Overdiagnosis Due to Prostate-Specific Antigen Screening: Lessons From U.S. Prostate Cancer Incidence Trends

Ruth Etzioni, David F. Penson, Julie M. Legler, Dante di Tommaso, Rob Boer, Peter H. Gann, Eric J. Feuer

Background: **Overdiagnosis of clinically insignificant prostate cancer is considered a major potential drawback of prostate-specific antigen (PSA) screening. Quantitative estimates of the magnitude of this problem are, however, lacking. We estimated rates of prostate cancer overdiagnosis due to PSA testing that are consistent with the observed incidence of prostate cancer in the United States from 1988 through 1998. Overdiagnosis was defined as the detection of prostate cancer through PSA testing that otherwise would not have been diagnosed within the patient's lifetime.** ***Methods:*** **We developed a computer simulation model of PSA testing and subsequent prostate cancer diagnosis and death from prostate cancer among a hypothetical cohort of two million men who were 60–84 years old in 1988. Given values for the expected lead time—that is, the time by which the test advanced diagnosis—and the expected incidence of prostate cancer in the absence of PSA testing, the model projected the increase in population incidence of prostate cancer associated with PSA testing. By comparing the model-projected incidence with the observed incidence derived from the National Cancer Institute's Surveillance, Epidemiology, and End Results (SEER) registry data, we determined the lead times and corresponding overdiagnosis rates that were consistent with the observed data.** ***Results:*** **SEER data on prostate cancer incidence from 1988 through 1998 were consistent with overdiagnosis rates of approximately 29% for whites and 44% for blacks among men with prostate cancers detected by PSA screening.** ***Conclusions:*** **Among men with prostate cancer that would be detected only at autopsy, these rates correspond to overdiagnosis rates of, at most, 15% in whites and 37% in blacks. The observed trends in prostate cancer incidence are consistent with considerable overdiagnosis among PSA-detected cases. However, the results suggest that the majority of screen-detected cancers diagnosed between 1988 and 1998 would have presented clinically and that only a minority of cases found at autopsy would have been detected by PSA testing. [J Natl Cancer Inst 2002;94: 981–90]**

Since the mid-1980s, dramatic swings have been observed in prostate cancer incidence in the United States. Between 1986 and 1992, the overall age-adjusted incidence rate for prostate cancer increased by over 100%, from 86 to 179 per 100 000 per year *(1).* Although similar patterns of prostate cancer incidence were observed for both white and black men, the incidence among whites peaked in 1992 whereas that among blacks peaked in 1993. Thereafter, incidence in both groups of men declined steadily for several years; recent figures for 1996 through 1998 show prostate cancer incidence returning to pre-1988 levels *(2,3).*

The incidence trends observed since the mid-1980s coincided with the rapid dissemination of the prostate-specific antigen (PSA) test in the population. The PSA test was first approved by the Food and Drug Administration in 1986 as a way to monitor prostate cancer progression; however, its use as a screening test for prostate cancer increased dramatically beginning in 1988 *(4,5),* despite the lack of definitive information regarding its efficacy.

The pattern of cancer incidence following the introduction of a screening test depends on four factors *(6,7)*: 1) the rate of dissemination of the screening technology in the population; 2) the lead time associated with the test (i.e., the time by which the test advances the diagnosis of the disease); 3) the background level of incidence, or the secular trend in incidence, that would be expected in the absence of screening, which is important to consider because other factors besides screening may also affect incidence; and 4) the extent of overdiagnosis due to the test, where overdiagnosis is defined as the detection, through screening, of disease that would never have been diagnosed in the absence of such screening.

Information on some of these factors is available from a number of sources. For instance, annual PSA testing rates may be estimated from administrative data on claims for medical procedures including screening tests *(4,5)* as well as from population surveys conducted in the past decade *(8).* Retrospective studies of PSA testing *(9–11)* have suggested that a range of lead times is associated with the test. Trends in practice patterns, particularly changing approaches to the management of benign prostatic hyperplasia *(12–17),* have provided some clues about the direction of the secular trend in prostate cancer incidence. However, the extent of prostate cancer overdiagnosis due to PSA testing remains unknown. This information is of great importance because considerable morbidity can be associated with treatment for the disease *(18).* Randomized trials of PSA screen-

Affiliations of authors: R. Etzioni, D. di Tommaso, Fred Hutchinson Cancer Research Center, Seattle, WA; D. F. Penson, Veterans Affairs Medical Center, Seattle; J. M. Legler, E. J. Feuer, Applied Research Branch, Cancer Surveillance Research Program, Division of Cancer Control and Population Sciences, National Cancer Institute, Bethesda, MD; R. Boer, RAND Corporation, Santa Monica, CA; P. H. Gann, Northwestern University Medical School, Chicago, IL.

Correspondence to: Ruth Etzioni, Ph.D., Program in Biostatistics, Fred Hutchinson Cancer Research Center, 1100 Fairview Ave. North, MP-665, Seattle, WA 98109–1024 (e-mail: retzioni@fhcrc.org).

See "Notes" following "References."

© Oxford University Press

ing *(19,20)* will presumably, with sufficient follow-up, yield estimates of the expected frequency of overdiagnosis. However, these results are not expected for a number of years.

Given the lack of information from clinical trials about overdiagnosis, the potential use of alternative data sources for estimating the extent of prostate cancer overdiagnosis is of great interest. In particular, population incidence may provide some clues as to the expected rate of overdiagnosis in the population. Therefore, we have asked the following question: Given the dissemination of PSA testing throughout the U.S. population and the expected lead time and the projected secular trend associated with such testing, what extent of prostate cancer overdiagnosis would yield the incidence patterns observed from 1988 through 1998? Throughout this study, we defined a prostate cancer case as an individual diagnosed with the disease and the rate of overdiagnosis as the fraction of cases detected by PSA screening that, in the absence of the test, would not have been diagnosed within the individuals' lifetimes.

METHODS

Overview of the Model

We developed a computer model of PSA testing and subsequent prostate cancer diagnosis and all-cause mortality in men who were aged 60–84 years in 1988. The model was programmed in GAUSS *(21)*; an in-depth description of the model logistics was reported by Etzioni et al. *(22).* Briefly, the model identified the cases of prostate cancer whose diagnosis was advanced by PSA screening; we focused on these cases because they account for all of the observed effects of PSA screening on disease incidence. For each case of prostate cancer that was detected through PSA screening, the model independently generated dates of other-cause death and of clinical diagnosis of prostate cancer, the latter of which was determined by adding the lead time to the date of screen detection. The date of clinical diagnosis is the date a case of prostate cancer would have been diagnosed in the absence of PSA testing, provided the patient did not die of other causes in the interim. The model estimated overdiagnosis as the proportion of case patients whose cancer was detected through PSA screening but who did not survive long enough to have their prostate cancer clinically diagnosed.

