Impact of Age, Benign Prostatic Hyperplasia, and Cancer on Prostate-Specific Antigen Level

Rinaa S. Punglia, M.D., M.P.H.[1]
Anthony V. D'Amico, M.D., Ph.D.[1]
William J. Catalona, M.D.[2,3]
Kimberly A. Roehl, M.P.H.[2]
Karen M. Kuntz, Sc.D.[4]

[1] Department of Radiation Oncology, Brigham and Women's Hospital and the Dana-Farber Cancer Institute, Boston, Massachusetts.

[2] Division of Urologic Surgery, Washington University School of Medicine, St. Louis, Missouri.

[3] Department of Urology, Northwestern University, Fineberg School of Medicine, Chicago, Illinois.

[4] Department of Health Policy and Management, Harvard School of Public Health, Boston, Massachusetts.

This screening study was supported in part by a grant from Beckman Coulter, Inc., and the Urological Research Foundation.

Presented in part at the American Society for Clinical Oncology, Orlando, Florida, May 18-21, 2005.

Address for reprints: Karen Kuntz, Sc.D., Harvard Center for Risk Analysis, Harvard School of Public Health, 718 Huntington Ave., Boston, MA 02115-5924; Fax: (617) 432-0190; E-mail: kmk@hsph.harvard.edu

Received 10 May 2005; revision received 21 September 2005; accepted 20 October 2005.

BACKGROUND. The distribution of prostate-specific antigen (PSA) values for men with or without prostate carcinoma are confounded because of verification bias. Correcting for verification bias, the means and variances of PSA values were estimated in specific clinical scenarios.

METHODS. Existing receiver operating characteristic (ROC) curves, adjusted for the presence of verification bias in a screening population, were used to estimate the mean and variance of PSA values for men with or without prostate carcinoma, stratified by age and the presence or absence of benign prostatic hyperplasia. Men with a suspicious digital rectal exam (nodular) were excluded from analysis.

RESULTS. Among men with cancer and the absence of benign prostatic hyperplasia, mean PSA values were 2.05 ng/mL and 2.66 ng/mL for younger (<60 yr) and older (≥60 yrs) men, respectively. These estimates were 2.56 ng/mL and 3.90 ng/mL in the presence of benign prostatic hyperplasia for younger and older men, respectively. For men without prostate carcinoma, these values were 0.78 ng/mL and 1.23 ng/mL for younger and older men, respectively, among those without benign prostatic hyperplasia, and 0.97 ng/mL and 1.75 ng/mL for younger and older men, respectively, among those with benign prostatic hyperplasia.

CONCLUSIONS. Accurate estimates of the mean and variance of PSA values for men with or without cancer may provide PSA thresholds for biopsy that are specific for age and prostate size as assessed by digital rectal exam. Therefore, the current threshold of 4.0 ng/mL should not be applied indiscriminately. ***Cancer*** **2006;106: 1507–13.** *© 2006 American Cancer Society.*

KEYWORDS: prostate-specific antigen, mass screening, bias (epidemiology), prostatic neoplasms.

For the prostate-specific antigen (PSA) test, the threshold for recommending a prostate biopsy has traditionally been PSA level greater than 4.0 ng/mL. However, recently published prospective data from the Prostate Cancer Prevention Trial (PCPT) have revealed a significant number of prostate carcinomas under this threshold.[1] Indeed, many have advocated that the threshold for biopsy should be lower.[2,3] Those who may benefit the most from such a reduction are healthy men of younger age whose life expectancy is long compared with the natural history of prostate carcinoma.[4,5] The utility of the PSA test is limited by the finding that it is a prostate marker and not a prostate-carcinoma-specific marker. Therefore, benign prostatic hyperplasia, a disease more commonly found in older men, confounds the clinical utility of the serum PSA test, as it can raise the PSA nonspecifically. Because of this, the concept of 'PSA density,' which attempts to adjust PSA values to prostate volume, has been proposed.[6] The presence of benign prostatic hyperplasia can be detected by digital rectal examination, which reveals a prostate gland to be enlarged but not suspicious for cancer. Whereas the focus of advocating different PSA thresholds has been age-based,[7]

DOI 10.1002/cncr.21766
Published online 3 March 2006 in Wiley InterScience (www.interscience.wiley.com).

