Computational Statistics and Data Analysis 56 (2012) 362–369

Contents lists available at SciVerse ScienceDirect

Computational Statistics and Data Analysis

journal homepage: www.elsevier.com/locate/csda

Bayesian piecewise mixture model for racial disparity in prostate cancer progression

L. Zhao [a,*], M. Banerjee [b]

[a] Biostatistics Unit, University of Michigan Comprehensive Cancer Center, University of Michigan, Ann Arbor, United States

[b] Department of Biostatistics, University of Michigan, Ann Arbor, United States

ARTICLE INFO

Article history:
Received 13 April 2011
Received in revised form 15 July 2011
Accepted 15 July 2011
Available online 29 July 2011

Keywords:
Random change point
Mixture distribution
PSA profiles
MCMC
DIC

ABSTRACT

Racial differences in prostate cancer incidence and mortality have been reported. Several authors hypothesize that African Americans have a more rapid growth rate of prostate cancer compared to Caucasians, that manifests in higher recurrence and lower survival rates in the former group. In this paper we propose a Bayesian piecewise mixture model to characterize PSA progression over time in African Americans and Caucasians, using follow-up serial PSA measurements after surgery. Each individual's PSA trajectory is hypothesized to have a latent phase immediately following surgery followed by a rapid increase in PSA indicating regrowth of the tumor. The true time of transition from the latent phase to the rapid growth phase is unknown, and can vary across individuals, suggesting a random change point across individuals. Furthermore, some patients may not experience the latent phase due to the cancer having already spread outside the prostate before undergoing surgery. We propose a two-component mixture model to accommodate patients both with and without a latent phase. Within the framework of this mixture model, patients who do not have a latent phase are allowed to have different rates of PSA rise; patients who have a latent phase are allowed to have different PSA trajectories, represented by subject-specific change points and rates of PSA rise before and after the change point. The proposed Bayesian methodology is implemented using Markov Chain Monte Carlo techniques. Model selection is performed using deviance information criteria based on the observed and complete likelihoods. Finally, we illustrate the methods using a prostate cancer dataset.

Published by Elsevier B.V.

1. Introduction

Prostate cancer is the most frequently diagnosed cancer in men, and the second leading cause of cancer-specific death among men in the United States. For reasons that remain unclear, prostate cancer incidence rates are significantly higher in African Americans (AA) than in Caucasians (CC). Mortality from prostate cancer is also two to three times greater among AA than among similarly aged CC (American Cancer Society, 2010).

Since the advent of the PSA screening era, men are more often diagnosed at an earlier stage of disease. This has led to an increase in surgery over the past two decades (American Cancer Society, 2010; Lu-Yao et al., 1993). However, approximately 30%–35% of these surgically treated patients will have pathologically detected disease outside the prostate and 25%–30% will ultimately develop disease progression (Freedland et al., 2003; Lu-Yao et al., 1996). Outcome data of men treated with surgery demonstrate more advanced tumors and higher recurrence rates among AA compared to similarly aged CC (Moul et al., 1996; Powell et al., 1999). The observed differential in the recurrence rates following surgery could be a manifestation of an earlier or a more aggressive disease progression in AA compared to CC.

* Corresponding author. Tel.: +1 7347636898.
E-mail addresses: zhaolili@umich.edu (L. Zhao), mousumib@umich.edu (M. Banerjee).

0167-9473/$ – see front matter. Published by Elsevier B.V.
doi:10.1016/j.csda.2011.07.011

In the medical literature, recurrence of prostate cancer after surgery is defined as a progressive or sustained elevation of serum prostate-specific antigen (PSA); and is based on an elevated PSA level of 0.4 ng/ml. Current studies in the medical literature are based on analyses of time to recurrence, focusing only on the time from surgery to the time when the PSA measurement exceeded the threshold (0.4 ng/ml). Given that PSA is a moving target, serial measures of PSA in time describe a dynamic process rather than a static one, and can contribute toward understanding the pattern by which the threshold was attained. This paper develops a Bayesian piecewise mixture model to characterize PSA progression over time in AA and CC. Each individual's PSA trajectory is hypothesized to have a latent phase immediately following surgery followed by a rapid increase in PSA indicating regrowth of the tumor. The true time of transition from the latent phase to the rapid growth phase is unknown, and can vary across individuals; suggesting a random change point across individuals. The PSA profiles of AA may be characterized by an earlier change point or a steeper rise during the tumor regrowth phase or both.

