cases of prostate cancer p is zero; thus, the testing and cancer detection rates in Table 1 are used. In Fig. 2, b, the assumed value for p is 0.016; thus, the rates of first screens used are 98.4% of the rates in Table 1, and the cancer detection rates are adjusted downward as in Table 2. We consider this value for p as representing moderate use of PSA for diagnostic as opposed to screening purposes. In Fig. 2, c, the assumed value for p is 0.032; thus, the rates of first screens used are 96.8% of those in Table 1, and cancer detection rates are adjusted downward once more, as in Table 2. We consider this value for p as representing relatively frequent use of PSA for diagnostic as opposed to screening purposes.

Fig. 2 provides an analysis of the sensitivity of the model's results to assumptions about the relative frequencies of screening and diagnostic tests. The only circumstances under which the inflated mortality curve shows a convincing increasing trend beyond 1987 are for a mean lead time of 3 years and for a diagnostic testing frequency of zero (Fig. 2, a; uppermost curve). These are extreme circumstances favoring the impact of screening on prostate cancer mortality. The curve representing a mean lead time of 3 years and moderate diagnostic testing frequency (Fig. 2, b) also shows a somewhat increasing trend, but this is not monotonic after 1987. For other mean lead times, the inflated curves show a level or even decreasing trend regardless of the frequency of diagnostic tests, suggesting that, for these lead times, screening in the population has not been sufficient to explain all of the recent declines in prostate cancer mortality.

Fig. 2 restricts attention to a value of 0.5 for the relative risk r, which is commensurate with the projected disease-specific mortality reduction posited when designing the PLCO trial. Fig. 3 presents curves similar to those in Fig. 2 for values of r representing greater mortality reductions as a result of screening. In Fig. 3, a, r is set to 0.4, and in Fig. 3b, r is set to 0.3. In both cases, the value of p is 0.016, representing a moderate level of diagnostic use of the PSA test. Fig. 3 suggests that, for hazard ratios lower than 0.5, the inflated mortality curve is effectively increasing for mean lead times of 5 years and of 3 years.

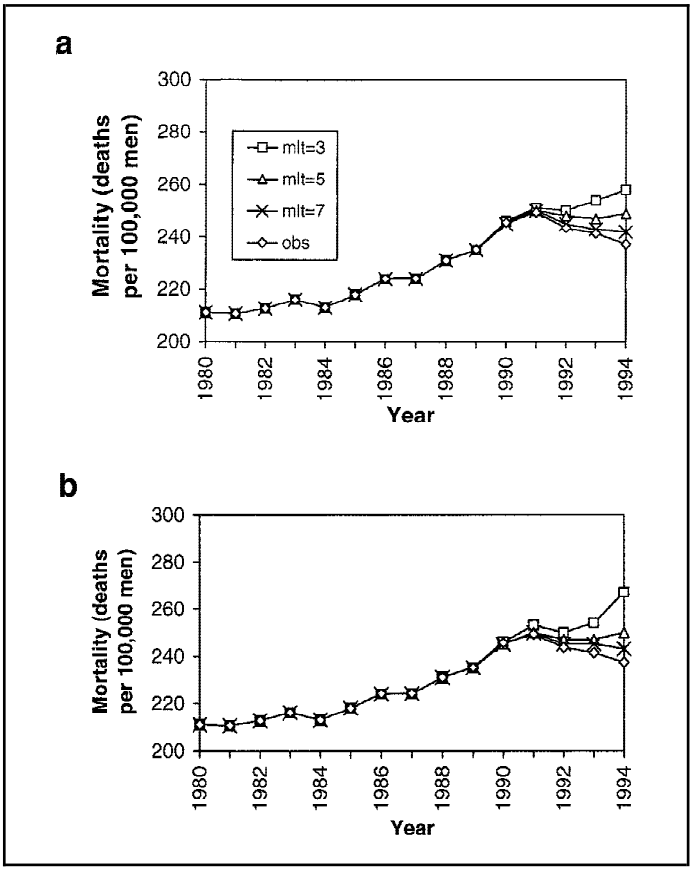

Fig. 3. U.S. prostate cancer mortality for white men 70–84 years old, observed (1980–1994) and inflated (1988–1994) by prostate-specific antigen (PSA)-prevented deaths. Relative frequency of symptomatic cases among those tested for the first time (p) is 0.016, representing moderate diagnostic use of the PSA test. **(a)** Relative risk is 0.4. **(b)** Relative risk is 0.3. obs = observed data; mlt = mean lead time.

