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