there is evidence that the presence or absence of benign prostatic hyperplasia may also be an important determinant of the optimal threshold for prostate biopsy.[8]

The receiver operating characteristic (ROC) curve is a plot of sensitivity versus one minus specificity for various thresholds for designating a test positive (e.g., PSA value $>$ 4.0 ng/mL). The shape of an ROC curve describing the PSA test depends on the distribution of PSA values among men with prostate carcinoma compared with the distribution of PSA values among men without prostate carcinoma. The degree to which these two distributions overlap describes the ability of the PSA test to discriminate between a patient with prostate carcinoma and one without it. The range of values in the overlap region of these distributions dictates the range of thresholds describing points along the ROC curve. Therefore, having unbiased estimates of the mean and standard deviation (SD) of each of these distributions stratified by age group and the presence or absence of benign prostatic hyperplasia could provide insight into the interpretation of PSA test values. However, determining accurate estimates of the distribution of PSA values in men with and without cancer cannot be directly performed in the presence of verification bias. Verification bias occurs when only a selected subset of screened men undergoes prostate biopsy and when the recommendation for biopsy is a function of PSA level and/or other clinical variables. If biopsy is more likely among men with higher PSA values, the directly estimated mean PSA level among the population of men with prostate carcinoma is artificially inflated.

Previously, we described estimates of ROC curves for PSA adjusted for verification bias[8] by using a mathematical method described by Begg and Greenes[9] to correct for this bias. We found that men who had an abnormal (enlarged but not suspicious) digital rectal exam result were significantly more likely to undergo biopsy compared with men with normal digital rectal exams after controlling for PSA level, age, race, and family history (OR = 1.547, P = .0002). This pattern was in contrast to the finding that among the biopsied subset, having an abnormal digital rectal exam was not a significant predictor of having cancer among these men after controlling for PSA and other covariates (OR = 0.691, P = .155). These findings provide evidence that some physicians may consider an abnormal (enlarged) digital rectal exam result as a risk factor for prostate carcinoma and not just a cause for a nonspecific rise in PSA level. By estimating unbiased distributions of PSA by cancer status, age, and digital rectal exam result, we aim to offer insights into the diagnostic characteristics and potential clinical utility of the PSA test.

MATERIALS AND METHODS

Patient Selection

The study population has been reported previously.[8] Briefly, between May 1995 and November 2001, 6,691 consecutive men were enrolled in a screening study at the Washington University School of Medicine and underwent PSA testing and digital rectal exam. To be enrolled, men had to be at least 50 years of age. If they had at least one family member with a history of prostate carcinoma or were African American, the minimum age was 40. Patients with PSA values greater than 2.5 ng/mL or findings suspicious for prostate carcinoma on digital rectal exam were recommended to undergo prostate biopsy. Any prostate carcinoma diagnosis up to 18 months of first PSA collection was considered evidence of disease to mitigate the effects of sampling error associated with prostate biopsy. Characteristics of the men who did and did not undergo prostate biopsy have been reported previously.[8] We used the PSA value from the initial enrollment visit to determine the sensitivity and specificity of the test for various cutoff values.

We performed logistic regression studying the probability of undergoing verification by prostate biopsy to study the influence of PSA level, age, family history, race, and digital rectal exam result. We also studied the influence of these variables on detection of prostate carcinoma in the subset of men who underwent biopsy.

ROC Estimation

Because of the continuous nature of the PSA test, thresholds were selected to allow for determination of discrete points on the ROC curve. PSA levels were categorized into ranges that allowed for adequate separation between the points on the ROC curve.

ROC curve analyses were performed by ROC Curve Analyzer software (developed by Centor and Keightley, University of Alabama, Birmingham, AL). Areas under the curve were calculated using the trapezoidal (nonparametric) method and compared using two-sided P values. An area under the curve of 1.0 describes a test with perfect discrimination between disease and no disease, whereas a test with an area under the curve of 0.5 has no discriminatory power.

Correction for Verification Bias

We used the method of Begg and Greenes[9] to correct for verification bias by adjusting for the verification process to estimate sensitivity and specificity in the entire population undergoing PSA testing and not just the subset that was verified by prostate biopsy. This method relies on the assumption that only the 'observable' clinical variables (e.g., PSA level, digital rectal exam result) af-

fected the likelihood of undergoing a biopsy and not the underlying or 'unobservable' cancer status.

