Biostatistics (2004), **5**, 3, *pp.* 483–500
doi: 10.1093/biostatistics/kxh003

Combining longitudinal studies of PSA

LURDES Y. T. INOUE*
University of Washington, Department of Biostatistics, F-600 Health Sciences Building, Campus Mail Stop 357232, Seattle, WA 98195 and Fred Hutchinson Cancer Research Center, 1100 Fairview Avenue North, MP 665, P.O. Box 19024, Seattle, WA 98109, USA
linoue@u.washington.edu

RUTH ETZIONI
Fred Hutchinson Cancer Research Center, 1100 Fairview Avenue North, MP 665, P.O. Box 19024, Seattle, WA 98109, USA

ELIZABETH H. SLATE
Medical University of South Carolina, Department of Biometry and Epidemiology, 135 Cannon Street, Suite 303, P.O. Box 250835, Charleston, SC 29425, USA

CHRISTOPHER MORRELL
Loyola College in Maryland, Mathematical Sciences Department, 4501 North Charles Street, Baltimore, MD 21210, USA

DAVID F. PENSON
Keck School of Medicine and University of Southern California/Norris Cancer Center, Department of Urology, 1441 Eastlake Avenue, Suite 7416, Los Angeles, CA 90033, USA

SUMMARY

Prostate-specific antigen (PSA) is a biomarker commonly used to screen for prostate cancer. Several studies have examined PSA growth rates prior to prostate cancer diagnosis. However, the resulting estimates are highly variable. In this article we propose a non-linear Bayesian hierarchical model to combine longitudinal data on PSA growth from three different studies. Our model enables novel investigations into patterns of PSA growth that were previously impossible due to sample size limitations. The goals of our analysis are twofold: first, to characterize growth rates of PSA accounting for differences when combining data from different studies; second, to investigate the impact of clinical covariates such as advanced disease and unfavorable histology on PSA growth rates.

Keywords: Bayesian hierarchical model; Interval-censored data; Longitudinal data; Meta-analysis; Prostate-specific antigen (PSA).

1. INTRODUCTION

Prostate-specific antigen (PSA) is a biomarker widely used for the early detection of prostate cancer. Since its introduction in the late 1980s, PSA has changed the way prostate cancer is diagnosed and

*To whom correspondence should be addressed.

 © Oxford University Press 2004; all rights reserved.

Downloaded from https://academic.oup.com/biostatistics/article/5/3/483/310204 by guest on 21 August 2022

managed. However, from the beginning, its use has been fraught with controversy. For example, since PSA is produced by the prostate epithelium, it can also be elevated by other prostate conditions, including benign prostatic hyperplasia and prostatitis. Many attempts have been made to improve the specificity of the test. These include use of information on PSA growth over time (Whittemore *et al.*, 1995) and thresholds for positivity that vary by age (Oesterling *et al.*, 1993). Recent publications (e.g. Punglia *et al.*, 2003) have cast doubt on the sensitivity of the test, suggesting that a non-trivial proportion of men with normal test results may harbor latent prostate cancer. Moreover, the survival benefits associated with PSA screening are still uncertain (Frankel *et al.*, 2003), and there are concerns that the test is overdiagnosing indolent cancers that would never have presented clinically (Etzioni *et al.*, 2002a). In spite of these uncertainties, use of the test has become widespread, particularly in the USA (Etzioni *et al.*, 2002b). The vast majority of prostate cancers detected today are diagnosed with early-stage, low-volume disease (Jemal *et al.*, 2003), making it difficult for the clinician to determine if a newly diagnosed patient has indolent disease or biologically aggressive cancer that would progress to an advanced stage if left untreated.

In determining diagnostic characteristics of PSA, retrospective analyses of serial PSA levels have been extremely valuable. To date, five studies have provided information about PSA growth from frozen serum samples that were stored prior to prostate cancer diagnosis (Carter *et al.*, 1992a,b; Pearson *et al.*, 1994; Morrell *et al.*, 1995; Whittemore *et al.*, 1995; Slate and Clark, 1999; Slate and Turnbull, 2000). Using the data from these studies, several authors have investigated PSA growth and the natural history of prostate cancer. All modeled PSA growth as either linear or exponential over time with an acceleration of growth at some point prior to disease onset. The results are, however, variable. Slate and Turnbull (2000) report estimates of annual percentage change in PSA from three individual studies, ranging from 13% to 20%. The variability of the estimates has important implications for clinical practice. For example, if it is known that PSA in cancer cases increases at an annual rate of 25% then a man with an annual increase of 20% will be treated quite differently than one with an increase of 13%. The results also have policy implications because decision analysis models (e.g. Etzioni *et al.*, 1999; Ross *et al.*, 2000) use PSA growth as input variables. Finally, we note that these analyses were based on single studies. The issue of estimating PSA growth by combining PSA data while accounting for study differences was never addressed.

When several studies are available to address the same scientific question, it is natural to consider combining information across studies in order to draw overall conclusions. The process of combining information from related studies is called *meta-analysis*. When combining studies, a natural concern is adequately capturing the study-to-study variability due to differences in study designs, protocols, and patient populations. Bayesian hierarchical models provide a natural modeling approach to address study-to-study variability. Such models have been used in the literature on a variety of problems; the earliest account is the paper by DuMouchel and Harris (1983). For a review of current meta-analytic methods and applications of hierarchical models see, for example, Stangl and Berry (2000).

This article presents a meta-analysis of PSA growth in prostate cancer patients using three of the above mentioned retrospective studies. The combined dataset represents the largest data resource available today with information about serial changes in PSA levels prior to a prostate cancer diagnosis. Our goal is to produce estimates of PSA growth that reconcile between-study differences while gaining greater insight into these differences. To do this, we propose a Bayesian hierarchical model that explicitly accounts for between-study variability, while also estimating the overall mean PSA growth trajectory across studies. Like previous studies, we use a change-point model to accommodate the non-linearity of PSA growth over time; the change-point is assumed to represent a latent point of transition from normal PSA growth to a point at which PSA is emitted at an accelerated rate. Since PSA is proportional to volume of prostate cancer (Kabalin *et al.*, 1995; Stamey *et al.*, 1987), we anticipate that this point will be highly correlated with onset of malignancy.

The combined dataset also provides a unique opportunity to address questions that could not be

Downloaded from https://academic.oup.com/biostatistics/article/5/3/483/310204 by guest on 21 August 2022

Table 1. *Summary statistics of the three retrospective longitudinal studies of PSA. Means and standard deviations* (*std*) *are provided for the age at last follow-up* (*FU*), *length of follow-up* (*in years*), *number of longitudinal measurements, and first and last log-transformed PSA measurements*

STUDY	GROUP	Individuals	Age at last FU		Length of FU		Measurements		First ln(PSA+1)		Last ln(PSA+1)	
			mean	(std)	mean	(std)	mean	(std)	mean	(std)	mean	(std)
CARET	Normal	88	67.33	(5.36)	7.78	(2.20)	6.40	(2.46)	0.77	(0.31)	0.93	(0.48)
	Local	53	60.00	(6.10)	5.34	(1.82)	4.55	(2.21)	1.42	(0.62)	2.04	(0.78)
	Metastasis	6	65.67	(5.43)	5.31	(2.13)	4.17	(2.04)	1.60	(0.88)	3.43	(1.87)
BLSA	Normal	282	60.05	(12.42)	16.15	(8.55)	6.34	(2.28)	0.52	(0.35)	0.82	(0.49)
	Local	41	71.76	(7.52)	19.27	(6.37)	8.42	(3.00)	0.67	(0.35)	1.98	(0.70)
	Metastasis	8	77.20	(8.11)	16.95	(8.71)	8.13	(3.04)	0.97	(0.70)	3.82	(0.97)
NPCT	Normal	898	76.98	(14.98)	12.48	(14.73)	7.09	(4.17)	0.77	(0.49)	0.90	(0.57)
	Local	51	72.27	(4.72)	3.68	(2.22)	7.31	(4.08)	1.63	(0.72)	2.11	(0.78)
	Metastasis	10	71.33	(5.31)	3.04	(1.84)	7.40	(3.69)	1.77	(0.92)	2.75	(0.84)

answered by previous studies because of their limited sample sizes. Indeed, the distinguishing feature of the combined dataset is that it has the largest collection of longitudinal PSA measures taken prior to diagnosis on disease cases. We exploit this strength to address the following question: does PSA provide information about whether a newly diagnosed prostate cancer case has disease that is likely to progress to an advanced stage if left untreated? To do this, we use our models to answer the following questions: (1) are PSA growth rates after the change-point different for cases diagnosed with early stage disease relative to those diagnosed with late-stage disease? and (2) are PSA growth rates after the change-point different for cases diagnosed with high-grade disease than for those diagnosed with low-grade disease? We allow post-change point growth rates to differ by stage and grade at diagnosis by incorporating information on these variables at the second level of the hierarchy.

