Annals of Epidemiology 26 (2016) 36–42

Contents lists available at ScienceDirect

Annals of Epidemiology

journal homepage: www.annalsofepidemiology.org

Original article

The use of multiphase nonlinear mixed models to define and quantify long-term changes in serum prostate-specific antigen: data from the Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial

Azza Shoaibi BPharm, MPH, PhD [a,b,*], Gowtham A. Rao MD, PhD, MPH [b,c], Bo Cai PhD [b], John Rawl MD [d], James R. Hébert MSPH, ScD [a,b,**]

[a] South Carolina Statewide Cancer Prevention and Control Program, Arnold School of Public Health, University of South Carolina, Columbia
[b] Department of Epidemiology and Biostatistics, Arnold School of Public Health, University of South Carolina, Columbia
[c] Department of Family and Preventive Medicine, School of Medicine, University of South Carolina, Columbia
[d] Columbia Urological Associates, P.A., Columbia, SC

ARTICLE INFO

Article history:
Received 19 May 2015
Accepted 4 October 2015
Available online 28 October 2015

Keywords:
High-risk prostate cancer
Mixed-effect model
PSA growth curves
PSA kinetics
PSA velocity
PSA change rate
Prostate cancer detection

ABSTRACT

Purpose: To test the hypothesis that the pattern of prostate-specific antigen (PSA) change in men diagnosed with high-risk prostate cancer (PrCA) differs from the pattern evident in men diagnosed with low-risk PrCA or those with no evidence of PrCA.
Methods: A retrospective cohort study from which PSA measures were taken before PrCA diagnosis from the Prostate, Lung, Colorectal and Ovarian Cancer Screening Trial. Data were fitted using a nonlinear regression model to estimate the adjusted absolute and relative (%) change of PSA.
Results: Data on 20,888 men with an average age of 61.61 years were included in the analysis. Of these, the 324 (1.55%) diagnosed with high-risk PrCA had a steeper and earlier transition into an exponential pattern of PSA change than the 1368 men diagnosed with low-risk cancer. At 1 year before diagnosis and/or exit, the average absolute PSA rates were 0.05 ng/mL/year (0.05–0.05), 0.59 (0.52–0.66), and 2.60 (2.11–3.09) for men with no evidence of PrCA, men with low-risk PrCA and those with high-risk PrCA, respectively.
Conclusions: The pattern of PSA change with time was significantly different for men who develop high-risk PrCA from those diagnosed with low-risk PrCA. Further research is required to validate this method and its utilization in PrCA screening.

Introduction

To improve the performance of prostate-specific antigen (PSA)–based screening for prostate cancer (PrCA), researchers have suggested using serial measures of PSA [1–3]. These measures would allow for computing parameters such as PSA kinetics and PSA velocity (PSAV), whose ability to improve PrCA detection has been the subject of debate [1,2,4–6]. Confusion exists because of multiple definitions and computation methods for PSAV [1], lack of a single threshold value for PSAV to predict PrCA, and changes in PSA associated with biological and biobehavioral characteristics, such as body weight, race, and age [7–12]. Evidence suggests that the pattern of PSA change over time differs between men with PrCA versus others [5,6], and that even among men with PrCA—PSA change over time may differ by disease aggressiveness [13]. These differences may be evident in the magnitude of PSA change over time (velocity) and in the rate of change (acceleration). However, there is a lack of consensus regarding specific use of PSAV to predict PrCA [1,2].

Some evidence shows that PSA change over time is nonlinear, and the pattern may vary according to disease aggressiveness [14]. PSAV, as typically derived from linear regression or calculated as the simple

The authors have no financial or other conflicts to disclose.

* Corresponding author. South Carolina Statewide Cancer Prevention and Control Program, University of South Carolina, 915 Greene, Columbia, SC 29208.
** Corresponding author. Department of Epidemiology and Biostatistics, Arnold School of Public Health, University of South Carolina, 915 Greene, Columbia, SC 29208.
E-mail addresses: shoaibi@email.sc.edu (A. Shoaibi), JHEBERT@mailbox.sc.edu (J.R. Hébert).

http://dx.doi.org/10.1016/j.annepidem.2015.10.003
1047-2797/

average difference of multiple PSA measures, may be too limited to describe the pattern of change seen in PrCA [13,15]. By contrast, nonlinear mixed models [16] provide an alternative to describing PSA rate of change in men subsequently diagnosed with PrCA. These models, which take into account repeated measures and allow for linear and nonlinear functions [17], are ideal to model PSA repeated measures without assuming a linear pattern of change. In the early 1990s, Carter et al. [6] proposed a multiphase nonlinear model to compute PSAV to describe PrCA growth patterns. The current availability of algorithms to fit nonlinear mixed models to large data sets derived from screening trials creates new and unique opportunities to develop models to predict PrCA based on PSAV.

In this study, we aimed to fit nonlinear mixed models to data from the Prostate, Lung, Colorectal, and Ovarian Cancer Screening Trial (PLCO) [18]. The goal is to quantify and compare the trajectory of PSA change over time among men: (1) with no evidence of PrCA at trial end, (2) diagnosed with low-risk PrCA, and (3) diagnosed with high-risk PrCA. We hypothesize that the pattern of change in PSA is significantly different for men diagnosed with high-risk PrCA when compared to men in the two other groups.

Material and methods

In this retrospective cohort study, using data from the PLCO clinical trial, we retrospectively "followed" individuals' repeated PSA measures over time until they were confirmed to have been diagnosed with high-risk or low-risk PrCA or exited the study without a cancer diagnosis.

Setting

Analyses used data from 38,340 men randomized into the PrCA screening arm of the PLCO trial, details of which are described elsewhere [18]. Each man was expected to comply with up to six annual blood draws during the initial 6 years of active screening and then followed for an additional 7 years.

Participants are men aged between 50 and 75 years at baseline with four or more PSA measures. Potential sources of misclassification were excluded: men with reported unconfirmed diagnosis of PrCA, those who were classified as nonresponsive or lost-to-follow-up; and those who did not have complete diagnostic and/or biopsy information after a "positive" screening. Data from men with benign prostatic hyperplasia (BPH) at baseline or those with incomplete information on baseline age, body mass index (BMI), or race were excluded (Fig. 1).