To apply this method for correction of verification bias, we first estimated the probability of disease as a function of clinical variables using a logistic regression model from the sample of men who underwent verification of disease status by prostate biopsy. Variables included in the model were results of digital rectal exam, race (African American vs. other), family history, and PSA test result category. We used this regression model to predict the probability of prostate carcinoma in the entire group based on specified covariates (e.g., black race with normal digital rectal exam and no family history with a PSA between 4.1 and 6.0 ng/mL). We then summed these probabilities weighting by the frequency of the specified set of covariates in the entire population undergoing PSA testing to obtain adjusted sensitivity and specificity estimates for the PSA test. Results are presented separately for men below age 60 and those age 60 and above, further separated by normal and 'abnormal' digital rectal exam result. Abnormal digital rectal exam result was defined as synonymous with benign prostatic hyperplasia, i.e., an enlarged gland that is not suspicious for prostate carcinoma. Men with suspicious findings on digital rectal exam, such as a nodule, were not included in this analysis as they are likely to undergo prostate biopsy regardless of PSA level.

Estimation of Means and Standard Deviations

We assumed that the natural logarithm of PSA [ln(PSA)] is normally distributed in the population with cancer as well as in the population without cancer, stratified by age group (< 60; ≥ 60 yrs) and the presence of benign prostatic hyperplasia. A given set of values for the mean and SD for the men with cancer (μ_{ca}, σ_{ca}) and a set of values for the mean and SD for men without cancer (μ_{nc}, σ_{nc}) will produce a unique ROC curve. We used the true-positive fraction (sensitivity) plotted on the adjusted ROC curve to inform the mean and SD of ln(PSA) for the population with cancer. Specifically, we selected a mean ln(PSA) (μ_{ca}) and SD (σ_{ca}) such that the squared difference between the true-positive fraction (TPF) observed for each ln(PSA) cutpoint and that expected from the selection was minimized. For each ln(PSA) threshold *i*, we minimized the following:

$$\sum_i \left(TPF_{observed} - TPF\left(\mu_{ca}, \sigma_{ca}\right)_{expected}\right)^2$$

Where TPF represents the area to the right of the ln(PSA) threshold on the normal curve with mean (μ_{ca}) and SD (σ_{ca}):

TABLE 1
Number of Men in Each Age Group Stratified by Biopsy Status, PSA Level, and DRE Result

	Younger Men < 60 y				Older Men ≥ 60 y			
	Normal DRE		Abnormal DRE		Normal DRE		Abnormal DRE	
Biopsy	no	yes	no	yes	no	yes	no	yes
PSA 0–2.0	3552	96	316	10	1098	45	145	8
PSA 2.1–4.0	229	89	17	39	234	108	47	26
PSA 4.1–6.0	38	41	10	6	59	65	15	18
PSA ≥ 6.1	17	32	5	5	57	47	17	26

PSA: prostate-specific antigen; DRE: digital rectal exam.

$$\int_{\ln(PSA)}^{\infty} \frac{1}{\sigma_{ca}\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{\ln(PSA)-\mu_{ca}}{\sigma_{ca}}\right)^2}$$

In a similar manner, for the populations without cancer we used the false-positive fractions plotted on the adjusted ROC curve (1-specificity) to inform the mean and SD of ln(PSA) for the population without cancer, and then selected the mean ln(PSA) (μ_{nc}) and SD (σ_{nc}) that minimized the squared differences between the false-positive fractions observed and expected from the selection. All ln(PSA) thresholds on the adjusted ROC curve, except those yielding either 0 or 100% positive fractions, were used to describe the observed distribution.

We compared the mean and SD determined for each population from the adjusted ROC curves to those estimated directly from the verified (biased) sample. We used linear regression modeling of ln(PSA) versus age, digital rectal exam result, race, and disease status and their interactions to test whether there was a significant increase in ln(PSA) with abnormal digital rectal exam results, independent of disease status. In addition, we compared the derived means among men with and without cancer for each of the populations using a two-sample Student *t*-test with the derived SDs and the number of men who underwent biopsy in each group as the number of observations.