DISCUSSION

The simultaneous occurrence of two potentially related events naturally leads one to surmise about the connection between them. For PSA testing and prostate cancer declines, the possibility of a causal relationship is especially tantalizing because of its implications for future control of the disease. However, even before one begins to address the question of causality, the question of consistency between the two events needs to be resolved. In this case, could recorded levels of PSA testing in the population possibly be sufficient to bring about a drop in prostate cancer mortality as soon as that observed and of the magnitude of that observed? The answer depends on what one believes about the lead time and decreased risk of death from prostate cancer as a result of screening. In this article, we have identified values of these key parameters under which the population screening patterns and observed mortality patterns might be said to be consistent in the sense that one could explain the other.

We drew on a number of data sources to estimate model parameters. We had few sources of data on which to base a choice of lead time and a hazard ratio reflecting a screening benefit. We examined several choices for the mean lead time that reflect published data and our belief that the published figures likely underestimate the mean lead time among screen-detected patients. We derived our baseline estimate of the hazard ratio from the effect predicted to design the PLCO study but assessed the sensitivity of our results to various values of this input parameter.

Our test utilization and cancer detection rates were estimated from a linked SEER–Medicare database. Although the linkage between these databases allowed us to eliminate tests performed after diagnosis, we found that it was not possible to easily distinguish between clinically identified patients with cancer, whose diagnosis would have occurred regardless of their use of the test, and patients with a genuine early diagnosis. We have presented a method for obtaining approximate numbers of patients with an early diagnosis that requires an estimate of the prevalence of symptomatic patients among all those tested for the first time. However, because precise knowledge of this quantity is lacking for the years under consideration, we are limited to considering values that represent moderate and relatively high levels of use of the PSA test for diagnostic as opposed to screening purposes. This effectively constitutes a sensitivity analysis on the frequency of true screening tests in the population. It is plausible that the relative frequencies of diagnostic and screening tests vary over time in the population; however, this variation is probably less marked among first-time users than among

all users of the PSA test. Given the lack of precise information about diagnostic versus screening use of the test over time, we did not model this phenomenon. Because Medicare did not reimburse for PSA screening during the modeled calendar period, it is possible that our rates are conservative in that they may underestimate the true frequency of PSA screening.

Our results suggest that, if PSA screening works with an effect on the order of that postulated for the PLCO trial and if the mean lead time is close to 3 years, then PSA testing in the population could explain most or all of the decline in prostate cancer mortality since 1991, as long as diagnostic use of the PSA test in clinically apparent cases of prostate cancer does not account for the majority of the prostate cancers diagnosed in recent years. However, the true effect of PSA screening on prostate cancer mortality is not known, and 3 years would appear to be a lower bound for mean lead time. Indeed, it is likely that the true mean lead time associated with prostate cancer screening is somewhat higher, especially for second or later screens. It is possible that mean lead time is close to 3 years in a subset of patients, and our results suggest that these are the patients who would be responsible for any impact on mortality observed to date. We note that Hankey et al. *(1)* identified a statistically significant decline in the incidence of distant stage disease, beginning in 1990. The patients contributing to this decline, whose disease would have been diagnosed in distant stage but whose disease was presumably diagnosed earlier because of screening, are precisely those who would tend to exhibit shorter lead times. Thus, the analysis of Hankey et al. suggests that screening appears to have detected a substantial number of patients with shorter lead times. The present analysis suggests that these are the patients who would have contributed to mortality declines by the end of 1994. However, based on the analyses in the first two articles of this series *(1,2),* alternative explanations for the mortality declines cannot be discounted. Thus, our findings suggest that the conclusion that PSA testing is wholly responsible for the recent declines in prostate cancer mortality is unwarranted at this time, but PSA testing may provide a partial explanation.

Naturally, computer models are no substitute for results from well-controlled, randomized studies *(15).* However, when an intervention is being applied in a complex, uncontrolled setting, computer models are the most efficient (and possibly the only way) to consolidate the relevant information and to clarify the process by which the intervention might affect outcomes of interest. By replicating the population experience with and without PSA testing, we have identified the circumstances under which the screening that took place could have affected prostate cancer mortality in a manner consistent with that observed.

APPENDIX

1) Adjusting Testing and Cancer Detection Rates to Eliminate Symptomatic Patients

Among all individuals undergoing a first test, let a denote the number of asymptomatic cancer patients, b denote the number of asymptomatic men without prostate cancer, c denote the number of symptomatic patients, and d denote the number of symptomatic men without prostate cancer. Here, a cancer patient is a detected patient, i.e., identified solely through the test or through symptoms or other clinical means at this time. Thus, the group of men without prostate cancer may include some patients with occult disease not detected at this time.