The overdiagnosis frequency estimated by the model is critically dependent on the lead time. To identify mean lead times that were consistent with the Surveillance, Epidemiology, and End Results (SEER)[1] incidence data, the model generated the expected prostate cancer incidence for several different mean lead times and then selected the one for which the expected incidence best matched the observed (SEER) incidence.

The expected incidence of prostate cancer generated by the model consists of the sum of two terms. The first term is the secular trend, which is the incidence that would have been expected in the absence of any PSA testing. This term was provided as an input into the model. The second term is the amount of incidence in excess of the secular trend that may be attributed to PSA testing. This excess incidence was produced as an output of the model as follows. Each individual whose prostate cancer was detected by PSA screening was considered a "diagnosis increment" in the year that a PSA test detected his cancer and a "diagnosis decrement" in the year that he would have been clinically diagnosed with prostate cancer in the absence of any PSA testing. For any given year, the excess incidence of prostate cancer was defined as the difference between the number of diagnosis increments and the number of diagnosis decrements in that year. Note that excess incidence is not synonymous with overdiagnosis; even if there were no overdiagnosis, the introduction of a sensitive screening test in a population would generally cause an initial increase in incidence.

Study Population Used in the Model

The model used a hypothetical population that consisted of two million men who were 60–84 years old in 1988, from which the cohort of screen-detected cases of prostate cancer arose. The age distribution of the study population in 1988 and the age-specific, all-cause mortality rates for that population were derived from census data *(23,24).* We used a sample size of two million men to provide a high degree of precision while preserving reasonable model run times on a personal computer. We chose 60 years as the lower limit of the age range of the study population in 1988 for two reasons. First, because data on PSA test utilization were available only for men aged 65 years and older, we were not comfortable extrapolating those rates of use to men who were younger than 60 years. Second, we used a lower age limit of 60 years rather than 65 years because we wanted to base our results on the cohort of men who were alive for the time period encompassed by our study (i.e., those aged 70–84 years). Thus, we wanted to maximize the inclusive cohort in terms of age.

Values Entered Into the Model

Testing rates. The model used PSA testing rates reported by Etzioni et al. *(5),* who updated through 1998 rates from a previous analysis by Legler et al. *(4),* to determine the number of individuals who were tested each year from 1988 through 1998. Annual PSA testing rates among men aged 65 years and older were obtained from a linkage between the SEER registry of the National Cancer Institute *(2)* and Medicare claims files from the Health Care Finance Authority *(25).* Claims data were available for all SEER-registered cases diagnosed with prostate cancer as well as for a random sample of men without prostate cancer who resided in the same SEER areas between 1988 and 1998 inclusive. The SEER–Medicare linkage allowed us to exclude men who had PSA tests after they were diagnosed with prostate cancer. Table 1 presents annual PSA test utilization rates by race, age, and calendar year of the test. PSA testing rates among men aged 60–84 years were assumed to be similar to those among men aged 65–84 years. This assumption is supported by a recent study that found no association between age and utilization of prostate cancer screening among men over the age of 50 *(26).*

Rates of screen-detected prostate cancer. The model used cancer detection rates derived from the SEER–Medicare linked database *(4,5)* to identify the cases whose prostate cancer was detected by screening (Table 1). The SEER–Medicare database contained prostate cancer diagnosis information through 1996 and therefore provided cancer detection rates through this time; we assumed that the cancer detection rates for 1997 and 1998 were the same as those for 1996.

The prostate cancer detection rate for a given year was defined as the number of men who were diagnosed with prostate cancer within 3 months after having a PSA test conducted in that year divided by the number of men who had at least one PSA test

Table 1. Annual rates of PSA testing among men aged 65 and older by race and calendar year and cancer detection rates following a PSA test by race, age group, and calendar year*

Calendar year	PSA testing rates†		Cancer detection rates‡							
	Whites	Blacks	Whites					Blacks		
	65–84 y	65–84 y	65–69 y	70–74 y	75–79 y	80–84 y	⩾85	65–74 y	75–84 y	⩾85 y
1988	1.20	0.88	3.74	7.69	8.00	10.53	3.70	20.00	28.57	28.57
1989	4.00	2.92	7.02	7.64	6.20	9.93	11.11	12.06	18.45	18.45
1990	8.20	5.99	4.21	6.17	7.10	7.31	8.54	4.12	8.33	10.00
1991	19.41	14.40	4.12	4.85	5.30	4.97	5.12	6.83	7.19	7.45
1992	30.42	21.83	3.48	4.10	4.26	4.34	4.03	6.58	6.85	6.96
1993	35.94	26.40	2.62	2.77	2.84	2.85	3.00	5.37	5.91	6.29
1994	38.03	28.48	2.07	2.24	2.04	2.02	2.01	4.50	4.27	3.59
1995	39.03	30.10	1.86	1.90	1.85	1.78	1.62	3.63	3.02	2.40
1996	39.80	31.01	1.86	2.04	1.89	1.70	1.85	3.48	3.62	2.65
1997	38.90	31.88	1.86	2.04	1.89	1.70	1.85	3.48	3.62	2.65
1998	39.18	32.81	1.86	2.04	1.89	1.70	1.85	3.48	3.62	2.65

*PSA = prostate-specific antigen.

†Annual PSA testing rates are given by the percentage of men who were alive, who did not have a prostate cancer diagnosis at the start of the given year, and who had received at least one PSA test in that year.

‡The cancer detection rate is the percentage of men tested at least once in the given year who were diagnosed with prostate cancer within 90 days of a test conducted in that year. Cancer detection rates for black men are given by 10-year rather than 5-year age groups and were smoothed across calendar years because of small sample sizes and high variability across years in this group.

in that year. Because PSA test results were not available, all men who were diagnosed with prostate cancer within 3 months after having a PSA test were included in estimates of the cancer detection rate. We refer to these cases of prostate cancer as PSA-associated cases and describe below how the cancer detection rates were adjusted to exclude cases whose PSA tests were used to confirm their disease status in the presence of symptoms and who, therefore, were not *bona fide* screen-detected case patients. We assumed that the cancer detection rates for men aged 60–64 were similar to those for men aged 65–69.

Because the administrative claims data did not distinguish between screening tests and confirmatory diagnostic tests, we introduced a parameter, p, that denotes the proportion of PSA-associated cases whose prostate cancer was detected by screening rather than by clinical examination. For a given value for p, we derived adjusted cancer detection rates that excluded patients whose prostate cancer was clinically detected but who had had a PSA test to confirm their diagnosis *(22)*.

Obtaining an unbiased estimate of p is generally not possible without performing a full medical record review and, even with such a review, is extremely challenging. Therefore, we performed a sensitivity analysis to determine how the model results would vary across a range of values for p. We chose values for p that increased over time to reflect the increased use of the PSA test for screening. Those values reflected high, moderate, and low frequencies of early prostate cancer detection as follows: For a high frequency of detection, p was 0.7 in 1988 and increased to 0.9 in 1998; for an intermediate frequency of detection, p was 0.5 in 1988 and increased to 0.8 in 1998; and for a low frequency of detection, p was 0.3 in 1988 and increased to 0.7 in 1998. Fig. 1 shows the incidence of screen-detected prostate cancer implied by each of these values, as computed by the product of the annual PSA testing and cancer detection rates and adjusted according to the different values for p. Fig. 1 also shows the total PSA-associated incidence, which corresponded to a value of 1 for p and was estimated by the product of the annual PSA testing and cancer detection rates.