RESULTS

Unadjusted Analyses

Table 1 shows the numbers of men who underwent biopsy and who did not undergo biopsy at various PSA thresholds. By using only the biopsied sample, a 'naive' estimate of the mean and SD for those with and without cancer in each of the age groups and for each

TABLE 2
Mean PSA Values in Men with and without Prostate Carcinoma Separated by DRE Derived from Men Who Underwent Biopsy without Bias Adjustment

		Age < 60 y		Age ≥ 60 y	
	DRE	Mean ln · PSA (SD)	Geometric mean PSA ng/mL	Mean ln · PSA (SD)	Geometric mean PSA ng/mL
With cancer	Normal	1.40 (0.60)	4.06	1.55 (0.56)	4.71
	Abnormal*	1.05 (0.62)	2.86	1.76 (0.79)	5.81
Without cancer	Normal	0.77 (0.89)	2.16	1.15 (0.67)	3.16
	Abnormal*	1.08 (1.61)	2.94	1.43 (0.80)	4.18

DRE: digital rectal exam; PSA: prostate-specific antigen; SD: standard deviation.
* Abnormal DRE indicates enlarged prostate size but not suspicious for cancer.

TABLE 3
Mean PSA Values in Men with and without Prostate Carcinoma Separated by DRE Derived from Adjusted ROC Curves

		Age < 60 y		Age ≥ 60 y	
	DRE	Mean ln · PSA (SD)	Geometric mean PSA ng/mL	Mean ln · PSA (SD)	Geometric mean PSA ng/mL
With cancer	Normal	0.72 (0.68)	2.05	0.98 (0.85)	2.66
	Abnormal*	0.94 (0.65)	2.56	1.36 (0.86)	3.90
Without cancer	Normal	−0.25 (0.72)	0.78	0.21 (0.91)	1.23
	Abnormal*	−0.03 (0.84)	0.97	0.56 (1.03)	1.75

ROC: receiver operating characteristic; PSA: prostate-specific antigen; DRE: digital rectal exam; SD: standard deviation.
*Abnormal DRE indicates enlarged prostate size but not suspicious for cancer.

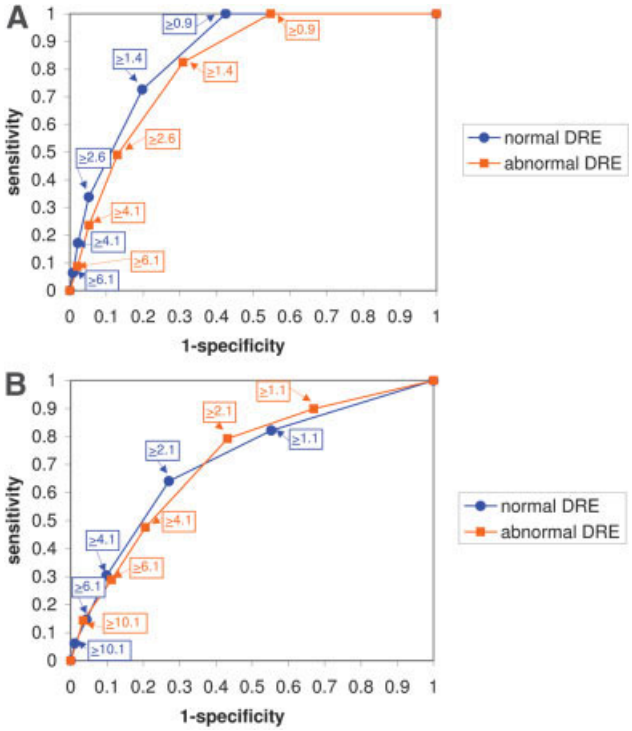

FIGURE 1. Adjusted ROC curves separating normal from abnormal digital rectal exam results. Shown are the ROC curves separating men with abnormal digital rectal exam results from normal digital rectal exam results, after correcting for verification bias. The overall shape of the ROC curve did not appear to change among the two digital rectal exam groups (A) in men younger than 60 years and (B) in men 60 years or older. However, the sensitivity and specificity at a given threshold to biopsy did change. For each threshold, the normal digital rectal exam group had a lower sensitivity and higher specificity compared with the abnormal digital rectal exam group. Reprinted with permission of the Massachusetts Medical Society from Punglia RS, D'Amico AV, Catalona WJ, Roehl KA, Kuntz KM. Effect of verification bias on screening for prostate cancer by measurement of prostate-specific antigen. *N Engl J Med.* 2003;349:335-342.

digital rectal exam result was calculated and is presented in Table 2.