The cancer detection rate (CDR) estimated from the Surveillance, Epidemiology, and End Results (SEER)–Medicare database is given by $(a + c)/(a + b + c + d)$. Let ACDR denote the adjusted CDR, i.e., the CDR associated with a first testing episode once clinical patients have been eliminated. Thus, ACDR is given by $a/(a + b + d)$. Let p denote the proportion of symptomatic patients among men having first tests; $p = c/(a + b + c + d)$. Then

$$\mathrm{CDR} = p + (1 - p)\mathrm{ACDR},$$

so that

$$\mathrm{ACDR} = \frac{\mathrm{CDR} - p}{1 - p}.$$

This gives an expression for the ACDR. The adjusted testing rates are given by the original testing rates multiplied by $(1 - p)$.

The adjusted rates effectively exclude the clinical patients and approximate the screening rates more closely than the unadjusted rates. When p has a value that is different from zero, the adjusted rates are used by the model.

2) Translating the Decreased Risk of Death From Prostate Cancer From the Prostate, Lung, Colorectal, and Ovarian (PLCO) Cancer Screening Trial Into a Direct Benefit Due to Screening

The PLCO Cancer Screening Trial sample size was determined to detect a 20% reduction in the 10-year mortality due to prostate cancer in the intervention arm versus the control arm of the trial. In the absence of contamination (screening among control participants) at specific predicted levels, the corresponding reduction is approximately 27%. Participants randomly assigned to the intervention arm of the trial are scheduled to undergo a baseline PSA test followed by three annual screenings. All subjects will be followed from baseline, and prostate cancer deaths will be noted. Unless otherwise noted, in what follows, "death" refers to prostate cancer death and "survival" refers to the survival time until death from prostate cancer. It is assumed that approximately 40% of the deaths will come from clinically detected patients, and 60% will be from screen-detected patients.

Let:

t = time in years from the beginning of the trial;
$S_1(t)$ = probability a clinically detected patient survives t years from the start of the trial;
$S_2(t)$ = probability a screen-detected patient survives t years from the start of the trial;
$S_2^*(t)$ = probability a patient in the intervention arm survives t years from the start of the trial
$= 0.4S_1(t) + 0.6S_2(t)$.

The decreased risk of death from prostate cancer posited for the PLCO trial is

$$\frac{1 - S_2^*(10)}{1 - S_1(10)} = 0.73. \quad [1]$$

This is equivalent to

$$\frac{1 - S_2(10)}{1 - S_1(10)} = 0.55. \quad [2]$$

In our simulations, we use a value for relative risk r, which is defined as follows:

$$S_2'(s) = [S_1'(s)]^r, \quad [3]$$

where S' represents survival starting from the time of clinical diagnosis and s denotes time from clinical diagnosis; recall that the model projects the time of clinical diagnosis for screen-detected patients by adding the lead time to the time of screen detection.

Since the simulation model's definition of decreased risk of death from prostate cancer is in terms of time from clinical diagnosis and the

PLCO trial definition is in terms of time from the start of the trial, it is necessary to estimate a value for s in equation 3 that corresponds approximately to 10 years from the start of the trial.

Given the expected mixture of clinically detected and screen-detected cases of cancer, the expected time from the start of the trial until a prostate cancer is detected in the intervention arm is

$$E(t_d) = 0.4E(t_{cd}) + 0.6E(t_{cd} - L), \quad [4]$$

where E denotes mathematical expectation, t_{cd} is the time from the start of the trial to clinical detection, and L is the lead time. Preliminary computations performed during the design of the PLCO trial estimated the expected time to detection in the intervention arm at 3.5 years. Given this result, we have

$$E(t_{cd}) = 3.5 + 0.6E(L). \quad [5]$$

Thus, 10 years from the start of the trial corresponds, on average, to $10 - E(t_{cd})$ years from clinical detection, where $E(_{cd})$ depends on the assumed mean lead time. In mathematical terms, $S_1(10)$ in equation 2 is, on average, equivalent to $S_1'[10 - E(t_{cd})]$ in equation 3 and similarly for S_2. Thus, we can rewrite equation 2 as

$$\frac{1 - S_2'[10 - E(t_{cd})]}{1 - S_1'[10 - E(t_{cd})]} = 0.55, \quad [6]$$

and, substituting equation 3 in equation 6, we have

$$r = \frac{\log\{0.45 + 0.55S_1'[10 - E(t_{cd})]\}}{\log\{S_1'[10 - E(t_{cd})]\}}. \quad [7]$$

Appendix Table 1 shows values for r for selected estimates of the mean lead time L. Values for $S_1'(t)$ were estimated from SEER data *(8)*.

