

To appear in the *Journal of Applied Statistics*  
Vol. 00, No. 00, Month 20XX, 1–15

## Frequentist and Bayesian approaches for a joint model for prostate cancer risk and longitudinal PSA data

Carles Serrat<sup>a,\*</sup>, Montserrat Rué<sup>b,c</sup>, Carmen Armero<sup>d</sup>, Xavier Piulachs<sup>e</sup>, Hèctor Perpiñán<sup>d</sup>, Anabel Forte<sup>f</sup>, Álvaro Páez<sup>g,h</sup> and Guadalupe Gómez<sup>i</sup>

<sup>a</sup>*Dept. of Applied Mathematics I, Universitat Politècnica de Catalunya-BarcelonaTECH, Avda. Dr. Marañón, 44–50, 08028-Barcelona, Spain;* <sup>b</sup>*Dept. of Basic Medical Sciences, Institut de Recerca Biomèdica de Lleida, Universitat de Lleida, Avda. Rovira Roure, 80, 25198-Lleida, Spain;* <sup>c</sup>*Health Services Research Network in Chronic Diseases (REDISSEC);* <sup>d</sup>*Dept. of Statistics and OR, Universitat de València, Doctor Moliner, 50, 46100-Burjassot, Spain;* <sup>e</sup>*Dept. of Econometrics. Riskcenter-IREA, Universitat de Barcelona, Avda. Diagonal, 690, 08034-Barcelona, Spain;* <sup>f</sup>*Dept. of Economics, Universitat Jaume I, Avda. de Vicent Sos Baynat, s/n, 12071-Castelló de la Plana, Spain;* <sup>g</sup>*Urology Dept., Hospital Universitario de Fuenlabrada, Camino del Molino, 2, 28942-Fuenlabrada, Madrid, Spain;* <sup>h</sup>*Spanish Branch of the European Randomized Study of Screening for Prostate Cancer;* <sup>i</sup>*Dept. of Statistics and OR, Universitat Politècnica de Catalunya-BarcelonaTECH, Jordi Girona, 1–3, 08034-Barcelona, Spain*

(Original Manuscript May 2014)

The paper describes the use of frequentist and Bayesian shared-parameter joint models of longitudinal measurements of prostate specific antigen (PSA) and the risk of prostate cancer (PCa). The motivating dataset corresponds to the screening arm of the Spanish branch of the European Randomized Screening for Prostate Cancer (ERSPC) study. The results show that PSA is highly associated with the risk of being diagnosed with PCa and that there is an age-varying effect of PSA on PCa risk. Both the frequentist and Bayesian paradigms produced very close parameter estimates and subsequent 95% confidence and credibility intervals. Dynamic estimations of disease-free probabilities obtained using Bayesian inference highlight the potential of joint models to guide personalized risk-based screening strategies.

**Keywords:** joint models; linear mixed models; prostate cancer screening; relative risk models; shared-parameter

*Classification codes:* 62N01; 62P10

### 1. Introduction

Joint models for longitudinal and time-to-event data are increasingly used to assess relationships between serial measurements of markers and time to an event of interest. Joint models were introduced during the 90s [7, 33, 37] and since then have been applied to a great variety of studies in epidemiological and biomedical areas. In turn, these studies have fed a wide methodological research on the subject, with models focused on event times, longitudinal patterns or both ([34] and [19] are excellent reviews up to date).

In a setting of dependent longitudinal and time-to-event data, shared-parameter models consider that the longitudinal and survival processes depend jointly on a common set of random effects. Given the random effects, the two processes are assumed independent. Recently, Rizopoulos has made a great contribution facilitating the use of the joint modeling methodology,

---

\*Corresponding author: e-mail: carles.serrat@upc.edu, Phone: +34 93 405 4486, Fax: +34 93 401 6306

first by means of an overview of the theory and applications of joint modeling [26] and secondly by developing the `JM` [23] and `JMbayes` [27] R packages for the frequentist and Bayesian shared-effects' approaches, respectively.

In statistical science, the long-lasting dissension between frequentists and Bayesians is evolving into a compromise that capitalizes on the strengths of both systems of inference [16]. In fact, it seems that an integrated vision of both methodologies, from both the frequentist [14] and Bayesian worlds [2, 5], is in progress. Joint modeling illustrates how classical and Bayesian statistics can be combined to reach a complex goal, taking advantage of the strengths of both approaches.

Longitudinal and time-to-event data have been combined in different studies of prostate cancer (PCa). Some authors have shown the usefulness of longitudinal measures of the prostate specific antigen (PSA), a protein produced by cells of the prostate gland, to study the natural history of prostate cancer (PCa) [3, 13, 15]. PCa has a highly variable natural history ranging from indolent to very aggressive, with a long preclinical period. Although controversial, opportunistic screening with PSA, has been widely used in Western countries [6]. Two facts have raised questions about the present PSA-based screening paradigm: a) Conflicting results from the two largest clinical trials assessing the effectiveness of PCa screening [1, 30]; and b) Significant harms associated to screening that include biopsy complications and, more seriously, overdiagnosis and overtreatment of indolent tumors. Current guidelines recommend against universal PCa screening and propose risk-based or personalized screening [4]. Joint models could serve this purpose by dynamically updating the estimated probabilities for the event time outcome as additional information is collected for each subject.