The clinical importance of distinguishing aggressive localized prostate cancer from indolent disease cannot be overstated. Prior studies have estimated the overdiagnosis rate due to PSA testing to range between 30% and 50%, (Etzioni *et al.*, 2002a; Draisma *et al.*, 2003). This implies that 30% or more of patients diagnosed by PSA screening do not require aggressive therapy for the condition, although the majority of them will still undergo these invasive procedures. Given that the aggressive therapies for localized prostate cancer (surgery or radiation) are commonly associated with complications such as urinary incontinence and sexual and bowel dysfunction, all of which have been shown to significantly diminish quality of life (Penson *et al.*, 2003; Potosky *et al.*, 2000, 2002), overdiagnosis is therefore associated with considerable morbidity that might be avoided if a reliable method were available that could distinguish aggressive from indolent disease given information available at the time of diagnosis.

The paper is organized as follows. In Section 2 we provide a brief description of the studies in our meta-analysis. We describe Bayesian hierarchical change-point models in Section 3. We present model fit results in Section 4. Finally, in Section 5 we present some conclusions and directions for further work.

2. Retrospective studies

Our models combine data from three studies: namely, the Nutritional Prevention of Cancer Trials (NPCT), the Beta-Carotene and Retinol Efficacy Trial (CARET) and the Baltimore Longitudinal Study of Aging (BLSA). We provide a brief description of these studies below and details are presented in Table 1.

Downloaded from https://academic.oup.com/biostatistics/article/5/3/483/310204 by guest on 21 August 2022

2.1 *The Nutritional Prevention of Cancer Trials (NPCT)*

The NPCT (Clark *et al.*, 1996; Duffield-Lillico *et al.*, 2003) was a multi-center, double-blind, randomized, placebo-controlled cancer prevention trial aimed at determining whether a nutritional supplement of selenium would decrease the incidence of cancer.

The trial accrued subjects with a history of two or more basal cell carcinomas or one squamous cell carcinoma of the skin within the prior accrual year. Moreover, the eligibility criteria required that subjects had a five year life expectancy and no internal malignancies treated within the previous five years, and no histories of significant liver or kidney disease.

Subjects were randomized from 1983 through 1992 to receive either placebo or a nutritional dosage of selenium daily. Participants were scheduled to return to the clinic every six months for dermatological examinations, a blood draw, and evaluation for signs of selenium toxicity. The primary outcome was incidence of basal and squamous cell carcinomas of the skin. Secondary endpoints, established in 1990, were all-cause mortality, total cancer mortality, total cancer incidence, and incidences of lung, prostate and colorectal cancers. In particular, PSA levels were determined retrospectively from frozen plasma samples using the Abbott Diagnostics IMx PSA assay (Abbott Park, IL).

2.2 *The Beta-Carotene and Retinol Efficacy Trial (CARET)*

This double-blind trial conducted from 1988–96 was aimed at evaluating the chemopreventive efficacy and safety of beta-carotene and retinol in a population at risk for lung cancer. Blood was obtained at randomization and yearly during a pilot period and every two years thereafter. Prostate cancer screening or management of prostate cancer was not part of CARET; thus all diagnoses of prostate cancer were made within the patients' regular health care system.

For the PSA substudy (Ellis *et al.*, 2001), cases were selected from the group of 336 CARET participants diagnosed with prostate carcinoma between January 1985 and May 1997. From this group 90 men were identified who were between 50 and 65 years old at randomization with more than one serum specimen obtained before cancer diagnosis. Controls were men 50 to 65 years old at randomization without a subsequent diagnosis of prostate cancer or lung cancer. All controls had serum available from a pre-randomization visit and at least two subsequent blood specimens. Serum was stored at $-70°$C until assay. Total and free serum PSA was determined by the Abbott Diagnostics AxSYM PSA assay (Abbott Park, IL).

Chart abstraction was performed by a single investigator. Chart data consisted of medical records forwarded to CARET investigators by treating physicians. These data included clinical reports indicating the initial events leading to diagnosis as well as surgical, pathological and radiological reports describing tumor staging. From these records tumor clinical and pathological stage, and Gleason grade were recorded. Staging was done according to 1997 American Joint Commission on Cancer TNM criteria.

In 70 cases at least three measurements were made before diagnosis and in 62 controls at least three measurements were made before the matched case diagnosis. Among the initial events leading to the diagnosis of prostate carcinoma was abnormal digital rectal examination in 60 men, abnormal PSA in 70, voiding symptoms in 17 and other events in four. In some patients multiple events led to the diagnosis of cancer. In 18 cases PSA and in 14 digital rectal examination was recorded as the only event leading to diagnosis. These values imply that screening for prostate cancer may have led to diagnosis in some cases.

2.3 *The Baltimore Longitudinal Study of Aging (BLSA)*

The BLSA began in 1958 to study normal human aging (Shock *et al.*, 1984). The BLSA is an ongoing multi-disciplinary study of community-dwelling volunteers with the continued recruitment of

Downloaded from https://academic.oup.com/biostatistics/article/5/3/483/310204 by guest on 21 August 2022

new participants to replace other participants who leave the study due to death or dropping out. Participants in the BLSA return to the study center every two years for three days of free biomedical and psychological examinations. To date, over 1500 men with an average of about eight visits to the study center and 13 years of follow-up and approximately 900 women (women were first studied in 1978) with an average of about four visits and seven years of follow-up have participated in the study. The study participants are mostly white (95%), well educated (at least 75% have a college degree), and financially comfortable (82%). Besides describing human physiological changes that occur with age, the study has also been beneficial in helping examine differences between normal aging and disease processes.

The serum PSA levels have been measured in male BLSA participants either at the time of routine subject visits (since 1991) or using a frozen serum bank for available retrospective samples (before 1991) donated at participant visits. This study uses PSA data on men prior to the development of prostate cancer. PSA levels were determined using Tandem-R (Hybritech) assays.

2.4 *Data aspects*

In all studies described above, retrospective longitudinal PSA measurements were available for all prostate cancer cases along with the clinical stage of the disease at diagnosis. Clinical stage is categorized into local or metastatic stage of the disease. Disease grade was available for a subset of prostate cancer patients. In our analysis, low grade reflects Gleason score lower than or equal to 6, while high grade means Gleason score higher than 6. All studies also included normal patients. Table 1 summarizes the information available from each study. Here we only consider the data for patients who had at least two longitudinal measurements of PSA. Figure 1 shows the longitudinal profiles of PSA (log-transformed) for prostate cancer patients as a function of age by study and stage of the disease. The data indicate that PSA levels increase, but that the growth rate may not be constant in prostate cancer patients. Though not shown here, longitudinal profiles of (log-transformed) PSA indicate a linear increase of PSA when considering normal patients.

Though we were not dealing with screening studies, because of the time period in which they took place, some men may have had PSA screening. However, our dataset does not contain information that identifies who had PSA screening. Moreover, in our dataset the last measurement of PSA in each patient, though closest to diagnosis, may be different from PSA measured at detection.

Finally, we note that these studies not only differ in their designs, but also in the way PSA levels were determined. Each study used a different reagent kit; the BLSA used the Tandem-R assay (produced by Hybritech), the NPCT used the IMx assay, and the CARET study used the AxSYM assay (both AxSYM and IMx are produced by Abbott laboratories). Several studies (Leewansangtong *et al.*, 1998; Oesterling *et al.*, 1995; Brawer *et al.*, 1997) have confirmed that the different assays produce highly correlated, although not always equivalent, results.

Our models, while borrowing the information across studies, account for the specific study differences through a hierarchical structure which we describe in the next section.

3. Model specification

Let S denote the total number of studies considered for meta-analysis. Any particular study is indexed by $s (s = 1, \ldots, S)$ and contains N_s individuals. Therefore, the overall number of patients is $N = \sum_{s=1}^{S} N_s$. Similarly, let $i (i = 1, \ldots, N_s)$ denote the subject index in study s. Subject i has n_i measurements indexed by $j (j = 1, \ldots, n_i)$.

Serial measurements of PSA are available along with individuals' age at which the serum was sampled. Let $y_{ij}^s = \log(PSA_{ij}^s + 1)$ denote the (log-transformed) PSA level observed at age t_{ij}^s. In particular, for prostate cancer cases, let t_{i0}^s denote the age closest to disease diagnosis for patient i in study s.