Lead time. Each screen-detected case of prostate cancer identified by the model had, by definition, a lead time greater than zero. The lead time was added to the date of screen detection to obtain the date at which a prostate cancer diagnosis would have occurred in the absence of PSA testing. We considered three values for mean lead time—3 years, 5 years, and 7 years—in accordance with prior estimates of this quantity *(9–11)*. The corresponding lead time distributions were gamma distributions with shape and scale parameters given by (3,1), (5,1), and (5,5/7) *(22)*.

Secular trend. The secular trend in cancer incidence is directly dependent on health-related behaviors and clinical practice patterns in the population. In the decade preceding the advent of PSA testing, the principal determinant of secular trend in prostate cancer was the frequency of transurethral resection of the prostate (TURP) for benign prostatic hyperplasia *(27,28)*. Fig. 2 shows that from 1973 through 1986, the overall incidence of prostate cancer almost exactly paralleled the incidence of TURP-detected prostate cancer.

Since the late 1980s, several reports have indicated that the frequency of TURP for the treatment of benign prostatic hyperplasia has declined dramatically *(12,16,29)*. These reports, together with surveys of the diagnosis and management of prostate cancer and benign prostatic hyperplasia *(13–14)* from the mid-1990s, indicate that the declines in referrals for TURP for benign prostatic hyperplasia occurred, in large part, because the surgical procedure was replaced by medical management of the condition through either androgen or alpha-adrenergic blockade *(15)*. Fig. 2 shows that the decline in the frequency of TURP among men aged 65 years and older translated directly into a decline in TURP-detected prostate cancer incidence from 1988 through 1993 *(27)*. The results of Wasson et al. *(17)* showed that declines in referrals for TURP among men in this same age group extended through 1997.

Given these observations, an intuitively reasonable projection of secular trend is one that would parallel the TURP-detected prostate cancer incidence from 1988 through 1998, as shown in

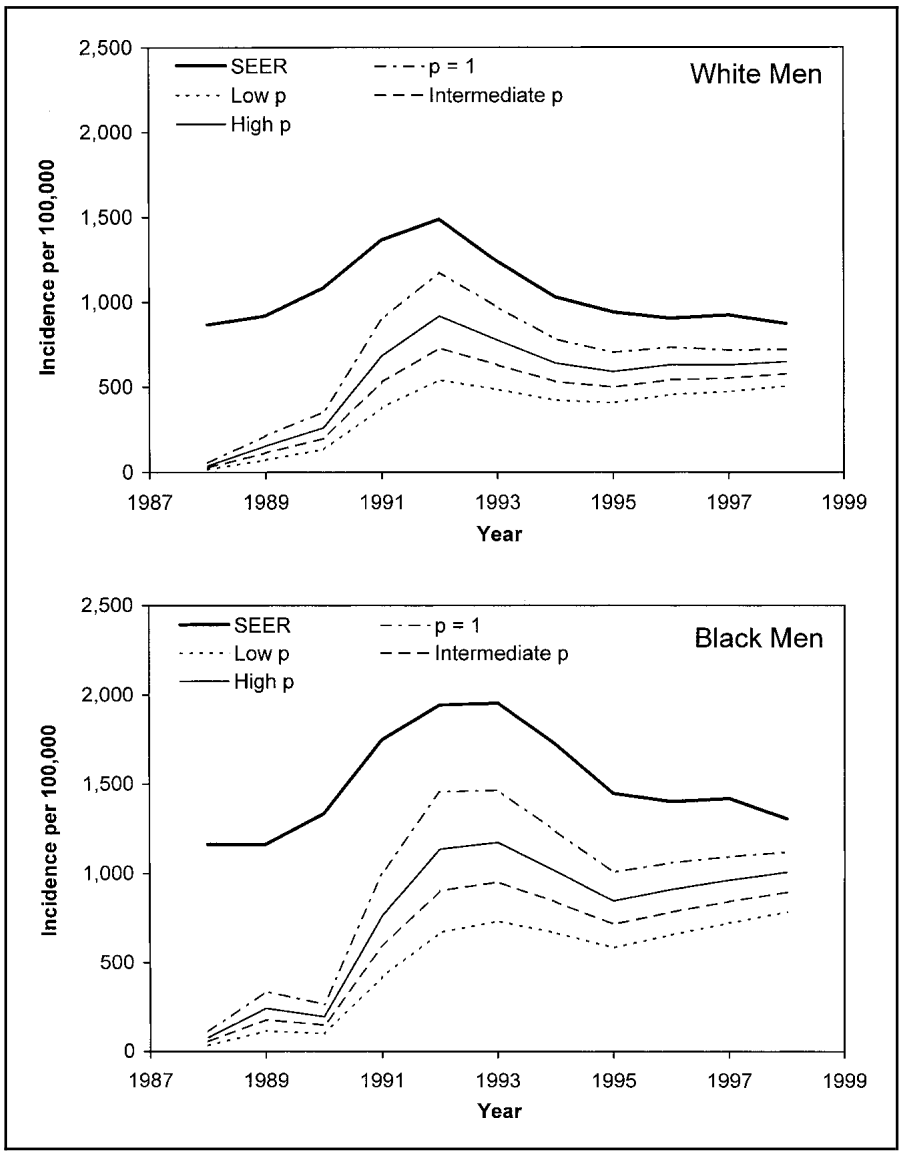

Fig. 1. Incidence of prostate cancer derived from SEER data per 100 000 for men who were alive for the study period and aged 65 and older in 1988 *(2),* with overall incidence detected by prostate-specific antigen (PSA) testing ($p = 1$) and approximations of screen-detected incidence for three assumptions about p, the proportion of screen-detected cases among all PSA-detected cases. High $p = 0.7$ in 1988, increasing to 0.9 in 1998; intermediate $p = 0.5$ in 1988, increasing to 0.8 in 1998; and low $p = 0.3$ in 1988, increasing to 0.7 in 1998.