Definition

The classification of PrCA into high or low biological risk was based on the prognostic stage introduced by The American Joint Committee on Cancer in 2010 [19]. Any PrCA that met one of these criteria was considered high biological risk: PSA level $\geq$20 ng/mL before or at the time of diagnosis; or a cancer that had invaded the prostate capsule, PrCA involving one or more lobe; or Gleason score (if available) $>$ 7. All other PrCAs were classified as low risk.

Statistical methods

Individual and mean trajectories of PSA were derived by plotting PSA as a function of time for each study group using a "spaghetti plot" for individual curves and locally weighted scatter-plot smoothing regression for the mean trajectory. These graphical tools were used to display the pattern that PSA changes over time and explore suitable functions that could be used in a statistical regression to model this pattern. These observed plots supported prior observations that PSA levels increase with age and/or time and that this change is not always constant and/or linear; within the PrCA group, an accelerating trend was observed sometime close to diagnosis.

Based on these preliminary analyses, we used multiphase nonlinear mixed regression to model PSA change over time. Two different modeling approaches were taken by considering: (1) PSA as a function of time (years to exit/diagnosis) and (2) the change of PSA over time on the natural log-transformed scale of the PSA measures.

(1) Linear–exponential piecewise PSA model allowed for estimating the individual PSA as a function of time (defined as years from entry to the diagnosis and/or exit). We used a two-phase function in the regression; a linear phase followed by rapid exponential increase. The phases were assumed to be connected through a transition point and/or change point (CP), unknown, *a priori*. We used this multiphase function to accommodate the observed PSA refectories. Figure 2 B and C represents this function. We fitted this model in two stages
   a. Because the pattern of change in PSA was hypothesized to be significantly different by study group, we started by fitting the same linear-exponential piecewise function for all participants, including an interaction term between "group" and all time-associated variables. This allowed for different coefficient estimates for each of the three groups. Fixed and random effects were included to estimate the mean and to allow for individual variation on the intercept, time coefficients, and the CP; that is, the number of years before diagnosis when the PSA pattern transitions from a linear to exponential growth pattern. The full mixed-effect model for the data is described in Appendix A. The most parsimonious model was determined by backward elimination of nonsignificant terms. As expected, cancer groups exhibit a significant exponential stage. The estimate of CP for the noncancer group was significantly lower (very close to zero) than the values for cancer groups. By backward elimination, a reduced model was introduced that allowed for transition to an exponential phase among the cancer groups only and reduced the function for noncancer group into a linear phase.
   b. We then used the resulting reduced model (allowing for transition to an exponential phase among the cancer groups only) to establish the PSA growth curve and estimate average PSAV as ng/mL/year per group while adjusting for baseline age (in three groups [$\leq$55, 55–65, $\geq$65], BMI (kg/m^2), PSA measure (ng/mL), and race (African American vs. others). To investigate and account for possible effect modification on PSA change over time by these variables, we included an interaction term between these variables and time. The simplified presentation of the reduced mixed-effect model is shown in Appendix A.

(2) Linear–Linear piecewise LOG PSA model allowed estimating the change of PSA over time on the natural log-transformed scale of the PSA measures. We regressed individual log (PSA + 1) as a function of time (years to diagnosis/exit). This transformation results in nonheterogeneous variances among errors and allows for a realistic linear assumption of PSAV and represents PSA change over time as an annual percent rate (change) instead of an absolute change. It replaces the observed linear-exponential relationship and/or function with a linear-linear function and simplifies derivation of PSAV by allowing for a single growth rate for all years post the CP. We replicated the two-model selection process described above:
   a. We started by fitting an initial model that allowed the same Linear–Linear piecewise function with unknown continuous CP for all groups. Fixed and random effects were included to estimate the mean, which allows for individual