Bayesian change point methods have been successfully applied to CD4 counts to predict the timing of HIV viral rebound (Kiuchi et al., 1995), to pre-diagnosis PSA profiles for predicting cancer onset (Bellera et al., 2008; Slate and Lark, 2001), and to model cognitive ability over time preceding diagnosis of dementia (Hall et al., 2003). Among patients undergoing surgery for prostate cancer, nearly one-third of the men will have pathologically detected disease that had already spread outside the prostate gland (Freedland et al., 2003). For these men, surgery will be useless and these patients may experience a rapid increase in PSA immediately following surgery, without going through a latent phase. We propose a two-component mixture model to accommodate patients with, as well as, without a latent phase. Within the framework of this mixture model, patients who do not have a latent phase are allowed to have different rates of PSA rise; patients who have a latent phase are allowed to have different PSA trajectories, represented by subject-specific change point, PSA value at the change point and rates of PSA rise before and after the change point. We use simulation-based approach that exploits recent advances in Markov Chain Monte Carlo (MCMC) techniques, to implement the proposed Bayesian methodology. Section 2 describes the model formulation. In Section 3 we present details of the Bayesian analysis, specifically prior models for the various parameters, and details of the MCMC technique for posterior analyses. Section 4 presents results of data analyses of a cohort of prostate cancer patients. Finally, Section 5 contains some concluding remarks and an appraisal of the approach adopted in the current article in contrast to other approaches used in this context.

2. Piecewise mixture model

2.1. Study description

The specific study involved patients who were surgically treated between 2001 and 2006 for clinically localized prostate cancer at Harper Hospital in Detroit, Michigan. Postoperative follow-up included serum PSA level measurements (along with digital rectal examination) every 3 months for the first two years, and every 6 months thereafter. We retained 94 men who had an undetectable serum PSA level (defined as an assay value of 0.05 ng/ml) right after surgery, no preoperative or postoperative hormonal and/or radiation therapy, race CC or AA, and who had experienced recurrence on or before 2009. In this study, 68% were CC; 47% had local, and 53% had regional stage disease. Mean pre-surgery PSA was 14.5 ng/ml (range 2.8–62.8 ng/ml). Median follow-up for these patients was 36 months (range 8–97 months). All PSA measurements recorded at or before the first recurrence were included in the analysis.

2.2. PSA trajectories

After surgery, PSA levels are undetectable and go through a latent phase (characterized by very slow change in PSA) and then start to increase again at variable rates across individuals, with close to exponential patterns once PSA rise begins. For this reason, we applied a base 2 logarithmic transformation to the PSA measurements. As pointed out by Bellera et al. (2008), this transformation has several advantages. First, it allows one to model the individual patient trajectories as piecewise linear before and after the change point within individuals. Second, the trajectories tend to be smoother than when plotted on the original PSA scale. The base 2 logarithmic transformation also has scientifically useful interpretation; the $\log_2$ PSA growth rate is equivalent to the number of PSA doublings per month, and its reciprocal corresponds to the PSA doubling time, a quantity of interest to clinicians.

Fig. 1 presents a $\log_2$ PSA trajectory over time for a given individual i; the time origin is at surgery. We are primarily interested in estimating the $\log_2$ PSA rate of change during the latent phase (i.e. slope of the first line), the $\log_2$ PSA rate of rise during the rapid phase (i.e. slope of the second line), the change point or time of transition, and the variability associated with these parameters in the population. Furthermore, we are interested in assessing whether these parameters differ by race, after adjusting for cancer stage at diagnosis and pre-surgery PSA.

2.3. Model formulation

A piecewise linear model was used to model the slow change in $\log_2$ PSA before the transition time and rapid change in $\log_2$ PSA after the transition time. Each patient is allowed to have a different $\log_2$ PSA trajectory, represented by a subject-specific slope before the transition point, a subject-specific transition point, a subject-specific $\log_2$ PSA value at the transition point and a subject-specific slope after the transition point. However, not all patients have a latent phase. Some patients may