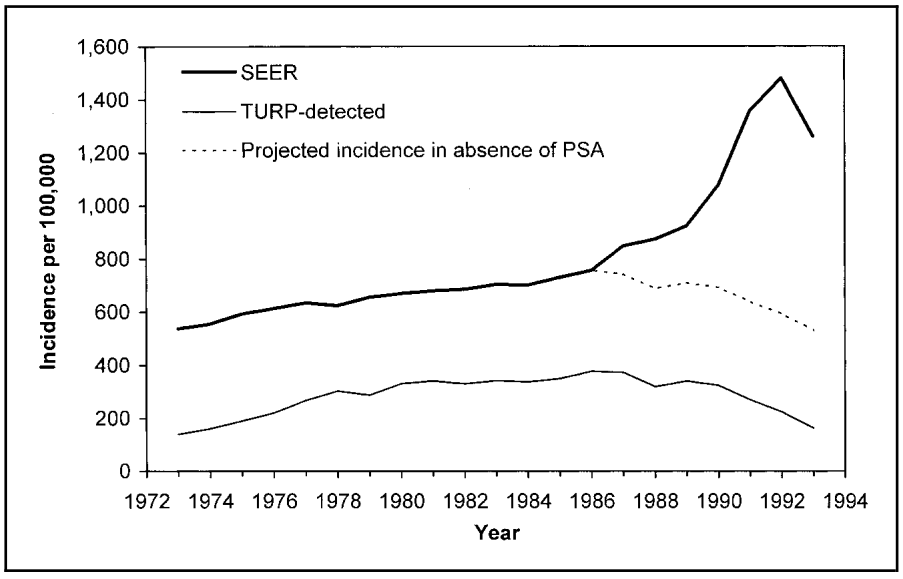

Fig. 2. Incidence of prostate cancer derived from SEER data. SEER incidence per 100 000 for men who were alive for the study period and aged 65 and older with incidence of prostate cancer detected by transurethral resection of the prostate (TURP) through 1993 from Merrill et al. *(24).* The projected declining secular trend is represented by the **dotted line.**

Fig. 2. However, this projection would assume that the declining incidence of TURP was independent of the increasing utilization of PSA testing, which may not have been the case *(16,30)*. In particular, patients who would have been surgically treated for their benign prostatic hyperplasia in the past are now frequently undergoing PSA screening as part of the diagnostic process *(30)*. In cases such as these, PSA testing may have superceded TURP as the mode of prostate cancer detection. To accommodate this lack of independence between the declining incidence of TURP and the increasing utilization of PSA testing, we present baseline results under a secular trend that balances these two trends and is constant after 1988. The message behind this constant secular trend is that, even in the absence of PSA testing, the increase in prostate cancer incidence observed prior to 1988 would probably not have been sustained, but it also would not have declined nearly as precipitously as suggested by the declines in TURP-detected prostate cancer incidence. In the sensitivity analysis, we also considered the declining secular trend as well as an increasing secular trend that continued the trend in SEER incidence that was observed prior to 1988. Fig. 3 illustrates the three secular trends we used in the model.

RESULTS

Baseline Conditions

Fig. 4 presents plots of prostate cancer incidence for the modeled population under baseline conditions, that is, for a population with a moderate use of prostate cancer screening (intermediate p) and a constant secular trend. The plots pertain to men aged 70–84 years because the men in this age group were alive for the entire study period *(1)*. Under baseline conditions, model-projected prostate cancer incidence rates corresponding to mean lead times of 5 years and 7 years were most consistent with the observed prostate cancer incidence rates from SEER data for white and black men, respectively (Fig. 4). The prostate cancer overdiagnosis rates associated with these mean lead times were 28.8% for white men and 43.8% for black men (Table 2).

Sensitivity Analysis

Our sensitivity analysis examined several secular trends in incidence as well as different settings for the relative frequency of screen-detected versus clinically-detected cases associated

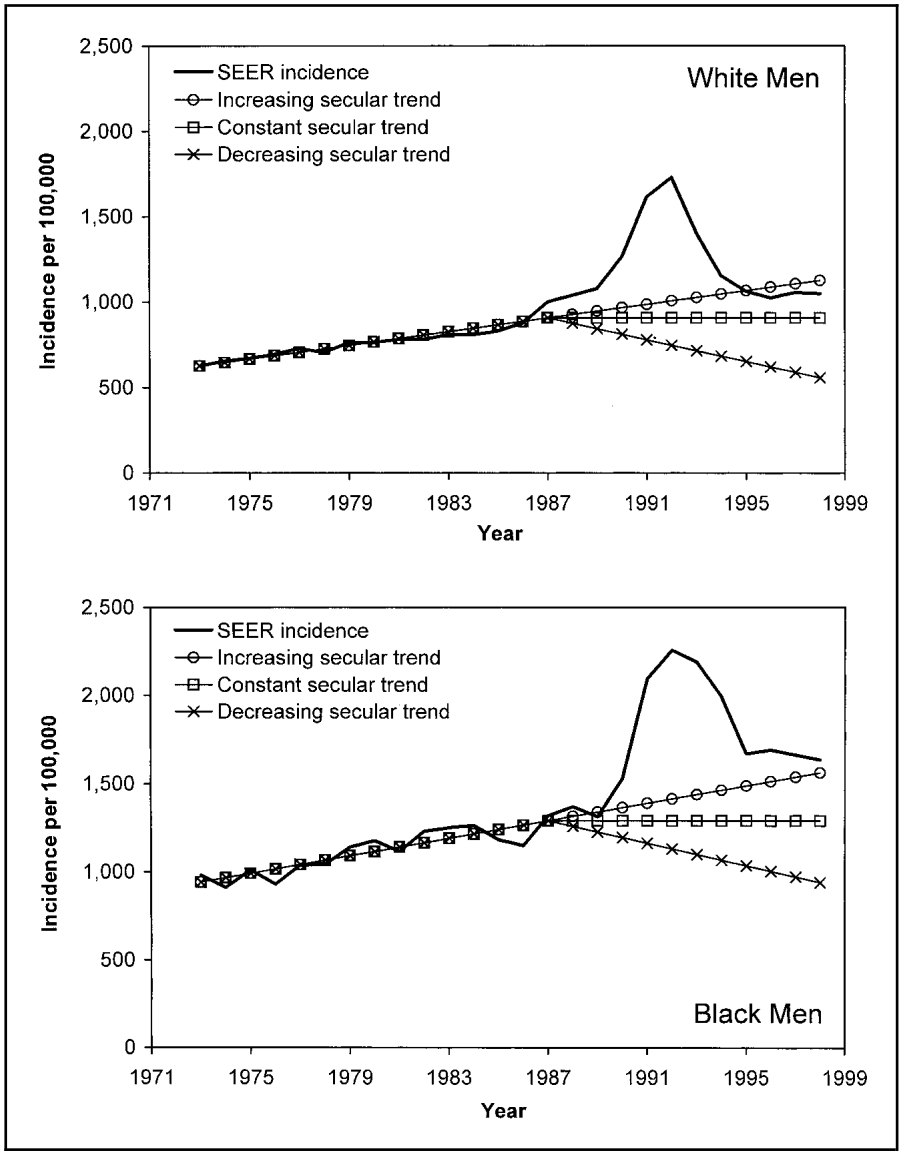

Fig. 3. Increasing, constant, and decreasing secular trends together with SEER incidence for white and black men aged 70–84 years. The declining secular trend decreases by 31.84 cases per 100 000 per year for both whites and blacks; the increasing secular trends increase by 20.07 cases per 100 000 per year for whites and by 24.81 cases per 100 000 per year for blacks.