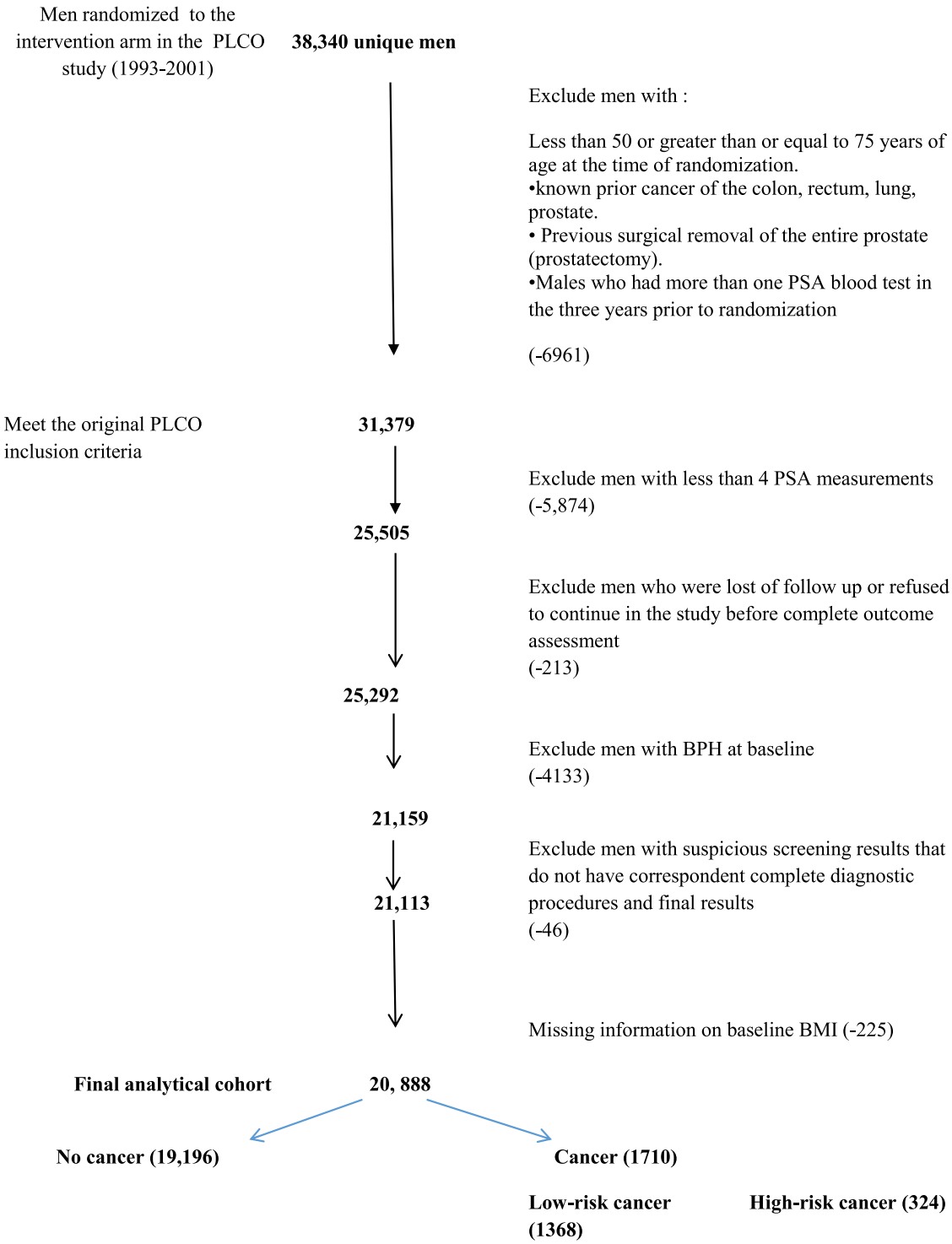

Fig. 1. The study cohort tree.

variation on the intercept, time coefficients, and the CP. The full mixed-effect model for log PSA is described in Appendix A. Because cancer groups were found to exhibit a significant second time coefficient that was not significant in the noncancer group, the most parsimonious model was chosen by backward elimination of nonsignificant terms that allowed for transition to a second linear phase among the cancer groups only and reduced the function for noncancer group into a one linear phase.

b. We then produced a reduced model to describe growth of log (PSA + 1) as a function of time to diagnosis/exit while adjusting for potential confounders. This allowed a transition to a second linear phase among the cancer groups only. The reduced mixed-effect model for log PSA is described in Appendix A.

PSA rate was estimated by taking the first derivative of the final equation in each model. We estimated PSAV at 1 and 2 year before

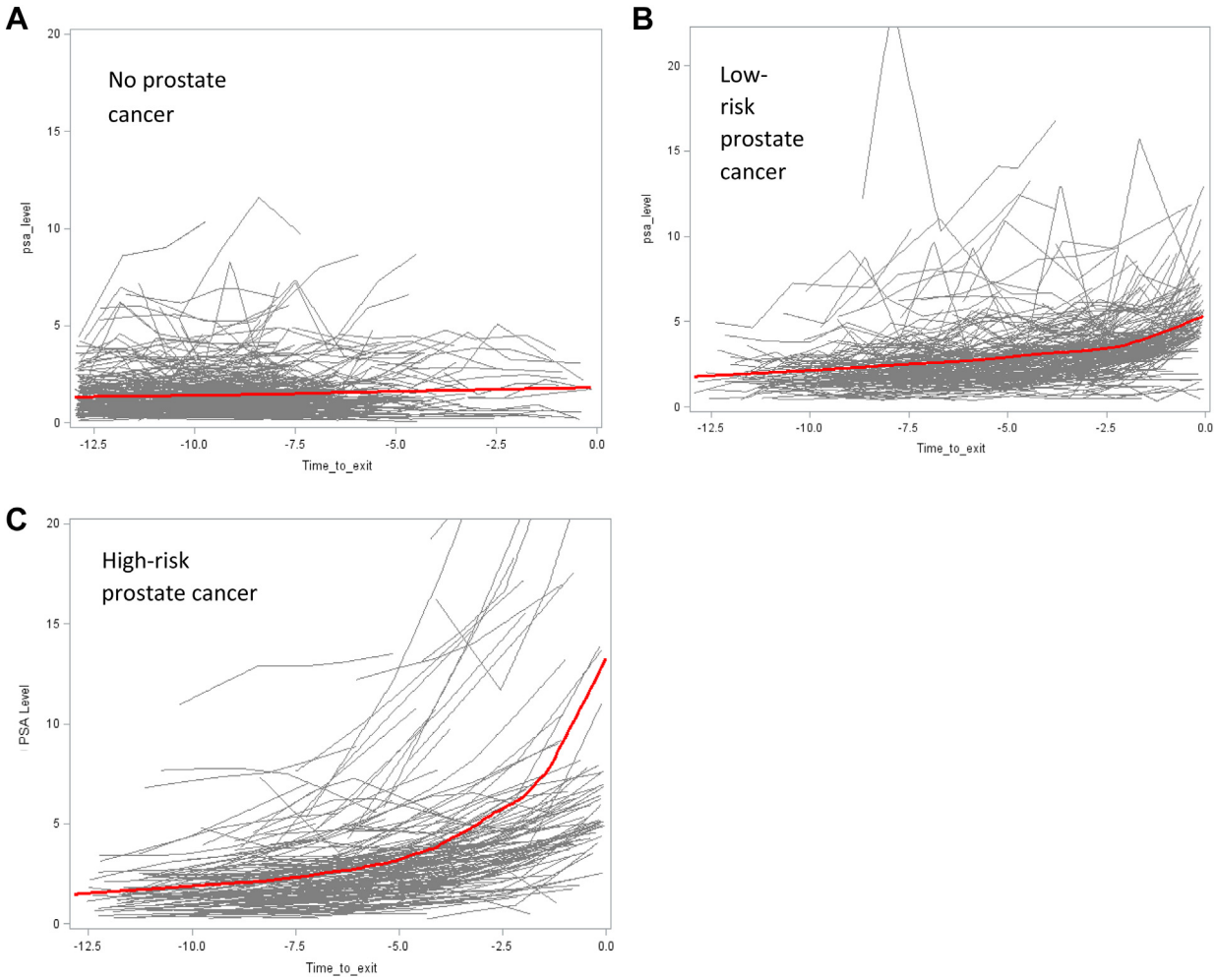

Fig. 2. Individual and mean longitudinal trajectories of PSA as a function of time (years from diagnosis/exit) in men with high-risk prostate cancer (*C*), low-risk prostate cancer (*B*), or no evidence of prostate cancer (*A*). The *black line* represents observed PSA trajectories of individual participants. The *red curve* represents the estimated mean curve in each obtained by locally weighted scatter-plot smoothing regression. (For interpretation of the references to color in this figure legend, the reader is referred to the Web version of this article.)