Downloaded from https://academic.oup.com/biostatistics/article/5/3/483/310204 by guest on 21 August 2022

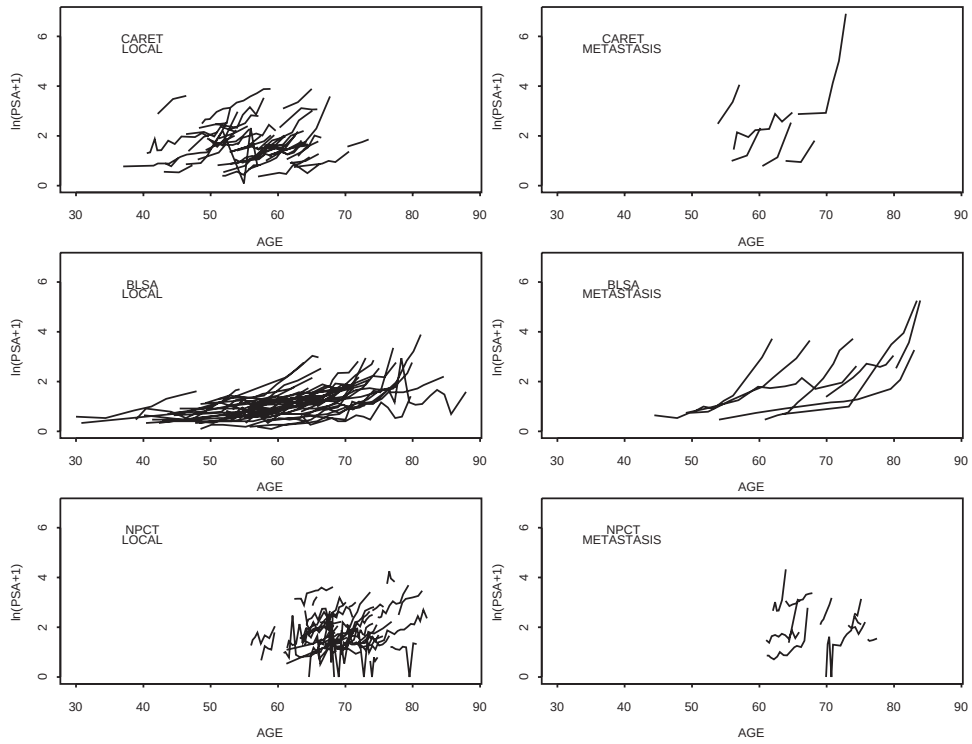

Fig. 1. Log-transformed PSA levels given age, group and study.

To address questions formulated in Section 1 we consider Bayesian hierarchical models. From a Bayesian perspective, such models are fully specified with (1) a model for PSA growth described by $p(y_{ij}^s|\boldsymbol{\theta}_i^s), s = 1,\ldots,S,\ i = 1,\ldots,N_s,\ j = 1,\ldots,n_i$, (2) a model for the inter-patient variation described by $p(\boldsymbol{\theta}_i^s|\boldsymbol{\theta}^s)$, (3) a model for the intra-study variation $p(\boldsymbol{\theta}^s|\boldsymbol{\theta})$, and finally (4) a model for the inter-study variability $p(\boldsymbol{\theta})$.

In what follows we describe the non-linearity of PSA growth rates with patient-specific change-point models. We formally combine the individual non-linear regression models with a hierarchical model. There is an extensive literature on Bayesian hierarchical change-point models. Earlier developments with discrete change-points include, for example, Booth and Smith (1982) and Carlin *et al.* (1992); while for continuous change points, see Stephens (1994); Slate and Cronin (1997) and Slate and Turnbull (2000).

Our models extend earlier approaches to the analysis of PSA growth in several directions; first, by formally addressing the study-to-study variability in PSA growth, and second, by allowing for different post-change point growth rates depending on subject-specific disease prognosis. Finally, we note that we model PSA as a function of age as opposed to years prior to diagnosis (Morrell *et al.*, 1995; Slate and Clark, 1999; Slate and Turnbull, 2000). This time-scale is flexible in that it allows us to present results as a function of years prior to diagnosis, but it also allows prediction of transitions on a scale that is appropriate when building natural history models.

Downloaded from https://academic.oup.com/biostatistics/article/5/3/483/310204 by guest on 21 August 2022

3.1 *A hierarchical model for PSA growth*

At the individual level, we propose the following non-linear model to describe growth of $\log(PSA+1)$ as a function of age:

$$y_{ij}^s = \begin{cases} b_{0i}^s + b_{1i}^s t_{ij} + \varepsilon_{ij}, & \text{for normal patients} \\ b_{0i}^s + b_{1i}^s t_{ij} + b_{2i}^s (t_{ij} - \tau_i)^+ + \varepsilon_{ij}, & \text{for prostate cancer patients} \end{cases} \quad (1)$$

where $x^+ = \max(0, x)$ and the error terms ε_{ij} are assumed independent and identically distributed with Normal$(0, \sigma^2)$: that is, a normal distribution with mean zero and variance σ^2. We assume that $1/\sigma^2 \sim$ Gamma(r_σ, s_σ), where r_σ and s_σ are known hyperparameters. (Here, Gamma(a, b) denotes a Gamma distribution with mean ab^{-1} and variance ab^{-2}.)

In words, equation (1) describes the growth of $\log(PSA+1)$. In normal patients the growth is constant over time, while in prostate cancer patients a piecewise model describes the growth, with τ_i denoting a change-point in the growth rate for patient i. Under this model, the growth of $\log(PSA+1)$ for patient i in study s is described by a linear model with intercept b_{0i}^s and slope b_{1i}^s. A prostate cancer patient has a transition at τ_i and experiences a change in the slope by b_{2i}^s after τ_i. In summary, PSA growth for a normal individual i in study s is characterized by a patient-specific parameter vector $\boldsymbol{\theta}_i^s = (b_{0i}^s, b_{1i}^s)'$ while for prostate cancer patients we have $\boldsymbol{\theta}_i^s = (b_{0i}^s, b_{1i}^s, b_{2i}^s, \tau_i)'$. Implicit in this formulation is the assumption that all prostate cancer patients experience a change-point in the growth rates and that normal patients do not experience a change-point at all. Moreover, the intercept is defined for 20 year-old men.

Though patients' change-point times are never observed, each prostate cancer patient contributes interval-censored information that a change-point occurred before diagnosis of the disease at age t_{i0}^s. This is formalized with a probability model that assigns probability mass one to the event $\{\tau_i < t_{i0}^s\}$.

3.2 *Hierarchical model*

To combine subject-specific non-linear regression models $p(y_{ij}^s|\boldsymbol{\theta}_i^s)$ defined in Section 3.1 we specify a model $p(\boldsymbol{\theta}_i^s|\boldsymbol{\theta}^s)$: that is, a model that explains the variation of the subject-specific parameters within studies. Let X_i be the subject-specific indicator variable for advanced stage disease. (Later, we specify a similar model for high- versus low-grade disease.) At the second stage of the hierarchical model we assume that

$$\begin{aligned} b_{0i}^s &\sim \text{Normal}(\beta_0^s, \sigma_0^2) \\ b_{1i}^s &\sim \text{Normal}(\beta_1^s, \sigma_1^2) \\ b_{2i}^s &\sim \text{Normal}(\beta_2^s + \gamma^s X_i, \sigma_2^2) \\ 1/\sigma_0^2 &\sim \text{Gamma}(r_{b_0}, s_{b_0}) \\ 1/\sigma_1^2 &\sim \text{Gamma}(r_{b_1}, s_{b_1}) \\ 1/\sigma_2^2 &\sim \text{Gamma}(r_{b_2}, s_{b_2}), \end{aligned} \quad (2)$$

where $i = 1, \ldots, N_s$ is the patient index, and $s = 1, \ldots, S$ is the study index. That is, we assume that within a study, all patients share the same mean intercept and mean pre-change point slope. However, patients with advanced disease (metastasis) at diagnosis may have a further increment in the mean post-change point slope. We note that the subject-specific parameters are allowed to vary from individual to individual in the same study. Here, the study-specific parameters are denoted by $\boldsymbol{\theta}^s = (\beta_0^s, \beta_1^s, \beta_2^s, \gamma^s)'$.