Appendix Table 1. Values for r for selected estimates of the mean lead time L

L,y	$E(t_{cd})$	$10 - E(t_{cd})$	$S_1'[10 - E(t_{cd})]$	r
3	4.7	5.3	0.54	0.47
5	5.5	4.5	0.55	0.48
7	6.3	3.7	0.57	0.48

Thus, the estimated value for r is approximately 0.5. This is the case also when the estimated percent of clinically detected cases of cancer is changed from 0.4 to 0.3 and when the expected time to detection in the intervention arm is changed from 3.5 to 4 years. In the first of these cases, the estimated value for r increases slightly, to approximately 0.54 for each lead time. Given these results, it appears that a value of 0.5 for r is consistent with the mortality reduction of 27% in the absence of contamination, posited for the PLCO trial.

REFERENCES

(1) Hankey BF, Feuer EJ, Clegg LX, Hayes RB, Legler JM, Prorok PC, et al. Cancer surveillance series: Interpreting trends in prostate cancer—Part I: Evidence of the effects of screening in recent incidence, mortality, and survival rates. J Natl Cancer Inst 1999;91:1017–24.

(2) Feuer EJ, Merrill RM, Hankey BF. Cancer surveillance series: Interpreting trends in prostate cancer—Part II: Cause of death misclassification and the recent rise and fall in prostate cancer mortality. J Natl Cancer Inst 1999; 91:1025–32.

(3) Legler JM, Feuer EJ, Potosky AL, Merrill RM, Kramer BS. The role of prostate-specific antigen (PSA) testing patterns in the recent prostate cancer incidence decline in the United States. Cancer Causes Control 1998;9: 519–27.

(4) Catalona WJ, Smith DS, Ratliff TL, Basler JW. Detection of organ-confined prostate cancer is increased through prostate-specific antigen-based screening. JAMA 1993;270:948–54.

(5) Richie JP, Catalona WJ, Ahmann FR, Hudson MA, Scardino PT, Flanigan PC, et al. Effect of patient age on early detection of prostate cancer with serum prostate-specific antigen and digital rectal examination. Urology 1993;42:365–74.

(6) GAUSS Mathematical and Statistical System Version 3.2.18. Copyright 1994–1995. Maple Valley (WA): Aptech Systems Inc.; 1995.

(7) National Center for Health Statistics. Vital statistics of the United States, 1986, Vol. II. Mortality Part A. DHHS Publ No. (PHS)88-1122. Washington (DC): US Govt Print Off; 1988.

(8) SEER 1973–1994 public-use CD-ROM. Bethesda (MD): National Cancer Institute; 1997.

(9) Potosky AL, Riley GF, Lubitz JD, Mentnech RM, Kessler LG. Potential for cancer related health services research using a linked Medicare–tumor registry database. Med Care 1993;31:732–48.

(10) Pearson JD, Luderer AA, Metter EJ, Partin AW, Chao DW, Fozard JL, et al. Longitudinal analysis of serial measurements of free and total PSA among men with and without prostatic cancer. Urology 1996;48:4–9.

(11) Morrell CH, Pearson JD, Carter HB, Brant LJ. Estimating unknown transition times using a piecewise nonlinear mixed-effects model in men with prostate cancer. JASA 1995;90:45–53.

(12) Gann PH, Hennekens CH, Stampfer MJ. A prospective evaluation of plasma prostate-specific antigen for detection of prostate cancer. JAMA 1995;273:289–94.

(13) Kalbfleisch JD, Prentice RL. The statistical analysis of failure time data. New York (NY): Wiley; 1980.

(14) Gohagan JK, Prorok PC, Kramer BS, Cornett JE. Prostate cancer screening in the prostate, lung, colorectal, and ovarian cancer screening trial of the National Cancer Institute. J Urol 1994;152:1905–9.

(15) Auvinen A, Rietbergen JB, Denis LJ, Schroder FH, Prorok PC. Prospective evaluation for randomised trials of prostate cancer screening. The International Prostate Cancer Screening Trial Evaluation Group. J Med Screen 1996;3:97–104.

(16) Yax LK (latest update—April 21, 1998). US Census Bureau, National Population Estimates [US Census Bureau homepage] [online]. Available: http://www.census.gov

NOTES

[1]*Editor's note:* SEER is a set of geographically defined, population-based, central cancer registries in the United States, operated by local nonprofit organizations under contract to the National Cancer Institute (NCI). Registry data are submitted electronically to the NCI on a biannual basis, and the NCI makes the data available for analysis.

[2]With adjustment for censoring of individuals entering the episode, e.g., because of other-cause death.

Supported by Public Health Service grant R29CA70227 and contract N01CN 67009 (to R. Etzioni) from the National Cancer Institute, National Institutes of Health, Department of Health and Human Services.

Manuscript received January 8, 1999, revised April 26, 1999; accepted April 27, 1999.