diagnosis/exit. We reported the results from each and compared the statistics of the derived PSAV (mean, median, percentile, and range) among the three study groups. The models included time variables, main effects of baseline characteristics, and corresponding interactions with the time variables (at the two phases). The time variable corresponded to PSA slope, and interaction of time with baseline characteristics corresponded to the influence of these characteristics on PSA slope/change. We compared the statistics of PSAV derived from a traditional formula for PSAV[1] and that derived from our proposed model.

Sensitivity, specificity, and the area under the curve of the receiver operator curve for predicting high-risk PrCA were based on derived PSA rates 1 year before diagnosis/exit. All analyses were conducted using SAS 9.4 (Cary, NC) with significance level of 0.05.

Results

Cohort demographics

After applying relevant exclusion and inclusion criteria, data from 20,888 men were available for analyses. Table 1 shows baseline characteristics for the three study groups. χ^2 tests and two-sided *t* tests were used for statistical comparisons of categorical and continuous variables. Compared to healthy men, those diagnosed with either high-risk or low-risk PrCA were found to be older at baseline, have shorter follow-up, higher PSA measures at baseline, fewer PSA measurements, and a shorter period between last PSA test and study exit. AA men and men with family history of PrCA were more likely to be diagnosed with PrCA.

Description of PSA changes over time

Figure 2 illustrates the observed trajectory of the three groups separately. For men in the noncancer group, we observed a linear trend of slight increase over time. A similar linear pattern was observed among the two cancer groups but only during the initial years of follow-up. Among the low-risk cancer group, the linear phase changes to exponential phase about 2–3 years before diagnosis. The linear-exponential pattern is more pronounced among high-risk cancer patients and the CP takes place earlier; around 4–5 years before diagnosis. Table 2 reports the unique CP statistics for the two cancer groups estimated from the final reduced models.

Table 3 summarizes PSA change and/or rate over time using different methods. Men diagnosed with high-risk PrCA had a statistically significantly higher estimate of absolute PSA change over time across different methods of estimation. The annual percent (%)

[1] Using the arithmetic equation PSAV $= (1/(n-1)) \times (\sum_{i=1}^{n}(p_i - p_{i-1})/(t_i - t_{i-1}))$, where n = total number of PSA tests, p = PSA value, t = time at PSA test.

Table 1
Characteristics of participants by study groups ($n = 20,888$)

Variable	Men with no cancer (19,196)	Men with LRC* (1368)	Men with HRC† (324)	Comparison (*P* for difference between study groups by characteristic)		
				No cancer versus LRC	No cancer versus HRC	LRC versus HRC
Race, *n* (%)						
African American	742 (90.05)	62 (7.52)	20 (2.43)	.098	.03	.20
Others	18454 (91.98)	1306 (6.51)	304 (1.52)			
Ethnicity, *n* (%) (missing = 606)						
Non-Hispanic	18203 (91.84)	1310 (6.61)	308 (1.55)	.32	.88	.74
Hispanic	428 (92.84)	26 (5.64)	7 (1.52)			
Family history, *n* (%) (missing = 144)						
No	17773 (91.33)	1225 (7.06)	284 (1.62)	<.001	<.001	.33
Yes, immediate family member	1291 (87.09)	132 (9.96)	39 (2.94)			
Age, *n* (%) (y)						
≤ 55, *n* = 2228	2096 (94.08)	107 (4.8)	25 (1.12)	.0004	.006	.34
55–65, *n* = 13,658	12560 (91.96)	898 (6.57)	200 (1.46)			
>65, *n* = 5002	4540 (90.76)	363 (26.54)	99 (1.96)			
Mean (95% CI)	61.42 (61.34–61.49)	62.21 (61.96–62.46)	62.73 (62.17–63.29)	<.001	<.001	.08
BMI, *n* (%)						
≤30 kg/m^2	14431 (91.58)	1068 (6.78)	258 (1.64)	.016	.0655	.54
>30 kg/m^2	4765 (92.87)	300 (5.85)	66 (1.29)			
Mean (95% CI)	27.75 (27.67–27.81)	27.34 (27.14–27.54)	27.63 (27.22–28.05)	<.001	.6	.20
PSA at baseline (ng/mL) mean/median (95% CI)	1.05/1.06 (1.04–1.06)	2.51/2.16 (2.42–2.59)	2.91/1.94 (2.37–3.46)	<.001	<.001	.14
Years of follow-up (y) mean/median (95% CI)	11.49/11.51 (11.46–11.52)	7.52/7.47 (7.37–7.66)	8.24/7.85 (7.54–8.16)	<.001	<.001	.05
Number of PSA tests mean/median (95% CI)	5.59/6.00 (5.58–5.60)	5.28/6.00 (5.24–5.33)	5.21/5.00 (5.12–5.30)	<.001	<.001	.16
Years from last PSA to exit or diagnosis mean/median (95% CI)	6.56/7.17 (6.54–6.59)	2.92/2.57 (2.79–3.04)	3.36/3.32 (3.07–3.64)	<.001	<.001	.005

* LRC: Low-risk prostate cancer.
† HRC: high-risk prostate cancer.

rate is higher among men who developed PrCA but was comparable between high-risk and low-risk PrCA. PSA annual change estimated by our models illustrated a narrower 95% CI (less variability around the mean value). Also, traditional methods cannot capture second-order effects of PSA exponential growth after CP, whereas this model can.

PSAV curves across age, race, and study groups

Tables 4 and 5 show PSA rate of change and annual percent PSA at 1 year before diagnosis and/or exit. These rates are illustrated for all study groups, stratified by age and race and adjusted for baseline distribution of BMI and initial PSA value of 1.3 ng/mL. After the CP and at 1 year before diagnosis and/or exit, the absolute PSA rate among men in the high-risk cancer group appears to be significantly greater compared to no cancer and low-risk groups. The annual percent rate is higher among men who developed PrCA, but the differences between high and low-risk groups were smaller.