Downloaded from https://academic.oup.com/biostatistics/article/5/3/483/310204 by guest on 21 August 2022

We combine the study-specific parameters by specifying a model for $p(\boldsymbol{\theta}^s|\boldsymbol{\theta})$ where $\boldsymbol{\theta}$ denotes the population parameters, that is $\boldsymbol{\theta} = (\beta_0, \beta_1, \beta_2, \gamma)'$. We assume

$$\begin{aligned} \beta_k^s &\sim \text{Normal}(\beta_k, \psi_k^2) \\ \gamma^s &\sim N(\gamma, \sigma_\gamma^2) \\ 1/\psi_k^2 &\sim \text{Gamma}(r_k, s_k) \\ 1/\sigma_\gamma^2 &\sim \text{Gamma}(r_\gamma, s_\gamma) \end{aligned} \tag{3}$$

where, as before, s is the study index and $k = 0, 1, 2$ indexes, respectively, parameters associated with intercept, pre change-point slope and post change-point slope at either the population level or study level. Finally, at the fourth level of the hierarchy we assume that

$$\beta_k \sim \text{Normal}(m_k, v_k^2), k = 0, 1, 2 \text{ and } \gamma \sim N(m_\gamma, v_\gamma^2). \tag{4}$$

Completing our description of the hierarchical model, for the change-point variable τ_i, we assume that

$$\begin{aligned} \tau_i &\sim \text{LogNormal}(\mu, \sigma_\tau^2) I(\tau_i < t_{i0}^s) \\ \mu &\sim N(m_\mu, v_\mu^2) \\ 1/\sigma_\tau^2 &\sim \text{Gamma}(r_\tau, s_\tau). \end{aligned} \tag{5}$$

3.3 *Priors and posterior distributions*

For our analysis we considered proper but fairly non-informative priors in that all hyperpriors were initially set to have prior variance of 1000. Priors set on parameters associated with means were centered at zero, while those associated with precisions were centered at one.

Posterior distributions of the model parameters, while not available in closed analytical form, can be obtained with Markov chain Monte Carlo techniques. Full conditional distributions are available in closed form for all parameters, except for the subject-specific change-point parameters τ_i. We implemented a Metropolis-within-Gibbs algorithm; within a Gibbs sampling algorithm (Geman and Geman, 1984; Gelfand and Smith, 1990), the updating of τ_i was performed using the Metropolis–Hastings algorithm (Metropolis *et al.*, 1953). To obtain samples from the posterior distribution of all parameters, we ran our Markov chain for a burn-in period of 1000 000 iterations, followed by additional 250 000 iterations from which we stored sampled values at every 50th iteration.

As we will see in the next section, we also consider posterior predictive distributions. Those can be obtained using our posterior samples and the composition method (Tanner, 1996, p. 2). For example, suppose one is interested in the posterior predictive distribution of τ for a patient who is not in our combined dataset. This is an out-of-sample prediction formalized by

$$p(\tau|D) = \int p(\tau|\mu, \sigma_\tau^2)\, p(\mu, \sigma_\tau^2|D)\, \mathrm{d}\mu\, \mathrm{d}\sigma_\tau^2,$$

where D denotes our combined data and we have assumed that τ is conditionally independent of D given (μ, σ_τ^2). Note that this calculation does not use information about PSA readings, nor age at diagnosis, for this new patient. Hence, this posterior predictive distribution applies to the population of all new patients, not any specific patient. Since, in particular, the age at diagnosis is unknown, we obtain samples from $p(\tau \mid D)$ by using posterior samples for (μ, σ_τ^2) and drawing τ given (μ, σ_τ^2) from the *untruncated* $\text{LogNormal}(\mu, \sigma_\tau^2)$.

Downloaded from https://academic.oup.com/biostatistics/article/5/3/483/310204 by guest on 21 August 2022

4. MODEL ASSESSMENT

4.1 *Comparing alternative models*

In this section we justify the model proposed in Section 3 which we denote by M_0. This model allows for post change-point slopes to depend on disease prognosis as defined by the patient's clinical stage of the disease.

We defined three alternative models to assess whether not only the post-changepoint slopes, but also the change-points varied with stage. In the first alternative model M_1 we assumed that both the change-point and the post change-point slopes do not depend on disease prognosis. In symbols, model M_1 assumes that $\gamma^s \equiv 0$. Our second alternative model M_2 allowed for stage-dependent change-points, but that post change-point slopes are not stage dependent, that is, only $\gamma^s \equiv 0$, but $\tau_i \sim N(\mu_0 + \mu_1 X_i, \sigma_\tau^2)$ where X_i is the indicator of advanced disease for patient i. We complete the hierarchical formulation by assuming that $\mu_0 \sim N(m_{\mu_0}, v_{\mu_0}^2)$, $\mu_1 \sim N(m_{\mu_1}, v_{\mu_1}^2)$. Finally, our third alternative model M_3 allowed for stage-dependent change-points and post change-point slopes, that is, it combines the model describing M_0 with the hierarchical structure for the change-point described above for M_2.

From a Bayesian viewpoint, one may use Bayes factors (Kass and Raftery, 1995) for model comparison. Let D denote the data. To compare models M_0 and M_1, the Bayes factor is the ratio $BF_{0,1} = P(D|M_0)/P(D|M_1)$. Using the output from our MCMC analysis, we can compute the Bayes factors using harmonic means to compute

$$P(D|M_k) = \frac{1}{\frac{1}{M}\sum_{m=1}^{M}\left(\frac{1}{P(D|\hat{\boldsymbol{\theta}}_i, M_k)}\right)} \tag{6}$$

where M is the number of posterior samples $(\hat{\boldsymbol{\theta}}_i, i = 1, \ldots, M)$ under model k. We obtain $2 \log BF_{0,1} = 12.26$. Kass and Raftery (1995) provide guidelines for the interpretation of Bayes factors. We found $2 \log BF_{0,1} > 10$, suggesting a decisive preference for M_0 over M_1. Similarly, we obtained $2 \log BF_{0,3} > 10$, again suggesting a decisive preference for model M_0 over M_3. When comparing M_0 and M_2 we obtain $2 \log BF_{0,2} = 7.30$, still showing a strong evidence favoring model M_0.

4.2 *Data analysis*

We fit the hierarchical model from Section 3 to the data from the three studies. Figure 2 shows 95% credible intervals for some subject-specific parameters. The left panel displays change-points in years prior to prostate cancer diagnosis and indicates that patients with metastatic disease are likely to have transitions occurring later in time (i.e. closer to disease diagnosis) as opposed to patients with local disease. A larger variability is, however, usually associated with the change-point in patients with local disease. The right panel of Figure 2 shows that the growth rates of PSA post change-points are somewhat higher for patients with metastatic disease than those with local disease. This plot also highlights between-study differences; the results suggest that post-change-point slopes for metastatic patients were highest for the CARET study, and considerably lower for the other two studies. This may be due to between-study differences in the definition of localized and metastatic disease. Figure 3 shows the posterior medians of the subject-specific change-points *versus* the posterior medians of the subject-specific post change-points slopes, indicating that in patients with metastatic prostate cancer, transitions occur at times closer to diagnosis and that the transition is followed by a higher growth rate. Though this plot ignores the variability in the posterior distributions, it shows a relatively good separation of the two groups of patients.

Figure 4 displays the 95% posterior predictive credibility intervals of PSA at patients' change-points. The horizontal dotted line displays the corresponding threshold value of 4.0 ng ml^{-1}. This value has been

Downloaded from https://academic.oup.com/biostatistics/article/5/3/483/310204 by guest on 21 August 2022

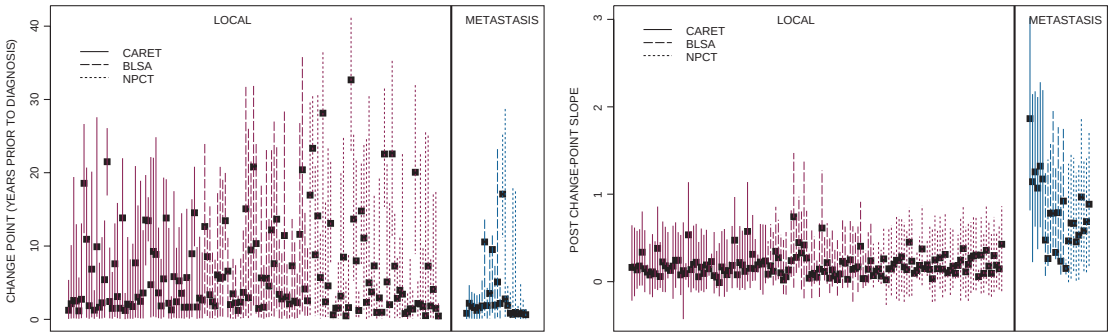

Fig. 2. 95% Posterior credibility intervals for subject-specific parameters by study and disease stage. The posterior medians are shown with small boxes. Left panel: Subject-specific change-points τ_i. Right panel: Subject-specific post change-point slopes b_{2i}^s.