Adjusted ROC Analyses

Figure 1 shows the adjusted ROC curves for those patients with normal digital rectal exam results and those with benign prostatic hyperplasia. Separating abnormal and normal digital rectal exam results did not change the resulting ROC curves significantly in the adjusted analyses, but the cutoff points between these groups did appear different, with increased sensitivity and decreased specificity for each PSA threshold in the abnormal exam group relative to the normal exam group, which could be explained by a shift in PSA levels in the men with abnormal digital rectal exams. From these adjusted curves, more accurate estimates of the mean and SDs of each group were determined, and distributions were plotted (Table 3; Fig. 2). As expected, using curves corrected for verification bias led to lower estimates of mean PSA relative to the unadjusted estimates.

For both age groups, mean ln(PSA) values among those with abnormal digital rectal exam results were greater than among those with normal digital rectal exam results in both the populations with prostate carcinoma and without prostate carcinoma (Table 3). By using a linear regression model, digital rectal exam result was found to be a significant predictor of ln(PSA) ($P = .0037$), independent of disease status, age, race, and family history. The lack of significance of the interaction term between disease status and digital rectal exam result ($P = .581$) further supports the hypothesis that both the men with cancer and the men without cancer have a similar shift in PSA level with an abnormal exam, resulting in ROC curves of similar shape but with altered cutoff points (Fig. 3).

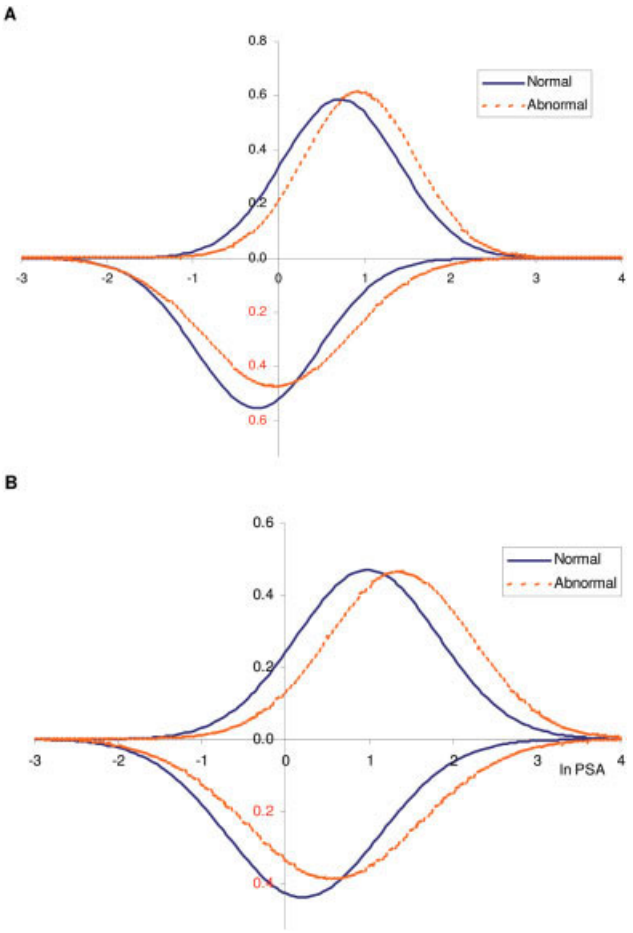

FIGURE 2. Distribution of ln(PSA) in men with and without prostate carcinoma derived from the adjusted ROC curves. Shown are the distribution of ln(PSA) values in men with prostate carcinoma (above the x-axis) and men without prostate carcinoma (below the y-axis) further separating men with abnormal digital rectal exam results from normal digital rectal exam results derived from the ROC curves of adjusted for verification bias in those (A) younger than 60 years and (B) those 60 years of age or older. These distributions would be expected to approach actual distributions of PSA within the population.