A threshold of 0.37 has the best combination of sensitivity (97.2%) and specificity (97.3%) to detecting high-risk PrCA in a window time of 1 year. A lower threshold of 0.10 has a sensitivity of 99.7% but a specificity of 91.0%. A threshold of 0.42 has a specificity of 99.7% but a sensitivity of 90.3%.

Discussion

Results using data from the PLCO trial to fit mixed models to describe and quantify PSA change over time before PrCA clinical diagnosis support the hypothesis that the pattern of PSA varies with age, time, or both, and is significantly different for men diagnosed with high-risk PrCA versus others. This pattern is evident starting 2–5 years before the clinical diagnosis of the disease. Both cancer groups demonstrated an acceleration of PSA values presenting an inflection in PSA trajectories transitioning from a linear into an exponential pattern. The high-risk cancer group had a more aggressive exponential pattern with the CP occurring earlier than in the low-risk PrCA group. Also, men in the high-risk cancer group had much higher absolute PSA rate compared to the two other groups within 1 year before diagnosis and a threshold of 0.37 ng mL/year produced the optimal combination of sensitivity and specificity to predict high-risk PrCA. The findings of this study build on past knowledge that PSA change over time can distinguish high-risk PrCA, which appears to be associated with a distinct PSA trajectory.

Table 2
Change point* mean, median by cancer risk groups

Model	Outcome	Function	Change point summary		
			Group	Mean (95% CI)	Median (25th, 75th)
Annual PSA rate	PSA	Linear–exponential	Low-risk prostate cancer	2.58 (2.58–2.58)	2.62 (2.31–3.02)
			High-risk prostate cancer	5.21 (4.85–5.58)	5.24 (4.75–5.59)
Annual % PSA rate model	Log PSA	Linear–linear	Low-risk prostate cancer	2.00 (2.00–2.00)	2.00 (2.00–2.00)
			High-risk prostate cancer	3.96 (3.61–4.31)	3.96 (3.70–3.97)

* change point is the number of years before/until diagnosis when PSA pattern transition into exponential increase.

Table 3
PSA rate over time (velocity) in the three study groups estimated by different methods

Method	Men with no cancer (19196), mean (95% CI)	Men with LRC (1368), mean (95% CI)	Men with HRC (324), mean (95% CI)	Comparison (*P* for difference between study groups)		
				No cancer versus LRC	No cancer versus HRC	LRC versus HRC
Arithmetic velocity* (ng/mL/year)	0.06 (0.06–0.07)	0.37 (0.34–0.39)	0.79 (0.55–1.03)	<.001	<.001	<.001
Annual rate before change point (ng/mL/year)	0.05 (0.05–0.05)	0.16 (0.15–0.17)	0.13 (0.11–0.16)	<.001	<.001	.21
Annual rate after change point (1 year before diagnosis) ng/mL/year	0.05 (0.05–0.05)	0.59 (0.52–0.66)	2.60 (2.11–3.09)	<.001	<.001	<.001
Annual % PSA rate before change point	1.63% (1.57%–1.68%)	5.56% (5.33%–5.78%)	5.06% (4.54%–5.57%)	<.001	<.001	.31
Annual % PSA rate after change point	1.63% (1.57%–1.68%)	10.85% (9.02%–12.68%)	12.10% (10.3%–14.17%)	<.001	<.001	.09

* using the arithmetic equation,PSAV $= (1/(n-1)) \times (\sum_{i=1}^{n}(p_i - p_{i-1})/(t_i - t_{i-1}))$, where n = total number of PSA tests, p = PSA value, t = time at PSA test.

Although producing excellent sensitivity and specificity, our findings are consistent with those from previous studies. Carter et al. [6], Pearson et al. [20], and Inoue et al. [21] described PSA pattern using a nonlinear mixed model approach. They also reported a CP at which PSA starts to accelerate among individuals who developed PrCA and provided evidence of higher and earlier progression of PSA change in metastatic PrCA. Past studies suggested thresholds of 0.40 [22] ng/mL/year and 0.75 ng/mL/year [23] to distinguish virulent PrCA, values within the lower range of those we report for high-risk PrCA.

Past studies that used PSA change summary statistics to predict PrCA reported high intraindividual variability within the comparison groups (cancer and no cancer)—making it difficult to find an optimum threshold to predict PrCA. Using our method to quantify PSA velocity or rate of change we were able to identify a distinct range of PSA rates when considering high-risk cancer versus low-risk cancer and no cancer. Our analysis is unique with respect to previous studies that have estimated the individual velocities using a linear model within a narrow time frame, using few PSA measures in close intervals. Our model is flexible, as it does not assume a monotonic rate of change, and it accounts for the actual pattern of PSA trajectory, uses five to six PSA measures taken annually across a time frame of 1–14 years before exit, accounts for baseline participant characteristics, and relies on a large sample size.

The PCLO data have unique strengths derived from its rigorous design as a screening trial. However, our retrospective analysis is limited by cohort characteristics of men participating in the trial. For example, there are few young men and AAs. Second, the PSA measures were collected over the first 6 years of enrollment and follow-up continued for up to 14 years, leaving a gap of up to 7 years when PSA was not measured. This gap period was significantly longer among men with no evidence of PrCA, creating a potential bias. However, given the slow linear pattern of PSA change in this noncancer group, it is unlikely that having PSA measures during the gap period would have influenced our findings. Third, our calculated velocities might be sensitive to the proposed piecewise model; it could be that lower rates among the noncancer group represent an underestimate secondary to the linear model that was used for this group. To test for this bias, and to ensure that the estimated velocities were independent of any preassumed pattern, we conducted a sensitivity analysis in which we used our first full model to estimate PSA rate of change. In this way, we allowed all individuals to either deviate into an exponential pattern or stay in a linear pattern, depending on which fit their observed PSA better. In the sensitivity analysis, the calculated PSA rates did not change, and the magnitude of the differences between the three groups remained the same, suggesting the robustness of the findings. Fourth is the threat from information bias and misclassification, especially among the noncancer group, for whom we did not have a biopsy to confirm their status. We limited this bias by restricting our analysis only to those with biopsy disconfirmation following a positive screen or those with a negative finding on screening. Finally, excluding men with less than three PSA measurements and those who were lost during the follow-up might have excluded lower or higher PSA values and thus may be of lower or higher risk for developing PrCA.