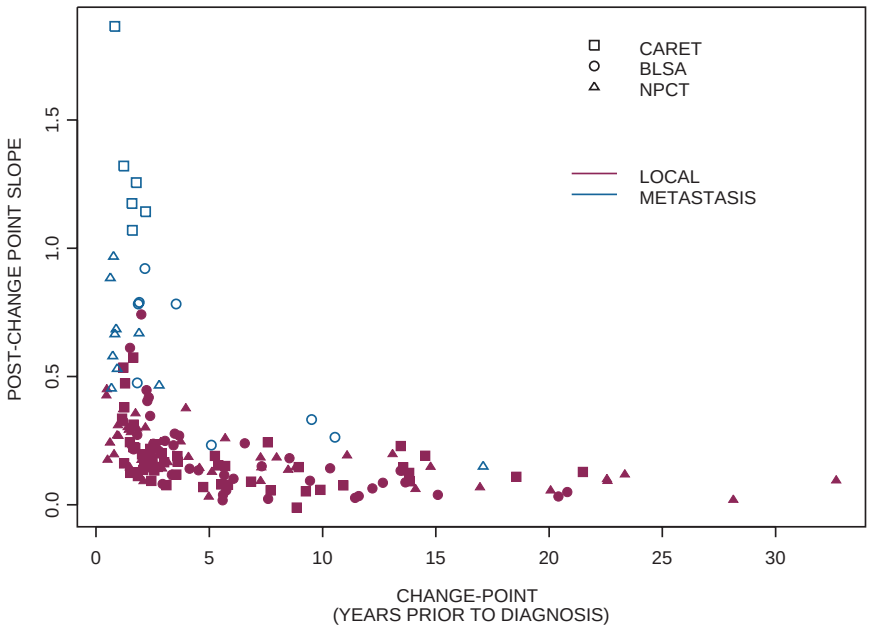

Fig. 3. Posterior medians of subject-specific change points *versus* posterior medians of subject-specific post-change point slopes by study and disease stage. Filled symbols denote local stage while open symbols denote metastasis.

used in clinical practice as indicative of a possible prostate cancer disease. This figure shows that most patients experience a transition at levels of PSA lower than the clinical threshold.

Posterior distributions of the study-level, post-change-point slopes for patients with local disease (that is, $\beta_2^s, s = 1, 2, 3$) are very similar between studies. However, for patients with metastatic disease (that is, $\beta_2^s + \gamma^s$) the posterior distributions show a much larger variability; moreover, between-study differences are clearly apparent. The posterior distributions of the population parameters β and γ place a large probability mass on the positive range (probabilities equal to 0.99 and 0.93, respectively). Table 2 summarizes the posterior distributions of the study-specific and population parameters.

Downloaded from https://academic.oup.com/biostatistics/article/5/3/483/310204 by guest on 21 August 2022

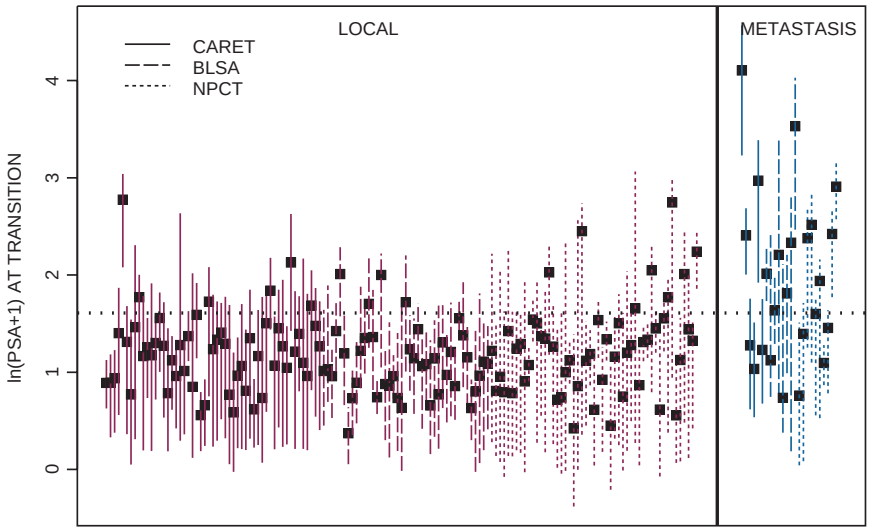

Fig. 4. Predictive subject-specific values of $\ln(PSA+1)$ levels at change-points by study and disease stage.

We chose to fit our models to $\log(PSA+1)$; this has been the scale used by others in analysis of PSA data, presumably to reduce skewness while avoiding indeterminate values (Morrell *et al.*, 1995; Slate and Clark, 1999; Slate and Turnbull, 2000). For interpretation purposes, $\log(PSA)$ may be preferable, because its inverse transformation provides estimated relative growth rates in the original measurement scale (i.e. as the annual percentage change in PSA). To estimate these relative growth rates, we fit our model to the response $y = \log(PSA)$, with zero values ($n = 10$) being replaced by the minimum value of PSA in the dataset. We found that PSA increases by an estimated 2.24% per year prior to the change-point; this is similar to the estimate of 3.2% from Oesterling *et al.* (1993) which was derived from a population without a prostate cancer diagnosis. After the change-point, PSA increases by 15.31% per year for patients with local disease, and by 63.09% per year for patients with advanced metastatic disease.

The posterior predictive distribution of age at transition is shown in the left panel of Figure 5. The posterior predictive mode of the distribution is at age 57.25 (the lower and upper boundaries of the 95% credibility intervals are 42.90 and 91.06 years, respectively). The right panel of Figure 5 shows the predictive PSA trajectory for a new patient by disease prognosis.

Figure 6 shows the posterior predictive PSA trajectories for a new patient within each of the studies. The full lines show the posterior predictive median trajectory, while the dotted lines show the boundaries of the 95% credibility intervals.

Figure 7 shows the 95% posterior predictive PSA trajectories for some randomly selected patients. Left side panels correspond to trajectories for some of the local patients, while right side panels correspond to trajectories for some of the metastatic patients in our combined dataset.

5. Discussion

In this analysis we used a non-linear Bayesian hierarchical model to combine longitudinal data on PSA growth from three different studies. Our model accounts for the study-to-study variability as well as the variability between and within patients. In the first level of the hierarchy—the individual level—we considered a change-point model to accommodate the non-linearity of PSA growth in cancer patients. We then incorporated information on disease stage at the second level of the hierarchical model. Our analysis

Downloaded from https://academic.oup.com/biostatistics/article/5/3/483/310204 by guest on 21 August 2022

Table 2. *Posterior estimates of the population and study specific growth parameters*

	Parameter	Posterior median	95% Credibility interval
POPULATION	β_0	0.26	(−0.52, 1.06)
	β_1	0.02	(−0.07, 0.09)
	β_2	0.21	(0.08, 0.35)
	γ	0.64	(−0.60, 1.89)
CARET	β_0^1	0.17	(−0.08, 0.41)
	β_1^1	0.02	(0.02, 0.03)
	β_2^1	0.20	(0.12, 0.30)
	γ^1	1.11	(0.51, 1.86)
BLSA	β_0^2	0.02	(−0.05, 0.08)
	β_1^2	0.02	(0.02, 0.03)
	β_2^2	0.21	(0.13, 0.33)
	γ^2	0.41	(0.08, 1.01)
NPCT	β_0^3	0.60	(0.56, 0.65)
	β_1^3	0.01	(0.00, 0.01)
	β_2^3	0.21	(0.13, 0.32)
	γ^3	0.40	(0.01, 0.95)

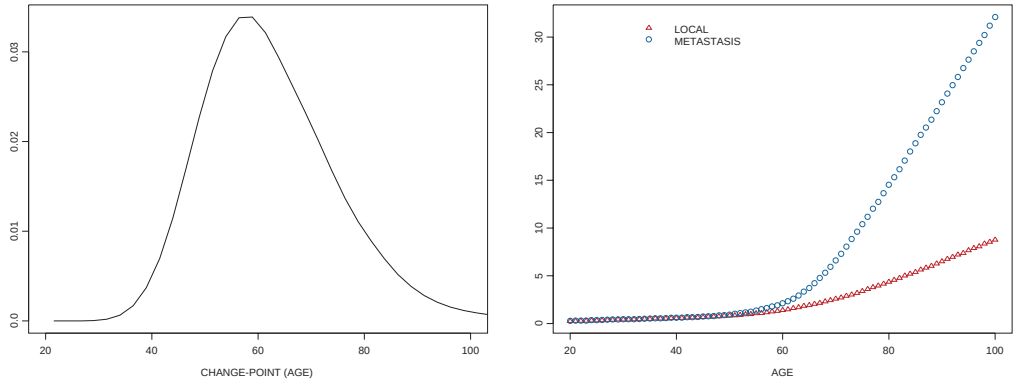

Fig. 5. Predictive distributions. Left panel: Predictive distribution of the age at transition. Right panel: Predictive PSA trajectory for a new patient.

allowed us to combine information across studies to obtain estimates of PSA growth that are more reliable than those based on a single study, while also identifying between-study differences.