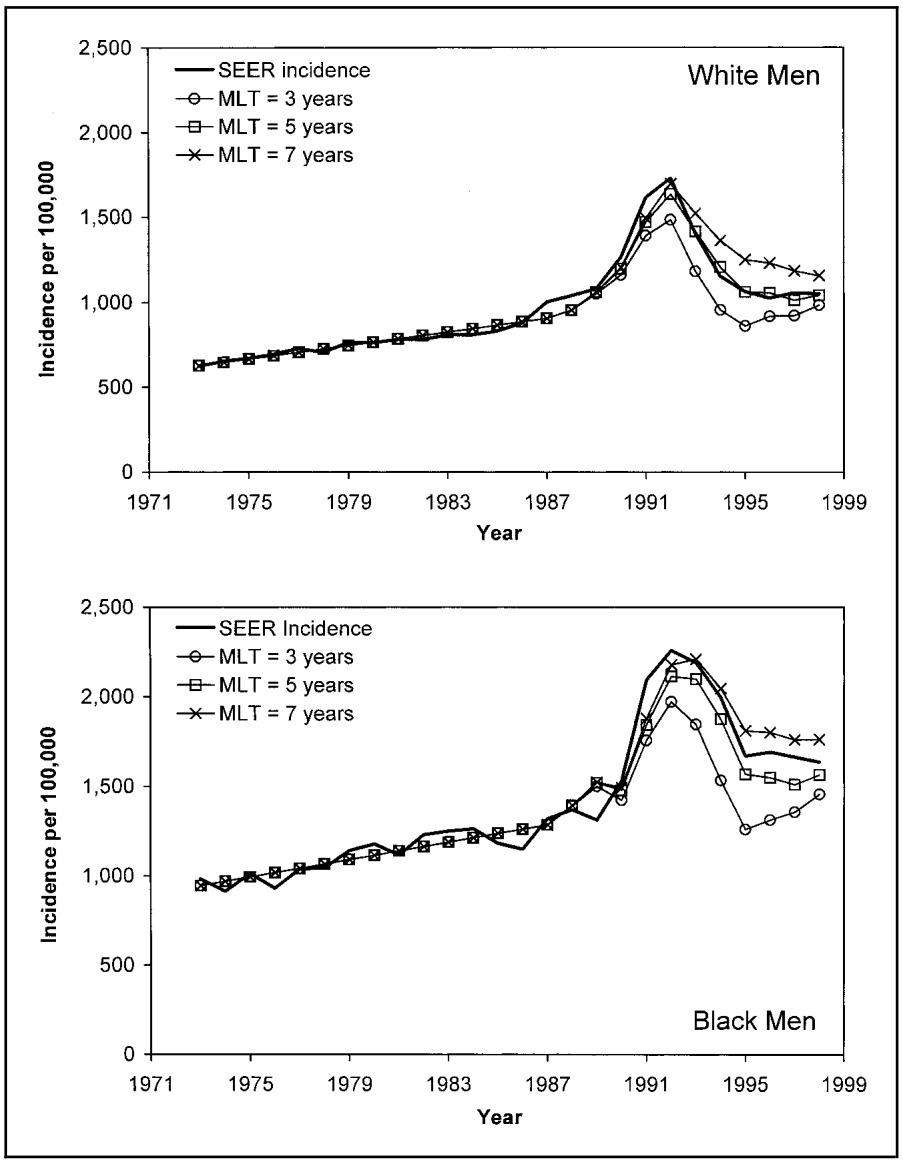

Fig. 4. Observed and model-projected prostate cancer incidence for men aged 70–84 under baseline conditions (constant secular trend and intermediate *p*). In each plot, the **bold curve** is the observed incidence from SEER data *(2)*. The remaining three curves represent the model-projected incidence; each corresponds to a specific mean lead time (MLT).

Table 2. Projected rates of overdiagnosis corresponding to each of three mean lead times and values for *p* considered in the model

Mean lead time, y	*p**	Overdiagnosis rates, %†	
		Whites	Blacks
3	Low	17.96	20.12
3	Intermediate	17.65	20.29
3	High	17.38	20.09
5	Low	29.33	32.36
5	Intermediate	28.77	32.61
5	High	28.59	32.31
7	Low	40.06	43.48
7	Intermediate	39.45	43.83
7	High	39.24	43.66

*p represents the proportion of prostate-specific antigen (PSA) test-associated diagnoses that were screen-detected.

†The overdiagnosis rates represent the proportion of screen-detected case patients whose death by causes other than prostate cancer would have preceded their date of clinical diagnosis and pertain to the modeled population, namely men aged 60–84 in 1988, with prostate cancer detected through PSA screening between 1988 and 1998 inclusive.

with PSA testing (*p*). Including the baseline analysis, we performed 27 different model runs for each racial group (i.e., one for each combination of secular trend, relative frequency of screen-detection [*p*], and mean lead time).

For a specific value of the mean lead time, we found that the estimated overdiagnosis rates were unchanged across the range of values for *p.* This finding is intuitively reasonable, given that overdiagnosis was expressed as a proportion of the screen-detected cancers. Simply changing the proportion of screen-detected cancers did not affect how frequently screen-detected cases were overdiagnosed. The relative frequency of screen detection did, however, affect how well the model-projected incidence of prostate cancer matched the observed (SEER) incidence. Under a constant secular trend, for example, the model results for low *p* did not match the observed data well. For high *p,* results for whites were similar to the baseline results, but a mean lead time of 5 years became the best-fitting projection for blacks, with a corresponding overdiagnosis frequency of 32.2% (Table 2).

The assumed choice of secular trend strongly influenced

which combination of lead times and overdiagnosis rates was most consistent with the observed incidence of prostate cancer obtained from SEER data. Under a declining secular trend, a mean lead time of 7 years for both whites and blacks was most consistent with the observed incidence of prostate cancer (Fig. 5). Under an increasing secular trend, a mean lead time of approximately 3 years for whites and 5 years for blacks yielded model-projected prostate cancer incidence rates that were very close to the observed incidence rates (data not shown). The corresponding overdiagnosis rates in this latter case were 17.7% for whites and 20.3% for blacks (Table 2).

Discussion

There is no doubt that PSA testing is the driving force behind the fluctuations in prostate cancer incidence that have been observed in the past decade. In this study, we used data from a representative survey of cancer incidence in the United States to identify the rates of prostate cancer overdiagnosis due to PSA testing that could be inferred from incidence patterns in the decade following its introduction. Under baseline conditions that reflected reasonable assumptions about the rates of screen-detected versus clinically-detected prostate cancer and secular trends in prostate cancer incidence, the estimated rates of overdiagnosis for men who were 60–84 years old in 1988—approximately 29% for whites and 44% for blacks—were consistent with the observed data in spite of sharp declines in prostate cancer incidence after 1992 to almost pre-1988 levels.