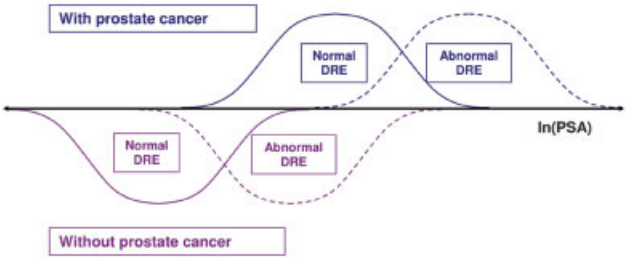

FIGURE 3. Schematic of the shift hypothesis with digital rectal exam result. Shown is a schematic distribution of frequency of ln(PSA) values among the men with prostate carcinoma (in blue) and those without (in purple) with normal (solid lines) and abnormal (dotted lines) digital rectal exam results. No single ln(PSA) value leads to complete discrimination between the cancer and non-cancer populations, and hence the imperfect nature of the PSA test. An abnormal digital rectal exam result leads to a shift of similar magnitude in ln(PSA) among both the group with cancer and that without cancer. Therefore, the relative distribution of cancer to no cancer test results within each of the digital rectal exam subgroups remains the same, leading to ROC curves of similar shape but with altered cutoff points.

Comparison of means using the number of men who underwent biopsy as the number of observations revealed that PSA values were significantly different between men with and without cancer in each of the four groups: younger men with normal digital rectal exam results ($P < .0001$), younger men with abnormal digital rectal exam results ($P = .015$), older men with normal digital rectal exam results ($P < .0001$), and older men with abnormal digital rectal exam results ($P = .0014$) (Table 4).

DISCUSSION

Correction for verification bias allows for a more accurate estimate of the performance characteristics of the PSA test in a screened population. Our adjusted analyses reveal no difference in PSA test performance among men with an enlarged prostate gland versus men with a normal-sized gland, and that there is a shift in the cutoff points along the ROC curve when separating by digital rectal exam result.

By using adjusted ROC curves to determine mean PSA values, we determined unbiased estimates of the distribution of PSA values within the population of men who have prostate carcinoma and those without cancer. In addition, we have shown that the derived mean PSA level is significantly different between men with and without cancer in each of the groups studied.

The geometric mean of the PSA levels derived from the adjusted ROC curves for men with prostate carcinoma ranged from 2.05-3.90 ng/mL, depending on age group and digital rectal exam status. Because the exponentiated mean of the ln(PSA) is mathematically equivalent to the geometric mean of PSA values, these numbers can be directly compared with the geometric mean among men with prostate carcinoma in prior published studies. One study found the geometric mean of PSA among men with prostate carcinoma to be between 6.28 and 7.46, much higher than our results.[7] This discrepancy further confirms the existence of strong selection bias in some prior studies, where men with higher PSA values were more likely to be diagnosed with disease.

Our unadjusted results (Table 2) also reveal the existence of verification bias, as mean PSA values among those men who underwent biopsy were higher

TABLE 4
Statistical Comparison of PSA Distributions Derived from the Adjusted ROC Curves

Age < 60 or ≥ 60 y	Cancer status	DRE Result	$P^{\dagger}$	Age < 60 or ≥ 60 y	Cancer Status	DRE Result
Younger	Cancer	Normal	< .00001	Younger	No cancer	Normal
Younger	Cancer	Abnormal*	.015	Younger	No cancer	Abnormal
Older	Cancer	Normal	< .00001	Older	No cancer	Normal
Older	Cancer	Abnormal*	0.0014	Older	No cancer	Abnormal

ROC: receiver operating characteristic; PSA: prostate-specific antigen; DRE: digital rectal examination.
*Abnormal DRE indicates enlarged prostate size but not suspicious for cancer.
†Comparison of columns 1–3 with columns 5–7.

than those among the entire screening population, and confirm the need for correction of the bias, as the numbers derived from the adjusted analysis more accurately describes the entire screened population. Although the unadjusted numbers for mean PSA values that included only those men who underwent prostate biopsy were higher than the better estimates gathered from the adjusted curves, they were not as high as those reported by Morgan et al.,[7] perhaps because of the lower threshold used to recommend prostate biopsy in this screening study. In addition to lower mean PSA values, the more accurate distributions reflect the heterogeneity of PSA values in the population as reflected by their greater variance than the corresponding distributions derived from the unadjusted analyses (Tables 2, 3).