The combination of data across studies also provided us with some key insights that are highly relevant from a clinical perspective. First, we found that for most patients, the change-point occurs at PSA levels below 4.0 ng ml^{-1}, which is the standard threshold for a positive test. Assuming that the change-point does indeed represent a biological transition from a healthy to a neoplastic state, this finding is in concurrence with recent studies that have observed a non-trivial frequency of prostate cancer cases with PSA levels below this standard threshold (e.g. Punglia *et al.*, 2003). Second, our comparison of PSA growth patterns between cases with localized and metastatic disease at diagnosis suggested that the PSA growth rates after

Downloaded from https://academic.oup.com/biostatistics/article/5/3/483/310204 by guest on 21 August 2022

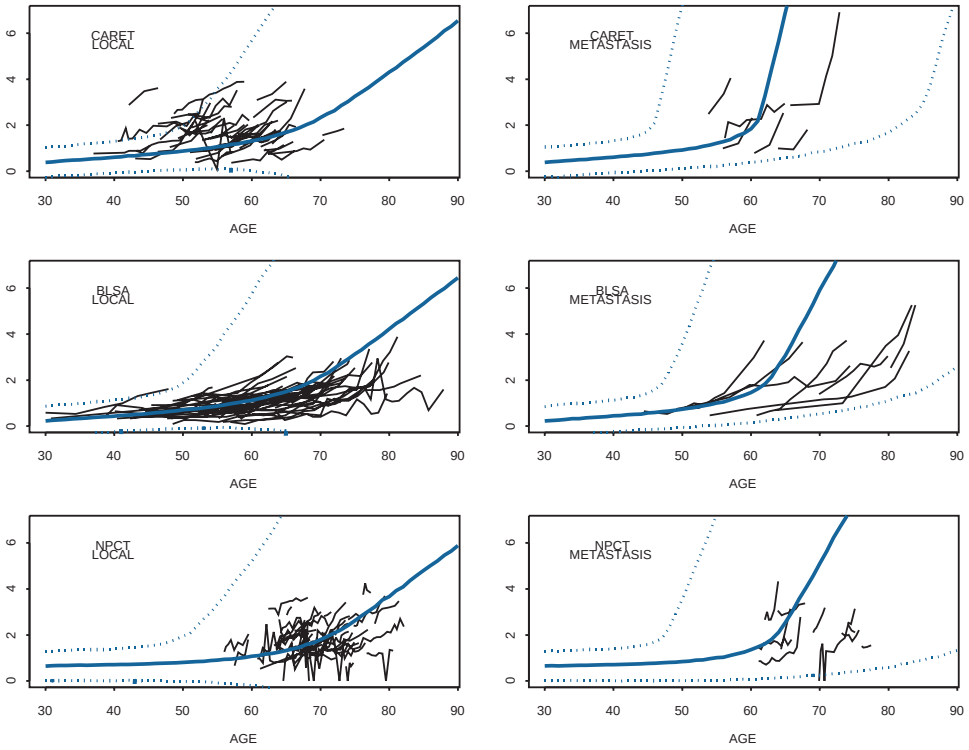

Fig. 6. Predictive PSA trajectories for a new patient in each study.

a transition point differ for localized and metastatic cases. In particular, the post-change-point growth rates appear to be considerably higher on average for cases destined to be diagnosed with metastatic disease. This finding has clear implications for determining prognosis among newly diagnosed cases, suggesting that patients with high PSA are more likely to be candidates for aggressive therapy than those with slowly increasing levels of PSA. Moreover, it also has important implications for early detection, because it suggests that metastatic disease may be biologically different to localized disease. The basic premise underlying early detection is that metastatic disease diagnosed while still localized is as treatable as localized disease. However, if metastatic disease is biologically different to localized disease from the time of disease onset, this premise may not always hold. It is possible that our higher growth rates among metastatic cases could reflect the presence of a second change-point with accelerated growth after that time. We considered a two-change-point model, but found our estimates under this model to be highly variable due to the limited number of metastatic cases in the dataset.

Our analysis compared cases staged clinically as localized or metastatic. It is known, however, that many cancers are clinically under-staged. Partin *et al.* (1997) report that of approximately 60% of men believed to have organ-confined disease and who have surgery, fewer than 50% of them are confirmed to have local disease on final pathological analysis. Using pathological stage instead of clinical stage could improve the estimation of PSA growth. However, information on pathological stage is only available for patients who undergo surgery and this information is not available for most of our patients. Our models could be expanded to deal with missing information on pathological stage or, alternatively, mis-classification of clinical stage. However, the results shown in Figure 3 do not indicate clear mis-classification patterns.

Downloaded from https://academic.oup.com/biostatistics/article/5/3/483/310204 by guest on 21 August 2022

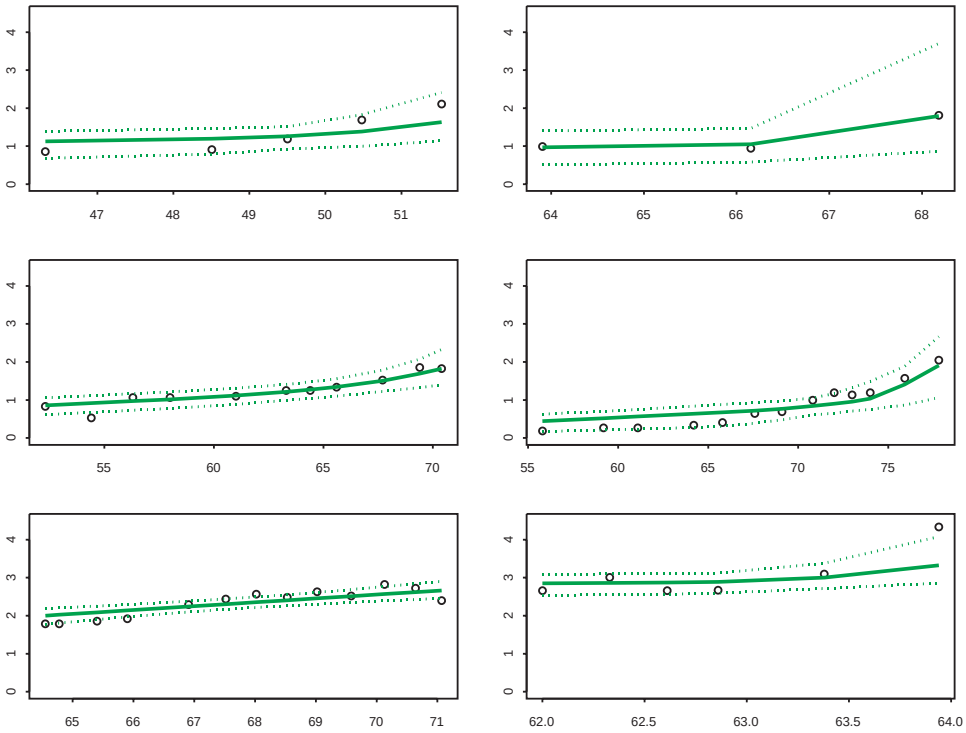

Fig. 7. Fitted trajectories for some randomly selected patients.

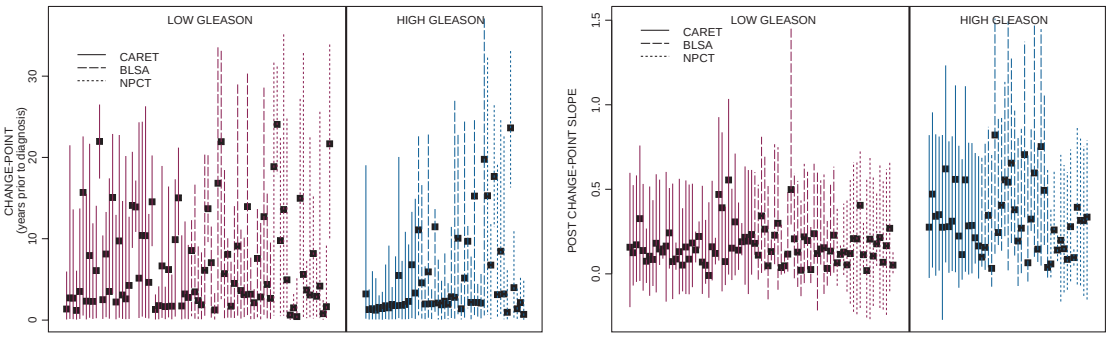

Fig. 8. 95% Posterior credibility intervals for subject-specific parameters by study and grade of disease. Left panel: Subject-specific change-points τ_i. Right panel: Subject-specific post change-point slopes b_{2i}^s.

As an alternative to stage, one could consider other prognostic variables for studying PSA growth such as the Gleason score. It is known that prostate cancer patients with low Gleason score (less than or equal to 6) have a considerably better disease prognosis than those with high Gleason score. Using our hierarchical Bayesian change-point model to investigate PSA growth by Gleason score we found that patients with a high Gleason score tended to exhibit higher post-change-point slopes than those with low Gleason score, but the difference in post-change-point slopes between cases with low- and high-grade disease was not as extreme as the difference observed between cases with early- and late-stage disease (see Figures 8 and 9).