The fact that the best-fitting lead time (and corresponding overdiagnosis rate) for blacks (7 years) was greater than that for whites (5 years) may seem to contradict prior evidence suggesting that prostate cancer tends to be a more aggressive disease in black men than in white men *(1,31)*. However, most of the evidence about relative disease aggressiveness pertains to patients whose prostate cancers were diagnosed clinically, in the absence of PSA testing. By contrast, the lead times identified by our model are among screen-detected cases, only a portion of which would have been diagnosed clinically in the absence of PSA testing. However, even if these clinically detected cases are more aggressive in blacks than in whites, the same is not necessarily true of the screen-detected cases. For instance, because of the phenomenon of length bias, whereby cases with longer disease natural histories tend to be the ones detected by screening, the most aggressive cases may not even be present in the screen-detected cohort. Note that, even under similar mean lead

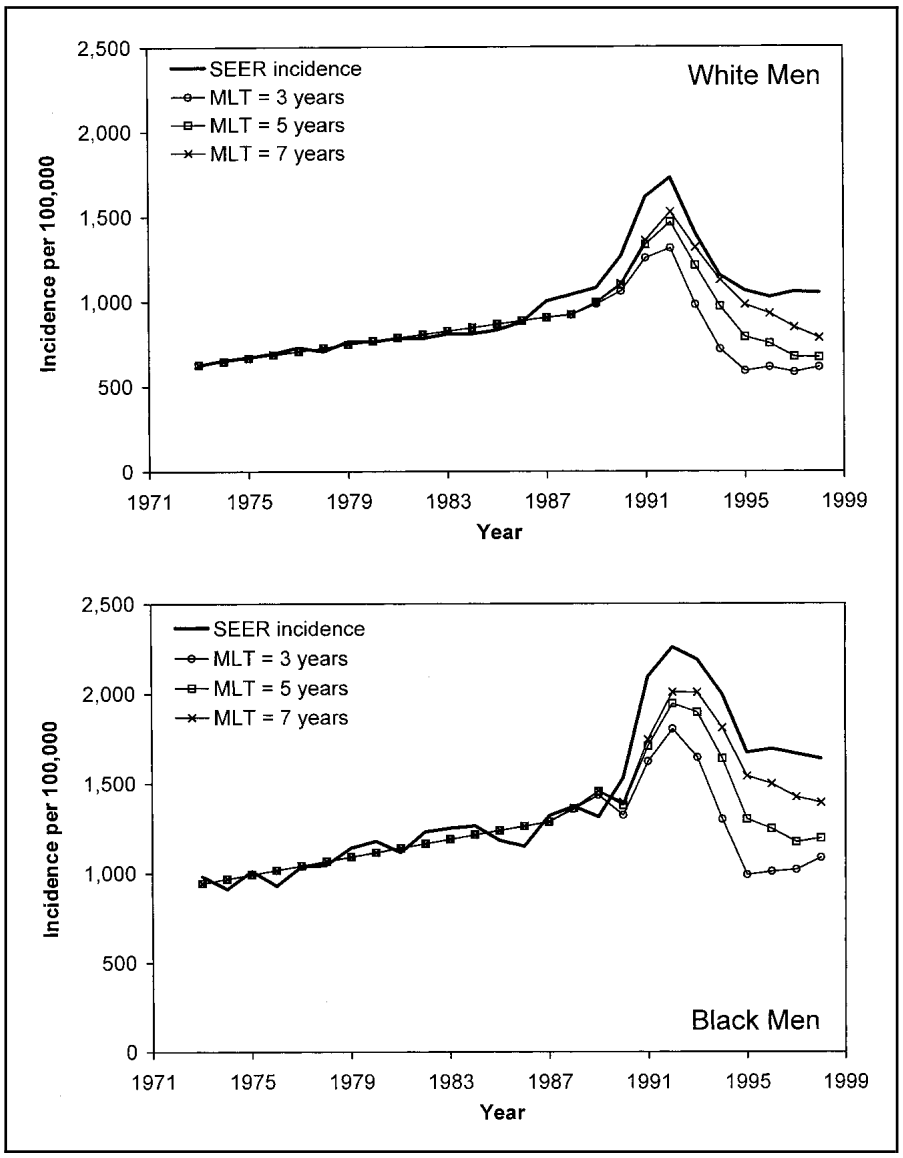

Fig. 5. Observed and model-projected prostate cancer incidence for men aged 70–84 under a declining secular trend. In each plot, the **bold curve** is the observed, delay-adjusted incidence from SEER data *(2)*. The remaining three curves represent the model-projected incidence; each corresponds to a specific mean lead time (MLT).

times for blacks and whites, the model projected higher overdiagnosis rates for blacks than for whites, probably because compared with whites, blacks have higher all-cause mortality rates and a distribution of age at PSA testing that is skewed toward higher ages *(5).*

It is important to distinguish between our use of the term *overdiagnosis* and other stated interpretations of this term. We have defined the overdiagnosis rate as the fraction of men whose prostate cancers were detected by PSA testing and who otherwise would not have been clinically diagnosed with prostate cancer in their lifetimes. The rationale for using this definition was to recognize the morbidity that results from a prostate cancer diagnosis, so that any diagnosis that would not have occurred in the absence of PSA testing would be considered a liability of PSA screening. As a comparison, McGregor and colleagues *(32)* defined overdiagnosis as the fraction of men whose prostate cancers were detected by screening who did not have their lives extended by screening. Their overdiagnosis rate includes some cases detected by PSA testing that would have been diagnosed clinically and, consequently, it may be substantially higher than our overdiagnosis estimate. The definition of overdiagnosis used by McGregor et al. *(32)* is relevant if the lifetime morbidity following an early diagnosis of prostate cancer is measurably greater than the lifetime morbidity following a later diagnosis. Morbidity following diagnosis is a potential issue in prostate cancer control, given the frequent occurrence of irreversible complications that can measurably affect quality of life following treatment for the disease *(18).*

Although our projected overdiagnosis rates for prostate cancer are nontrivial, they are far lower than the estimates that arise when comparing prostate cancer incidence in a cohort undergoing screening with that in an unscreened control group, as reported by Zappa et al. *(33).* We contend that such studies cannot provide a clinically meaningful estimate of the rate of overdiagnosis, because large increases in incidence are to be expected when a fairly sensitive screen, such as PSA testing, is introduced and because the relative increase in incidence cannot be interpreted without having an estimate of the lead time.

Our projected overdiagnosis rates are consistent with the views of Gann *(7),* who commented that the decline in incidence rates following the peak seen in the early 1990s "fits with the view that PSA does not reach so deeply into the preclinical pool so as to detect the huge reservoir of trivial, indolent tumors that can be seen on autopsy." Using the model-projected overdiagnosis rates presented herein, as well as results from Etzioni et al. *(34),* we can now quantify just how far PSA testing reaches into this reservoir.