Conclusion

Our PSA model showed clear differences in PSA pattern among those who were diagnosed with high-risk PrCA when compared to

Table 4
Estimated annual PSA rate of 1 year before exit stratified by race, age, and study groups and fixed at baseline BMI of 25 and initial PSA of 1.3

Race	Age (y)	Group	Mean (95% CI)	Median	25th percentile	75th percentile
Non-African American	Youngest (≤55)	No cancer	0.05 (0.04–0.05)	0.04	0.02	0.06
		Low-risk cancer	0.65 (0.53–0.77)	0.69	0.58	0.88
		High-risk cancer	2.82 (2.08–3.56)	1.95	1.63	3.57
	Middle (55–65)	No cancer	0.05 (0.05–0.05)	0.04	0.02	0.07
		Low-risk cancer	0.47 (0.41–0.54)	0.55	0.42	0.71
		High-risk cancer	2.10 (1.65–2.54)	1.88	1.25	2.68
	Older (≥65)	No cancer	0.06 (0.05–0.06)	0.04	0.02	0.07
		Low-risk cancer	0.92 (0.79–1.06)	1.07	0.81	1.40
		High-risk cancer	4.30 (3.50–5.11)	4.21	2.88	6.33
African Americans	Youngest (≤55)	No cancer	0.05 (0.04–0.07)	0.04	0.03	0.05
		Low-risk cancer	0.69 (0.50–0.88)	1.21	0.81	1.26
		High-risk cancer	3.04 (2.04–4.05)	1.90	1.89	1.91
	Middle (55–65)	No cancer	0.06 (0.05–0.07)	0.04	0.03	0.07
		Low-risk cancer	0.51 (0.36–0.65)	0.70	0.50	0.94
		High-risk cancer	2.26 (1.60–2.93)	2.50	1.75	3.71
	Older (≥65)	No cancer	0.06 (0.05–0.07)	0.04	0.03	0.07
		Low-risk cancer	0.98 (0.73–1.22)	1.00	0.78	1.54
		High-risk cancer	4.62 (3.28–5.95)	3.82	2.11	4.09

Table 5
Estimated annual % PSA rate 1 year before exit stratified by age, race, study group and fixed at baseline BMI of 25 and initial PSA of 1.3

Race	Age	Group	Mean (95% CI)	Median	25th percentile	75th percentile
Non-African American	Youngest (≤55)	No cancer	1.48% (1.32%–1.64%)	11.91%	10.62%	13.77%
		Low-risk cancer	11.67% (8.96%–14.38%)	12.20%	11.25%	13.60%
		High-risk cancer	12.91% (10.01%–15.81%)	13.21%	11.34%	15.39%
	Middle (55–65)	No cancer	1.61% (1.55%–1.68%)	11.88%	10.39%	13.76%
		Low-risk cancer	10.53% (8.64%–12.42%)	11.52%	10.25%	13.07%
		High-risk cancer	11.79% (9.66%–13.91%)	12.56%	10.81%	14.37%
	Older (≥65)	No cancer	1.68% (1.57%–1.78%)	11.87%	10.39%	13.76%
		Low-risk cancer	10.93% (8.61%–13.26%)	11.81%	10.53%	13.33%
		High-risk cancer	12.18% (9.68%–14.68%)	12.80%	11.08%	14.18%
African Americans	Youngest (≤55)	No cancer	1.82% (1.53%–2.12%)	11.85%	10.62%	13.25%
		Low-risk cancer	14.11% (10.31%–17.91%)	15.55%	14.43%	17.81%
		High-risk cancer	15.31% (11.52%–19.10%)	9.36%	4.67%	14.04%
	Middle (55–65)	No cancer	1.96% (1.70%–2.21%)	11.72%	10.33%	13.53%
		Low-risk cancer	13.00% (9.62%–16.39%)	13.92%	12.75%	15.88%
		High-risk cancer	14.22% (10.87%–17.57%)	15.93%	10.33%	20.55%
	Older (≥65)	No cancer	2.02% (1.75%–2.29%)	11.80%	10.46%	13.90%
		Low-risk cancer	13.40% (9.78%–17.01%)	15.77%	13.52%	17.58%
		High-risk cancer	14.61% (11.04%–18.18%)	13.64%	4.26%	13.91%

men with low-risk PrCA or no PrCA diagnosis. This pattern should be considered when estimating the rate at which PSA changes with time. The current methods of assuming a linear pattern attenuate the ability of PSAV to differentiate men with different types of PrCA. When estimating PSA rate while considering the transition to exponential increase of PSA measure, we found that the range of PSA rates among men in the low-risk cancer group may slightly overlap with those with no cancer across the different age and race groups; whereas those with high-risk PrCA have significantly different PSA change rates with no overlap. Moreover, this clear distinction takes place within a window of time before clinical diagnosis that is relevant to early detection. Further research is required to thoroughly investigate and validate the predictive value of this method of calculating PSA rate in predicting high-risk PrCA. This will help to inform clinical and public health practice. The implications for this work are important given the need to distinguish virulent cancer from indolent cancer to make competent treatment decisions.

Acknowledgments

We acknowledge the cancer data access system, CDAC, for providing us with the data to investigate our scientific question.

Funding: The funding was provided by the South Carolina Statewide Cancer Prevention and Control Program, University of South Carolina, Columbia, South Carolina. This work also was supported by an Established Investigator Award in Cancer Prevention and Control from the Cancer Training Branch of the National Cancer Institute to JR Hébert (K05 CA136975) and by a grant from the National Cancer Institute, Center to Reduce Cancer Health Disparities (Community Networks Program) to the South Carolina Cancer Disparities Community Network-II (SCCDCN-II) [1U54 CA153461-01 Hebert, JR (PI)].