While our analysis yields important new insights into the natural history of prostate cancer and PSA,

Downloaded from https://academic.oup.com/biostatistics/article/5/3/483/310204 by guest on 21 August 2022

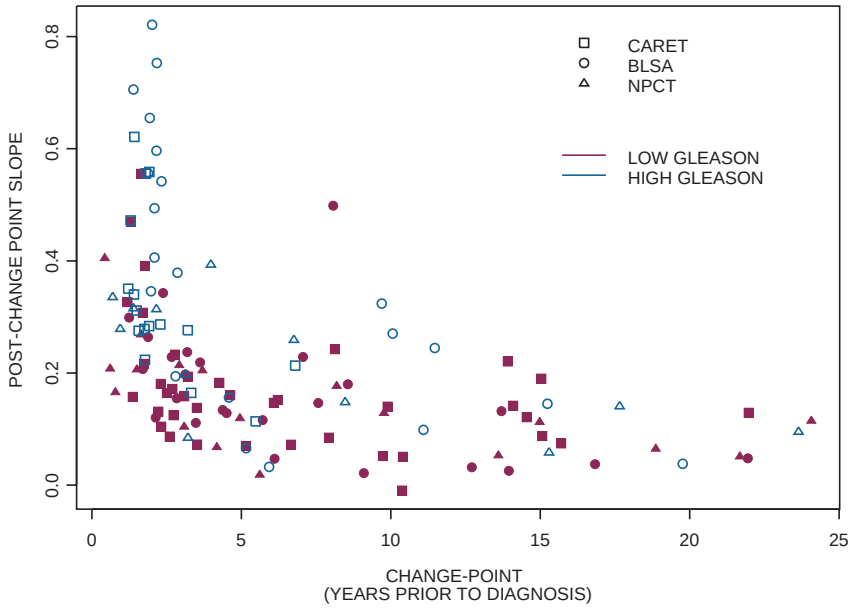

Fig. 9. Posterior medians of subject-specific change points *versus* posterior medians of subject-specific post-change point slopes by study and grade of disease. Filled symbols denote low Gleason while open symbols denote high Gleason.

it is subject to some limitations. The Bayesian hierarchical framework is based on a parametric model and imposes constraints on the correlation structure among observations from the same individual or study. However, such models also allow detailed estimates of within- and between-individual variation—which are needed for simulation modeling of prostate cancer. Moreover, the ability of hierarchical modeling to provide population-level parameter estimates with any degree of precision is limited by the number of studies (regardless of the estimation technique used). In the present analysis, data were available from three studies, which resulted in population estimates that were considerably less precise that the study-specific estimates. Consequently, we have presented many of our results by study which has also enabled us to understand study-to-study variability. We hope in the future to add at least one more study to the meta-analysis.

In conclusion, the Bayesian hierarchical framework has enabled us to combine data on PSA from multiple studies, yielding novel insights about the disease process in prostate cancer, as reflected by PSA growth. Findings from the current study provide insight into the evolving natural history of a patient's prostate cancer and may allow clinicians to more accurately determine which patients are more likely to have indolent disease and which have more biologically aggressive tumors. This, in turn, may result in more selective application of aggressive therapy, restricting the risk of adverse outcomes to men who are most likely to benefit from these treatments, and, therefore, improving quality of life for all men with prostate cancer.

Acknowledgements

The authors wish to thank the suggestions and comments by the editors, associate editor and reviewer. This work was supported by grant 5 U01 CA 88160 from the National Cancer Institute. E. H. Slate also acknowledges partial support from NIH grants 5 U01 CA079080-05, 1 R01 CA077789-04, DOD grant

Downloaded from https://academic.oup.com/biostatistics/article/5/3/483/310204 by guest on 21 August 2022

DAMD17-02-1-0138, the NIH COBRE grant 1 P20 RR 17696-01, and the USC-MUSC Collaborative research program.

REFERENCES

BLUTE, M. L., BERGSTRALH, E. J., PARTIN, A. W., WALSH, P. C., KATTAN, M. W., SCARDINO, P. T., MONTIE, J. E., PEARSON, J. D., SLEZAK, J. M. AND ZINCKE, H. (2000). Validation of partin tables for predicting pathological stage of clinically localized prostate cancer. *Journal of Urology* **164**, 1591–1595.

BOOTH, N. B. AND SMITH, A. F. M. (1982). A Bayesian approach to retrospective identification of change-points. *Journal of Econometrics* **19**, 7–22.

BRANT, L. J., SHENG, S. L., MORRELL, C. H., VERBEKE, G. N., LESAFFRE, E. AND CARTER, H. B. (2003). Screening for prostate cancer using random-effects models. *Journal of Royal Statistical Society: Series A* **166**, 51–62.

BRAWER, M. K., BANKSON, D. D., HAVER, V. M. AND PETTEWAY, J. C. (1997). Comparison of three commercial PSA assays: results of restandardization of the Ciba Corning method. *The Prostate* **30**, 269–273.

CARLIN, B. P., GELFAND, A. E. AND SMITH, A. F. M. (1992). Hierarchical Bayesian analysis of changepoint-problems. *Applied Statistics* **41**, 389–405.

CARTER, H. B., PEARSON, J. D., METTER, E. J., BRANT, L. J., CHAN, D. W., ANDRES, R., FOZARD, J. L. AND WALSH, P. C. (1992a). Longitudinal evaluation of prostate-specific antigen levels in men with and without prostate disease. *Journal of the American Medical Association* **267**, 2215–2220.

CARTER, H. B., MORRELL, C. H., PEARSON, J. D., BRANT, L. J., PLATO, C. C., METTER, E. J., CHAN, D. W., FOZARD, J. L. AND WALSH, P. C. (1992b). Estimation of prostatic growth using serial prostate-specific antigen measurements in men with and without prostate disease. *Cancer Research* **52**, 2232–3328.

CARTER, H. B., EPSTEIN, J. I., CHAN, D. W., FOZARD, J. L. AND PEARSON, J. D. (1997). Recommended prostate-specific antigen testing intervals for the detection of curable prostate cancer. *Journal of the American Medical Association* **277**, 1456–1460.

CLARK, L. C., COMBS, G. F., TURNBULL, B. W., SLATE, E. H., CHALKER, D. K., CHOW, J., DAVIS, L., GLOVER, R. A., GRAHAM, G. F., GROSS, E. G., KRONGRAD, A., LESHER, J. L., PARK, H. K., SANDERS, B. B., SMITH, C. L. AND TAYLOR, J. R. (1996). Effects of selenium supplementation for cancer prevention in patients with carcinoma of the skin: a randomized controlled trial. *Journal of the American Medical Association* **276**, 1957–1963.

DRAISMA, G., BOER, R., OTTO, S. J., VAN DER CRUIJSEN, I. W., DAMHUIS, R. A., SCHRODER, F. H. AND DE KONING, H. J. (2003). Lead times and overdetection due to prostate-specific antigen screening: estimates from the European randomized study of screening for prostate cancer. *Journal of the National Cancer Institute* **95**, 868–878.

DUFFIELD-LILLICO, A. J., SLATE, E. H., REID, M. E., TURNBULL, B. W., WILKINS, P. A., COMBS, G. F., PARK, H. K., GROSS, E. G., GRAHAM, G. F., STRATTON, M. S., MARSHALL, J. R. AND CLARK, L. C. (2003). Selenium supplementation and secondary prevention of nonmelanoma skin cancer in a randomized trial. *Journal of the National Cancer Institute* **95**, 1477–1481.

DUMOUCHEL, W. H. AND HARRIS, J. E. (1983). Bayes methods for combining the results of cancer studies in humans and other species with discussion. *Journal of the American Statistical Association* **78**, 293–315.

EASTHAM, J. A., RIEDEL, E., SCARDINO, P. T., SHIKE, M., FLEISHER, M., SCHATZKIN, A., LANZA, E., LATKANY, L. AND BEGG, C. B. (2003). Variation of serum prostate-specific antigen levels—an evaluation of year-to-year fluctuations. *Journal of the American Medical Association* **289**, 2695–2700.

ELLIS, W. J., ETZIONI, R., VESSELLA, R. L., HU, C. AND GOODMAN, G. E. (2001). Serial prostate specific antigen, free-to-total prostate specific antigen ratio and complexed prostate specific antigen for the diagnosis of prostate cancer. *Journal of Urology* **166**, 93–98,. discussion 98–99.

Downloaded from https://academic.oup.com/biostatistics/article/5/3/483/310204 by guest on 21 August 2022

ETZIONI, R., CHA, R. AND COWEN, M. E. (1999). Serial prostate specific antigen screening for prostate cancer: a computer model evaluates competing strategies. *Journal of Urology* **162**, 741–748.