Etzioni et al. *(34)* have estimated, based on historical autopsy data *(35),* that the lifetime probability (up to age 90) of autopsy-detectable prostate cancer is approximately 36% for white men and 28% for black men. However, just prior to the advent of PSA testing, the lifetime probability of a clinical prostate cancer diagnosis was only approximately 9% for both whites and blacks *(36).* This probability implies that Gann's "huge reservoir" amounts to a lifetime probability of latent and undiagnosed disease in the pre-PSA testing era of 27% in whites and 19% in blacks. Now, in the era of PSA testing, suppose that screening detects all (100%) future clinical cases. If we apply our estimates of the frequencies of overdiagnosis among screen-detected cases for whites (29%) and blacks (44%), we calculate that over their lifetimes, approximately 4% (29% × 9%/[100% – 29%]) of whites and 7% (44% × 9%/[100% – 44%]) of blacks will be screen-detected and overdiagnosed. Thus, at most, 15% (4%/27%) and 37% (7%/19%) of latent tumors present at death in whites and blacks, respectively, will be detected by PSA screening. These figures are upper boundaries because they assume that all future clinical diagnoses would be detected early by PSA screening. In the calendar period considered in our study, the proportion of autopsy-only tumors that were detected by PSA screening is likely to be far lower than the 15% and 37% estimates because not all men underwent testing and, among those who did, most were not being tested regularly *(5).*

There are several advantages to using SEER–Medicare data as a resource for statistics regarding the use of PSA testing. First, these data represent a broad segment of the U.S. population, namely the areas covered by the SEER registry. The second is the fact that medical claims are generally not subject to the types of biases that can arise when one relies on survey data concerning screening behavior *(37).* Although procedure codes for PSA screening were added to the Medicare data only in the latter part of the calendar period studied, codes for PSA diagnostic testing were available for the duration of that period, and we assume that the vast majority of PSA screens, in addition to those tests conducted for diagnostic confirmation of disease status, were captured by these codes.

The administrative claims data used herein also have several limitations. First, the data are restricted to older men. However, it is difficult to find reliable population-based data that provide similarly complete information on testing histories, particularly for younger men, over the time period of interest. Because the likelihood of overdiagnosis is dependent on age, it is important to note that our results pertain to the age group studied here and not to younger men. A second limitation is the lack of information on the reasons for PSA testing, which makes it impossible to distinguish between screen-detected and clinically detected cases of prostate cancer. This problem, which exists in practically all retrospective analyses of PSA testing utilization *(38),* severely complicates attempts to draw inferences about the effects of PSA screening on outcomes of interest. Despite the lack of published information on the relative frequency of PSA screening tests versus PSA diagnostic tests, it seems reasonable to assume that the relative frequency of screening tests has increased over time and that our analyses incorporating the parameter p reflect this assumption. In addition to the linear trends in p reported in the results, we also considered exponential increases in p over time and obtained similar results.

The computer model presented here does not represent a formal statistical approach to the problem of estimating lead time and overdiagnosis from cancer screening data. Such an approach has been developed in the context of cancer screening trials, where screening and incidence data are available at the level of the individual *(39).* Indeed, from a statistical point of view, the approach presented here is exploratory in the sense that it considers a small subset of possible lead-time distributions; the subset is based on published evidence concerning mean lead times. A more formal analysis would develop a likelihood function for the observed data and identify the best-fitting lead-time distribution through a formal optimization algorithm. It is not clear that the population data used here are amenable to such an approach, but this topic deserves further study.

This study provides the first quantitative analysis of the evidence concerning prostate cancer overdiagnosis due to PSA

screening from population data on prostate cancer incidence. We have shown that those data are consistent with a sizeable probability of overdiagnosis among screen-detected cases of prostate cancer. However, we found that the majority of cases of prostate cancer detected by screening in the population would still have presented clinically within the lifetime of the patient. This finding is consistent with results from clinical studies *(40,41)* of the histopathologic characteristics of PSA-detected prostate tumors, which show that these tumors appear to be clinically significant, and has important policy implications for PSA screening. However, this finding does not provide any information about the potential impact of PSA screening on survival or about the potential cost-benefit tradeoffs associated with the test. Although an investigation of these issues is beyond the scope of the present article, they have been explored elsewhere using similar computer modeling approaches *(42–45)*. Ongoing randomized trials will provide important evidence concerning the effects of PSA screening on survival, but computer models can provide useful insights while we await these results.

REFERENCES

(1) Stanford JL, Stephenson RA, Coyle LM, Cerhan J, Correa R, Eley JW, et al. Prostate cancer trends 1973–1995, Seer Program, National Cancer Institute. NIH Publ No. 99-4543. Bethesda (MD): 1999.

(2) Ries LA, Eisner MP, Kosary CL, Hankey BF, Miller BA, Clegg L, et al., editors. SEER cancer statistics review, 1973–1999, National Cancer Institute. Bethesda (MD): 2002. [Last accessed: 05/22/02.] Available at: http://seer.cancer.gov/csr/1973_1999.

(3) Kim HJ, Fay MP, Feuer EJ, Midthune DN. Permutation tests for joinpoint regression with applications to cancer rates. Stat Med 2000;19:335–51.

(4) Legler J, Feuer E, Potosky A, Merrill R, Kramer B. The role of prostate-specific antigen (PSA) testing patterns in the recent prostate cancer incidence decline in the USA. Cancer Causes Control 1998;9:519–57.

(5) Etzioni R, Berry KM, Legler JM. PSA testing in the US population: an analysis of Medicare claims from 1991–1998. Urology 2002;59:251–5.

(6) Feuer EJ, Wun LM. How much of the recent rise in breast cancer incidence can be explained by increases in mammography utilization: a dynamic population approach. Am J Epidemiol 1992;136:1423–36.

(7) Gann PH. Interpreting recent trends in prostate cancer incidence and mortality. Epidemiology 1997;8:117–20.

(8) National Center for Chronic Disease Prevention and Health Promotion. Behavioral Risk Factor Surveillance System Home Page. [Last accessed: 05/28/02.] Available at: http://www.cdc.gov/brfss.

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

(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) Hugosson J, Aus G, Becker C, Carlsson S, Eriksson H, Lilja H, et al. Would prostate cancer detected by screening with prostate-specific antigen develop into clinical cancer if left undiagnosed? BJU Int 2000;85:1078–84.

(12) Breslin DS, Muecke EC, Reckler JM, Fracchia JA. Changing trends in the management of prostatic disease in a single private practice: a 5-year followup. J Urol 1993;150:347–50.

(13) Gee WF, Holtgrewe L, Blute L, Miles ML, Naslund BJ, Nellans MJ, et al. 1997 American Urological Association Gallup Survey: changes in diagnosis and management of prostate cancer and benign prostatic hyperplasia, and other practice trends from 1994 to 1997. J Urol 1998;160:1804–7.

(14) Barry MJ, Fowler FJ Jr, Lin B, Oesterling JE. A nationwide survey of practicing urologists: current management of benign prostatic hyperplasia and clinically localized prostate cancer. J Urol 1997;158:488–92.