References

[1] Loughlin KR. PSA velocity: A systematic review of clinical applications. Urol Oncol 2014;32(8):1116–25.
[2] Vickers AJ, Thompson IM, Klein E, Caroll PR, Scardino PT. A commentary on PSA velocity and doubling time for clinical decisions in prostate cancer. Urology 2014;83(3):592–6.
[3] Vickers AJ, Brewster SF. PSA Velocity and Doubling Time in Diagnosis and Prognosis of Prostate Cancer. Br J Med Surg Urol 2012;5(4):162–8.
[4] Vickers AJ. Prostate cancer: why is PSA velocity such a sticky concept? Nat Rev Urol 2013;10(4):189–90.
[5] Orsted DD, Bojesen SE, Kamstrup PR, Nordestgaard BG. Long-term prostate-specific antigen velocity in improved classification of prostate cancer risk and mortality. Eur Urol 2013;64(3):384–93.
[6] Carter HB, Pearson JD. PSA velocity for the diagnosis of early prostate cancer. A new concept. Urol Clin North Am 1993;20(4):665–70.
[7] Liu M, Wang JY, Su HX, Wan G, Zhu L, Wang XM. A longitudinal study of PSA and its influential factors in a cohort of Chinese men with initial PSA levels less than 4 ng ml(-1). Asian J Androl 2013;15(4):483–6.
[8] Burton AJ, Martin RM, Donovan JL, Lane JA, Davis M, Hamdy FC, et al. Associations of lifestyle factors and anthropometric measures with repeat PSA levels during active surveillance/monitoring. Cancer Epidemiol Biomarkers Prev 2012;21(10):1877–85.
[9] Pejcic T, Hadzi-Djokic J, Topuzovic C, Basic D, Marjanovic A, Djurasic L. The analysis of some factors that influence on serum PSA level in localized prostate cancer patients: mathematical model. Acta Chir Iugosl 2011;58(1):81–7.
[10] Wright JL, Lin DW, Stanford JL. The effect of demographic and clinical factors on the relationship between BMI and PSA levels. Prostate 2011;71(15):1631–7.
[11] Martinez Sarmiento M, Espana F, Royo M, Estelles A, Aznar J, Peinado F, et al. [Factors affecting the PSA:alpha 1 antichymotrypsin/total PSA ratio] [in Spanish]. Actas Urol Esp 2001;25(5):350–6.
[12] Sanz Chinesta S, Martinez Jabaloyas JM, Boronat Tormo F, Martinez Sarmiento M, Jimenez Cruz JF. [Factors influencing variability of PSA blood concentration in patients with prostate benign hyperplasia] [in Spanish]. Actas Urol Esp 1996;20(10):867–72.
[13] Loeb S, Kettermann A, Ferrucci L, Landis P, Metter EJ, Carter BH. The optimal application of prostate-specific antigen (PSA) velocity to predict high-risk disease. Eur Urol 2008;54(5):978–9.
[14] Carter HB, Morrell CH, Pearson JD, Brant LJ, Plato CC, Metter EJ, et al. Estimation of prostatic growth using serial prostate-specific antigen measurements in men with and without prostate disease. Cancer Res 1992;52(12):3323–8.
[15] Connolly D, Black A, Murray LJ, Napolitano G, Gavin A, Keane PF. Methods of calculating prostate-specific antigen velocity. Eur Urol 2007;52(4):1044–50.
[16] Littell RC, Milliken GA, Stroup WW. SAS for mixed models. 2nd ed. Cary,NC,USA: SAS Institute Inc; 2006.
[17] Lindstrom ML, Bates DM. Nonlinear mixed effects models for repeated measures data. Biometrics 1990;46(3):673–87.
[18] Gohagan JK, Prorok PC, Hayes RB, Kramer BS, Prostate LC, Ovarian T, Cancer Screening Trial Project. The Prostate, Lung, Colorectal and Ovarian (PLCO) Cancer Screening Trial of the National Cancer Institute: history, organization, and status. Control Clin Trials 2000;21(6 Suppl):251S–72S.
[19] Edge SB, Compton CC. The American Joint Committee on Cancer: the 7th edition of the AJCC cancer staging manual and the future of TNM. Ann Surg Oncol 2010;17(6):1471–4.
[20] Pearson JD, Morrell CH, Landis PK, Carter HB, Brant LJ. Mixed-effects regression models for studying the natural history of prostate disease. Stat Med 1994;13(5-7):587–601.
[21] Inoue LY, Etzioni R, Slate EH, Morrell C, Penson DF. Combining longitudinal studies of PSA. Biostatistics 2004;5(3):483–500.
[22] Loeb S, Carter HB, Schaeffer EM, Kettermann A, Ferrucci L, Metter EJ. Distribution of PSA velocity by total PSA levels: data from the Baltimore Longitudinal Study of Aging. Urology 2011;77(1):143–7.
[23] Vickers AJ, Till C, Tangen CM, Lilja H, Thompson IM. An empirical evaluation of guidelines on prostate-specific antigen velocity in prostate cancer detection. J Natl Cancer Inst 2011;103(6):462–9.

Appendix A

1. A. The initial linear-exponential piecewise PSA model used an interaction term between the group type and time. To account for individual-level natural heterogeneity in the rate of growth, the transition point or change point and the intercept in each group, we included random effects for their corresponding estimates. The full mixed-effect model for the data can be written as follows:

$$PSA_{ij} = \begin{cases} \beta_0 + (\beta_g \times G) + (\beta_c \times C) + b_i + \left[(\beta_t + \beta_{tg} \times G + \beta_{tc} \times C + b_{ti}) \times \left((CP + \beta_{cpg} \times G + b_{cpi}) - x\right)\right], & x < cp \\ (\beta_0 + (\beta_c \times C) + (\beta_g \times G) + b_i) \times e^{(\beta_{t2} + \beta_{t2g} \times G + \beta_{t2c} \times C + b_{t2i}) \times (CP + \beta_{cpg} \times G + b_{cpi} - x)}, & x \geq cp \end{cases}$$

where PSA_{ij} is the PSA measure for i[th] individual j[th] occasion.