ETZIONI, R., PENSON, D. F., LEGLER, J. M., DI TOMMASO, D., BOER, R., GANN, P. H. AND FEUER, E. J. (2002a). Overdiagnosis due to prostate-specific antigen screening: lessons from U.S. prostate cancer incidence trends. *Journal of the National Cancer Institute* **94**, 981–990.

ETZIONI, R., BERRY, K. M., LEGLER, J. M. AND SHAW, P. (2002b). Prostate-specific antigen testing in black and white men: an analysis of medicare claims from 1991–1998. *Urology* **59**, 251–255.

FRANKEL, S., SMITH, G. D., DONOVAN, J. AND NEAL, D. (2003). Screening for prostate cancer. *The Lancet* **361**, 1122–1128.

FRYDENBERG, M., STRICKER, P. D. AND KAYE, K. W. (1997). Prostate cancer diagnosis and management. *The Lancet* **349**, 1681–1687.

GELFAND, A. E. AND SMITH, A. F. M. (1990). Sampling-based approaches to calculating marginal densities. *Journal of the American Medical Association* **85**, 398–409.

GEMAN, S. AND GEMAN, D. (1984). Stochastic relaxation, Gibbs distributions, and the Bayesian restoration to images. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **6**, 721–741.

JEMAL, A., MURRAY, T., SAMUELS, A., GHAFOOR, A., WARD, E. AND THUN, M. J. (2003). Cancer statistics, 2003. *CA: A Cancer Journal for Clinicians* **53**, 5–26.

KABALIN, J. N., MCNEAL, J. E., JOHNSTONE, I. M. AND STAMEY, T. A. (1995). Serum prostate-specific antigen and the biologic progression of prostate cancer. *Urology* **46**, 65–70.

KASS, R. E. AND RAFTERY, A. E. (1995). Bayes factors. *Journal of the American Statistical Association* **90**, 773–795.

LEEWANSANGTONG, S., GOKTAS, S., LEPOFF, R., HOLTHAUS, K. AND CRAWFORD, E. D. (1998). Comparability of serum prostate-specific antigen measurement between the hybritech Tandem-R and Abbott AxSym Assays. *Urology* **52**, 467–469.

METROPOLIS, N., ROSENBLUTH, A. W., ROSENBLUTH, M. N., TELLER, A. H. AND TELLER, E. (1953). Equations of state calculations by fast computing machine. *Journal of Chemistry in Physics* **21**, 1087–1091.

MORRELL, C. H., PEARSON, J. D., CARTER, H. B. AND BRANT, L. J. (1995). Estimating unknown transition times using a piecewise nonlinear mixed-effects model in men with prostate cancer. *Journal of the American Statistical Association* **90**, 45–53.

OESTERLING, J. E., JACOBSEN, S. J., CHUTE, C. G., GUESS, H. A., GIRMAN, C. J., PANSER, L. A. AND LIEBER, M. M. (1993). Serum prostate-specific antigen in a community-based population of healthy men. Establishment of age-specific reference ranges. *Journal of the American Medical Association* **270**, 860–864.

OESTERLING, J. E., MOYAD, M. A., WRIGHT, G. L. AND BECK, G. R. (1995). An analytical comparison of the three most commonly used prostate-specific antigen assays: Tandem-R, Tandem-E and IMx. *Urology* **46**, 524–532.

PARTIN, A. W., MANGOLD, L. A., LAMM, D. M., WALSH, P. C., EPSTEIN, J. I. AND PEARSON, J. D. (2001). Contemporary update of prostate cancer staging nomograms (Partin Tables) for the new millennium. *Urology* **58**, 843–848.

PARTIN, A. W., KATTAN, M. W., SUBONG, E. N. P., WALSH, P. C., WOJNO, K. J., OESTERLING, J. E., SCARDINO, P. T. AND PEARSON, J. D. (1997). Combination of prostate-specific antigen, clinical stage, and gleason score to predict pathological stage of localized prostate cancer. *Journal of the American Medical Association* **277**, 1445–1451.

PEARSON, J. D., MORRELL, C. H., LANDIS, P. K., CARTER, H. B. AND BRANT, L. J. (1994). Mixed-effects regression models for studying the natural history of prostate disease. *Statistics in Medicine* **13**, 587–601.

PENSON, D. F., LITWIN, M. S. AND AARONSON, N. K. (2003). Health related quality of life in men with prostate cancer. *Journal of Urology* **169**, 1653–1661.

Downloaded from https://academic.oup.com/biostatistics/article/5/3/483/310204 by guest on 21 August 2022

POTOSKY, A. L., LEGLER, J., ALBERTSEN, P. C., STANFORD, J. L., GILLILAND, F. D., HAMILTON, A. S., ELEY, J. W., STEPHENSON, R. A. AND HARLAN, L. C. (2000). Health outcomes after prostatectomy or radiotherapy for prostate cancer: results from the Prostate Cancer Outcomes Study. *Journal of the National Cancer Institute* **92**, 1582–1592.

POTOSKY, A. L., REEVE, B. B., CLEGG, L. X., HOFFMAN, R. M., STEPHENSON, R. A., ALBERTSEN, P. C., GILLILAND, F. D. AND STANFORD, J. L. (2002). Quality of life following localized prostate cancer treated initially with androgen deprivation therapy or no therapy. *Journal of the National Cancer Institute* **94**, 430–437.

PUNGLIA, R. S., D'AMICO, A. V., CATALONA, W. J., ROEHL, K. A. AND KUNTZ, K. M. (2003). Effect of verification bias on screening for prostate cancer by measurement of prostate-specific antigen. *The New England Journal of Medicine* **349**, 335–342.

ROSS, K. S., CARTER, H. B., PEARSON, J. D. AND GUESS, H. A. (2000). Comparative efficiency of prostate-specific antigen screening strategies for prostate cancer detection. *Journal of the American Medical Association* **284**, 1399–1405.

SHOCK, N. W., GREULICH, R. C., ANDRES, R., LAKATTA, E. G., ARENBERG, D. AND TOBIN, J. D. (1984). *Normal Human Aging: The Baltimore Longitudinal Study of Aging*, NIH Publication No. 84-2450. Washington, D.C.: U.S. Government Printing Office.

SILVERMAN, B. W. (1986). *Density Estimation for Statistics and Data Analysis*. London: Chapman and Hall.

SKATES, S. J., PAULER, D. K. AND JACOBS, I. J. (2001). Screening based on the risk of cancer calculation from Bayesian hierarchical changepoint and mixture models of longitudinal markers. *Journal of the American Statistical Association* **96**, 429–439.

SLATE, E. H. AND CLARK, L. C. (1999). Using PSA to detect prostate cancer onset: an application of Bayesian retrospective and prospective change-point identification. In *Case Studies in Bayesian Statistics* IV, New York: Springer, pp. 511–533.

SLATE, E. H. AND CRONIN, K. A. (1997). Change-point modeling of longitudinal PSA as a biomarker for prostate cancer. In *Case Studies in Bayesian Statistics* III, New York: Springer, pp. 435–456.

SLATE, E. H. AND TURNBULL, B. W. (2000). Statistical models for longitudinal biomarkers of disease onset. *Statistics in Medicine* **19**, 617–637.

STAMEY, T. A., YANG, N., HAY, A. R., MACNEAL, J. E., FREIHA, F. S. AND REDWINE, E. (1987). Prostate specific antigen as a serum marker for adenocarcinoma of the prostate. *New England Journal of Medcine* **317**, 909–916.

STANGL, D. K. AND BERRY, D. A. (2000). *Meta-Analysis in Medicine and Health Policy*. New York: Dekker.

STEPHENS, D. A. (1994). Bayesian retrospective multiple change-point identification. *Applied Statistics* **43**, 159–178.

SULLIVAN, M. P., ETZIONI, R., FENG, Z., POTTER, J. D., THOMPSON, M. L., THORNQUIST, M., WINGET, M. AND YASUI, Y. (2001). Phases of biomarker development for early detection of cancer. *Journal of the National Cancer Institute* **93**, 1054–1061.

TANNER, M. A. (1996). *Tools for Statistical Inference*, 3rd Edition. New York: Springer.

WHITTEMORE, A. S., LELE, C., FRIEDMANN, G. C., STAMEY, T., VOGELMAN, J. J. AND ORENTREICH, N. (1995). Prostate-specific antigen as predictor of prostate cancer in black man and white mean. *Journal of the National Cancer Institute* **87**, 354–360.

[*Received 14 October 2003; revised 5 February 2004; accepted for publication 18 February 2004*]

Downloaded from https://academic.oup.com/biostatistics/article/5/3/483/310204 by guest on 21 August 2022