(15) Beduschi MC, Beduschi R, Oesterling JE. Alpha-blockade therapy for benign prostatic hyperplasia: from a nonselective to a more selective alpha1A-adrenergic antagonist. Urology 1998;51:861–72.

(16) Collins MM, Barry MJ, Bin L, Roberts RG, Oesterling JE, Fowler FJ. Diagnosis and treatment of benign prostatic hyperplasia. Practice patterns of primary care physicians. J Gen Intern Med 1997;12:224–9.

(17) Wasson JH, Bubolz TA, Lu-Yao GL, Walker-Corkery E, Hammond CS, Barry MJ. Transurethral resection of the prostate among Medicare beneficiaries: 1984–1997. J Urol 2000;164:1212–5.

(18) Stanford JL, Feng Z, Hamilton AS, Gilliland FD, Stephenson RA, Eley JW, et al. Urinary and sexual function after radical prostatectomy for clinically localized prostate cancer: the Prostate Cancer Outcomes Study. JAMA 2000;283:354–60.

(19) 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.

(20) Beemsterboer PM, de Koning HJ, Kranse R, Trienekens PH, van der Maas PJ, Schroder FH. Prostate specific antigen testing and digital rectal examination before and during a randomized trial of screening for prostate cancer: European randomized study of screening for prostate cancer, Rotterdam. J Urol 2000;164:1216–20.

(21) GAUSS mathematical and statistical system, version 3.2.18. Copyright 1994–1995. Maple Valley (WA): Aptech Systems Inc.; 1995.

(22) Etzioni R, Legler JM, Feuer EJ, Merrill RM, Cronin KA, Hankey BF. 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. J Natl Cancer Inst 1999;91:1033–9.

(23) National Center for Health Statistics. Vital statistics of the United States, 1992. Vol II, Sec 6, life tables. Washington (DC): Public Health Service; 1996.

(24) Centers for Disease Control and Prevention (CDC). CDC WONDER home page. [Last accessed: 05/22/02.] Available at: http://wonder.cdc.gov/.

(25) 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.

(26) Steele CB, Miller DS, Maylahn CM, Uhler RJ, Baker CT. Knowledge, attitudes and screening practices among older men regarding prostate cancer. Am J Public Health 2000;90:1595–600.

(27) Merrill RM, Feuer EJ, Warren JL, Schussler N, Stephenson RA. The role of transurethral resection of the prostate in population-based prostate cancer incidence rates. Am J Epidemiol 1999;150:848–60.

(28) Potosky AL, Kessler L, Gridley G, Brown CC, Horm JW. Rise in prostatic cancer incidence associated with increased use of trans-urethral resection. J Natl Cancer Inst 1990;82:1624–8.

(29) Endrizzi J, Optenberg S, Byers R, Thompson IM Jr. Disappearance of well-differentiated carcinoma of the prostate: effect of transurethral resection of the prostate, prostate-specific antigen and prostate biopsy. Urology 2001;57:733–6.

(30) Meigs JB, Barry MJ, Giovannucci E, Rimm EB, Stampfer MJ, Kawachi I. High rates of prostate-specific antigen testing in men with evidence of benign prostatic hyperplasia. Am J Med 1998;104:517–25.

(31) Powell IJ, Banerjee M, Novallo M, Sakr W, Grignon D, Wood DP, et al. Prostate cancer biochemical recurrence stage for stage is more frequent among African-American than white men with locally advanced but not organ-confined disease. Urology 2000;55:246–51.

(32) McGregor M, Hanley JA, Boivin JF, McLean RG. Screening for prostate cancer: estimating the magnitude of overdetection. CMAJ 1998;159:1368–72.

(33) Zappa M, Ciatto S, Bonardi R, Mazzotta A. Overdiagnosis of prostate carcinoma by screening: an estimate based on the results of the Florence Screening Pilot Study. Ann Oncol 1998;9:1297–300.

(34) Etzioni R, Cha R, Feuer EJ, Davidov O. Asymptomatic incidence and duration of prostate cancer. Am J Epidemiol 1998;148:775–85.

(35) Carter HB, Piantadosi S, Isaacs JT. Clinical evidence for and implications of the multistep development of prostate cancer. J Urol 1990;143:742–6.

(36) DEVCAN: probability of developing or dying of cancer software. [Last accessed: 5/28/02.] Available at: http://srab.cancer.gov/devcan.

(37) Jordan TR, Price JH, King KA, Masyk T, Bedell AW. The validity of male patients' self-reports regarding prostate cancer screening. Prev Med 1999;28:297–303.

(38) Kramer BS, Brown ML, Prorok PC, Potosky AL, Gohagan JK. Prostate cancer screening: what we know and what we need to know. Ann Intern Med 1993;119:914–21.

(39) Pinsky PF. Estimation and prediction for cancer screening models using deconvolution and smoothing. Biometrics 2001;57:389–95.
(40) Humphrey PA, Keetch DW, Smith DS, Shepherd DL, Catalona WJ. Prospective characterization of pathological features of prostatic carcinomas detected via serum prostate specific antigen based screening. J Urol 1996; 155:816–20.
(41) Schwartz KL, Grignon DJ, Sakr WA, Wood DP Jr. Prostate cancer histologic trends in the metropolitan Detroit area, 1982 to 1986. Urology 1999; 53:769–74.
(42) Krahn MD, Mahoney JE, Eckman MH, Trachtenberg J, Pauker SG, Detsky AS. Screening for prostate cancer. A decision analytic view. JAMA 1994; 272:773–80.
(43) Barry MJ, Fleming C, Coley CM, Wasson JH, Fahs MC, Oesterling JE. Should Medicare provide reimbursement for prostate-specific antigen testing for early detection of prostate cancer? Part IV: estimating the risks and benefits of an early detection program. Urology 1995;46:445–61.
(44) Etzioni R, Cha R, Cowen ME. Serial prostate specific antigen screening for prostate cancer: a computer model evaluates competing strategies. J Urol 1999;162(3 Pt 1):741–8.
(45) Ross KS, Carter HB, Pearson JD, Guess HA. Comparative efficiency of prostate-specific antigen screening strategies for prostate cancer detection. JAMA 2000;284:1399–405.

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 without personal identifiers to the NCI on a biannual basis, and the NCI makes the data available to the public for scientific research.

Supported by the Cancer Intervention and Surveillance Modeling Network (CISNET) and by Public Health Service grants U01CA88160 and R29CA70227 from the National Cancer Institute, National Institutes of Health, Department of Health and Human Services (to R. Etzioni and D. di Tommaso). R. Boer's work was conducted at the Department of Public Health, Erasmus University, Rotterdam, The Netherlands.

We thank Dennis Fryback, Paul Pinsky, and Scott Ramsey for helpful comments on an earlier draft of this manuscript and Lauren Clarke for development of the model Web site at http://www.fhcrc.org/labs/etzioni/podx.html. This interface allows individuals to run the model with their own settings for key input parameters.

Manuscript received November 19, 2001; revised April 25, 2002; accepted May 15, 2002.