Coefficients of the linear phase:

β_0 is the value of PSA at the transition between linear and exponential phase

b_i is the random effect for β_0

β_g is the coefficient corresponding to the group effect

G is a categorical indicator of the group, and is replaced in the model by two binary dummy variables as follow:

$$\begin{cases} g_1 = 1 \text{ for low} - \text{risk prostate cancer}, & \text{else } g_1 = 0, \\ g_2 = 1 \text{ for high} - \text{risk prostate cancer}, & \text{else } g_2 = 0, \end{cases}$$

Coefficients of the exponential phase:

β_{t2} is the exponential rate constant during the exponential PSA phase

β_{t2g} is the coefficient corresponding to the effect of group on the exponential rate constant, that is, interaction between time and group in phase 2

β_{t2c} is a vector of coefficients corresponding to the effect of covariants on the exponential rate constant, that is, interaction between time and covariates in the second stage

b_{t2i} is the random effect on β_{t2}.

1. B. The reduced linear-exponential piecewise model (allowing a transition to an exponential phase among the cancer groups only) estimates average and individual PSAV as ng/mL/year per group while adjusting for baseline age, BMI (kg/m^2), PSA measure (ng/mL), and race (African American versus others). We included an interaction term between all these variables and time. The reduced mixed-effect model can be simplified to:

$$PSA_{IJ} = \begin{cases} \beta_0^* + b_i + \beta_t^* + b_{ti} \times (cp - x), & x < cp \text{ and for no cancer group} \\ \beta_0^* \times e^{\left(\beta_{t2}^* + b_{t2i}\right) \times \left(cp^* + b_{cpi} - x\right)}, & x \geq cp \end{cases}$$

where the set of (β_0^*, β_t^*, β_{t2}^*, cp^*) is adjusted for group and effect of age, BMI (kg/m^2), PSA measure (ng/mL), and race (AA vs. others).

β_0^* is the PSA at the transition point for cancer groups and at exit for no cancer group

b_i is the random effect on β_0^*

β_t^* is the linear time coefficient

b_{ti} is the random effect on β_t^*

β_{t2}^* is the exponential time coefficient

b_{t2i} is the random effect on β_{t2}^*

cp^* is the change point between linear and exponential phase

b_{cpi} is the random effect on CP

β_c is a vector of the coefficients corresponding to the effect of the set of covariates

C is a matrix representing the individual covariate values

β_t is the linear coefficient corresponding to the effect of time that is the linear rate of change

β_{tg} is the coefficient corresponding to the effect of the group on the linear rate of change; that is, interaction between time and group

β_{tc} is the coefficient corresponding to the effect of covariates on the linear rate of change; that is, interaction between time and covariates

b_{ti} is the random effect on β_t

X is time (years) before exit and/or diagnosis

CP is the change point (inflection point) between linear and exponential phase

2. A. The full mixed-effect model for log PSA

$$\log(PSA + 1) = \begin{cases} \beta_0 + (\beta_g \times G) + (\beta_c \times C) + b_i + \left[(\beta_t + \beta_{tg} \times G + \beta_{tc} \times C + b_{ti}) \times \left((CP + \beta_{cpg} \times G + b_{cpi}) - x\right)\right], & x < cp \\ \beta_0 + (\beta_c \times C) + (\beta_g \times G) + b_i + \left[(\beta_{t2} + \beta_{t2g} \times G + \beta_{t2c} \times C + b_{t2i}) \times \left((CP + \beta_{cpg} \times G + b_{cpi}) - x\right)\right], & x \geq cp \end{cases}$$

where

β_0 is the value of log(PSA) at the transition between the first and the second linear phases

β_g is the coefficient corresponding to the patient group effect on β_0

β_{cpg} is the coefficient corresponding to the effect of group on the change point

b_{cpi} is the random effect on cp

G is a categorical indicator of the group and is replaced in the model by two binary dummy variables as follows:
$\begin{cases} g_1 = 1 \text{ for Low} - \text{risk prostate cancer}, & else\ g_1 = 0, \\ g_2 = 1 \text{ for High} - \text{risk prostate cancer}, & else\ g_2 = 0, \end{cases}$

β_c is a vector of coefficient corresponding the effect of the set of covariate on β_0

C is a matrix representing the individual covariate values

b_i is the random effect for β_0

β_t is the first phase linear coefficient, that is, the linear rate of change at the first phase

β_{tg} is the coefficient corresponding to the effect of group on the first linear rate of change, that is, interaction between time and group at the first phase

β_{tc} is the coefficient corresponding to the effect of covariants on the first linear rate of change

b_{ti} is the random effect on β_t

X is time (years) before exit and/or diagnosis

CP is the change point between the first and the second linear phases

β_{cpg} is the coefficient corresponding to the effect of group on the change point

b_{cpi} is the random effect on cp

Coefficients of the second phase:

β_{t2} is the difference in rate of change between the first and the second phase

β_{t2g} is the coefficient corresponding to the effect of group on β_{t2}, that is, interaction between time and group in phase 2

β_{t2c} is a vector of coefficients corresponding to the effect of covariants on β_{t2}

b_{t2i} is the random effect on β_{t2}

Based on this model, the rate of change at the second phase is the addition of β_t and β_{t2}

2. B. The reduced mixed-effect model for log PSA

$$PSA_{IJ} = \begin{cases} \beta_0^* + b_i + \beta_t^* + b_{ti} \times (cp - x), & x < cp \text{ and for no cancer group} \\ \beta_0^* + \left(\beta_{t2}^* + b_{t2i}\right) \times \left(cp^* + b_{icp} - x\right), & x \geq cp \end{cases}$$

where the set of (β_0^*, β_t^*, β_{t2}^*, cp^*) is adjusted for group and all other coverlets effect.

β_0^*: log(*PSA*) *at the trsition point*

b_i: *random effect on* β_0^*

β_t^*: *linear time coeffcient*

b_{ti}: *random effect on* β_t^*

β_{t2}^*: *exponential time coeffcient*

b_{t2i}: *random effect on* β_{t2}^*

cp^*: *is the chamge point* between linear and exponential phase

b_{icp}: *random effect on CP*