Separating the analyses by digital rectal exam result (abnormal vs. normal) showed that the ROC curves had no difference in overall diagnostic performance, but had altered cutoff points. Such a model would be consistent with the schematic shown in Figure 3, where the relative distribution between the disease and nondisease populations does not change with an abnormal digital rectal exam, but instead only shifts to higher ln(PSA) levels, leading to higher cut points within the abnormal population for a given sensitivity/specificity, but the same overall ROC curve. This finding implies that the threshold value for recommending biopsy should be higher among men with abnormal digital exam results. This is especially important because we have shown that men with abnormal digital rectal exam findings are more likely to be referred for biopsy after controlling for PSA level. Although our study does provide better estimates of sensitivity and specificity at different PSA thresholds to biopsy as a function of age and digital rectal exam result, the ideal threshold for prostate biopsy depends on the relative tradeoff between false-positive and false-negative results. In addition, the power to detect small differences between men with enlarged versus normal-sized glands may be limited given the small number of men biopsied at certain PSA thresholds.

Our study was also limited by the use of prostate biopsy as the 'gold standard' for disease confirmation, which may have underestimated adjusted test sensitivity because of sampling error caused by the small amount of tissue removed at biopsy relative to the entire prostatic volume. To help mitigate this issue, we used all cancer diagnoses made within 18 months of PSA-based screening as our outcome. However, this method may have introduced some degree of selection bias, where more patients with rising PSA levels underwent late biopsies. Moreover, there may be additional variables (beyond age, PSA level, digital rectal exam result, race, and family history) not included in our analysis that may both predict for the chance of undergoing prostate biopsy and be related to underlying disease status, a problem with retrospective studies. The PCPT trial that recommended biopsy for all men enrolled prospectively studied PSA test characteristics.[1] Comparison of our previously published results[8] for older men to their cohort suggests that our procedure for adjusting for verification bias was indeed valid, as the two studies lead to comparable results among similar subgroups.

In conclusion, we have demonstrated a technique for finding the underlying mean and SD in the presence and absence of disease after accounting for verification bias, allowing for more accurate determination of these values even in a dataset with bias. The mean and SD of PSA levels among men with prostate carcinoma may have important implications regarding interpretation of PSA level for prostate carcinoma screening and biopsy recommendations. Our analysis may also inform modification of these recommendations for men with abnormal digital rectal exam results. The adjusted PSA distributions suggest that the current threshold of 4.0 ng/mL used for recommending prostate biopsy should not be applied indiscriminately.

REFERENCES

1. Thompson IM, Goodman PJ, Tangen CM, et al. The influence of finasteride on the development of prostate cancer. *N Engl J Med.* 2003;349:215-224.
2. Catalona WJ, Ramos CG, Carvalhal GF, Yan Y. Lowering PSA cutoffs to enhance detection of curable prostate cancer. *Urology.* 2000;55:791-795.
3. Catalona WJ, Smith DS, Ornstein DK. Prostate cancer detection in men with serum PSA concentrations of 2.6 to 4.0 ng/mL and benign prostate examination. Enhancement of specificity with free PSA measurements. *JAMA.* 1997;277:1452-1455.
4. Pound CR, Partin AW, Epstein JI, Walsh PC. Prostate-specific antigen after anatomic radical retropubic prostatectomy. Patterns of recurrence and cancer control. *Urol Clin North Am.* 1997;24:395-406.
5. Albertsen PC, Hanley JA, Gleason DF, Barry MJ. Competing risk analysis of men aged 55 to 74 years at diagnosis managed conservatively for clinically localized prostate cancer. *JAMA.* 1998;280:975-980.
6. Catalona WJ, Southwick PC, Slawin KM, et al. Comparison of percent free PSA, PSA density, and age-specific PSA cutoffs for prostate cancer detection and staging. *Urology.* 2000;56: 255-260.
7. Morgan TO, Jacobsen SJ, McCarthy WF, Jacobson DJ, McLeod DG, Moul JW. Age-specific reference ranges for prostate-specific antigen in black men. *N Engl J Med.* 1996; 335:304-310.
8. Punglia RS, D'Amico AV, Catalona WJ, Roehl KA, Kuntz KM. Effect of verification bias on screening for prostate cancer by measurement of prostate-specific antigen. *N Engl J Med.* 2003;349:335-342.
9. Begg CB, Greenes RA. Assessment of diagnostic tests when disease verification is subject to selection bias. *Biometrics.* 1983;39:207-215.