This paper has a double aim: on the one hand, to show how joint models can be used to incorporate past and current PSA levels into a predictive model of PCa. On the other hand, to illustrate how the frequentist and Bayesian approaches can complement each other. The rest of the paper is organised as follows. Section 2 presents a description of the motivating dataset, the Spanish screening arm of the European Randomized Screening for Prostate Cancer study (ERSPC). Section 3 contains the formulation, computational details and results of the joint model under a frequentist and Bayesian approach, respectively. Section 4 illustrates how to estimate dynamic disease-free probabilities. Finally, Section 5 contains a discussion focused on the impact of longitudinal PSA values and baseline covariates on PCa risk and also on the complementarity of frequentist and Bayesian reasoning in joint modeling.

## 2. Motivation ERSPC study and data description

Our motivating dataset consists of the screening arm of the Spanish subset of the European Randomized Screening for Prostate Cancer (ERSPC) study [17]. The ERSPC study [30] is a multicenter trial initiated in 1991 in the Netherlands and Belgium, with six more European countries joining during the 90s. The objective of the ERSPC was to evaluate whether PCa screening with PSA reduces PCa specific mortality in asymptomatic men. The Spanish subset recruited 4278 individuals between February 1996 and June 1999, aged 45 to 71 years, of whom 2415 were assigned to the screening arm and the remaining to the control arm. Those screened were given a blood test to measure PSA levels.

Serum PSA determinations, for individuals assigned to the screening arm of the study, were mostly performed at an interval of 4 years. A biopsy was indicated if  $PSA \geq 3$  ng/ml, and in case of a negative result the subject was recalled for an early visit within two years. Among the 2415 participants, a total of 4673 follow-up PSA measurements at the successive screening rounds were recorded. Screening ended in October 2005 but active surveillance of PCa incidence took place until December 31st, 2007 which corresponds to administrative closure.

The minimum age required to enter the study was 45 years. Age at randomization ranged from

Table 1. Distribution of the subjects in the screening arm of the Spanish ERSPC study, descriptive statistics of the PSA measurements and distribution of the PCa diagnosed cases, stratified by the number of visits and overall.

	Number of visits per subject				Overall
	1	2	3	$\geq 4$	
Sample size (%)	573 (23.7)	1499 (62.1)	293 (12.1)	50 (2.1)	2415 (100.0)
PSA descriptive					
Mean	2.40	1.40	1.92	5.46	1.78
StDev	5.67	1.67	2.25	2.79	3.20
Min	0.00	0.00	0.00	0.50	0.00
1st Q	0.60	0.60	0.60	3.79	0.63
Median	1.10	1.00	1.10	4.69	1.06
3rd Q	2.20	1.67	2.44	6.28	1.95
Max	68.90	42.30	20.36	15.53	68.90
PCa diagnosed (%)	51 (44.0)	50 (43.1)	13 (11.2)	2 (1.7)	116 (100.0)

45.5 to 71.0 years, with a mean and median of 57.7 and 56.9 yr, respectively. Concerning the number of PSA measurements per subject, Table 1 shows that the most frequent group (62.1%) corresponds to men with two PSA measurements and that nearly one fourth of subjects (23.7%) contributed to the analysis with only one PSA observation. The variability of PSA levels within the latter group is the highest (StDev = 5.67 ng/ml and Max = 68.90 ng/ml). Note that the 50 subjects with four or more follow-up visits had PSA values much higher than the rest of the subjects, with median and third quartile equal to 4.69 and 6.28 ng/ml, respectively. With regard to PCa diagnosed cases, 87.1% correspond to individuals with one or two follow-up visits.

The variable of interest in the longitudinal analysis is PSA, ranging from 0.00 to 68.90 ng/ml with mean and median 1.78 and 1.06 ng/ml., respectively (see Table 1). In order to account for the non-normality of the data, PSA measurements in this paper were considered in a double logarithmic transformed scale. We defined the *LLPSA* variable as the  $\log(1+\log(1+PSA))$ . Figure 1 shows the time-plots of *LLPSA* according to presence or absence of PCa diagnosis, where the consecutive observations on each man are connected through line segments. The top plot shows the *LLPSA* trajectories for the 116 men diagnosed with PCa and the bottom one shows a random sample of 116 men for the not-diagnosed group. Interesting features in both graphics are the wide variation in the initial values of *LLPSA* thus indicating different intercepts for the different individuals and a general, which few exceptions, positive trend with a gentle slope, as well as a potential mildly nonlinear component. However, the PCa group had higher mean values of *LLPSA*, in comparison with the non-PCa group. This information gives clear support to the relevance of PSA measurements to explain time to PCa diagnosis.

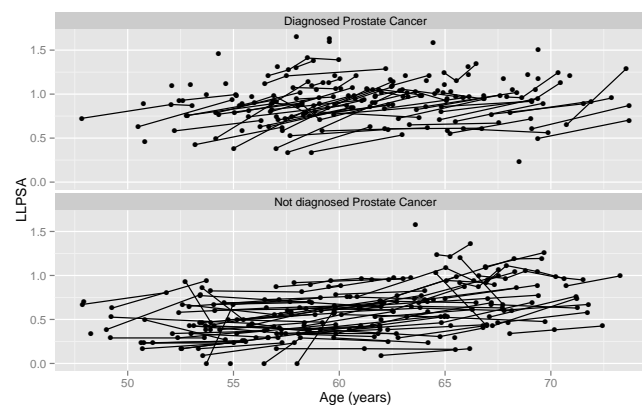


Figure 1. *LLPSA* profiles across time (age in years) for men who developed PCa during the follow-up period (top panel) and for 116 randomly selected men without a PCa diagnosis (bottom panel), for the participants in the screening arm of the Spanish ERSPC study.

We defined the time-to-event,  $T$ , as the time elapsed from the protocol screening start (age 45) to diagnosis of prostate cancer. We observe that more than 90% of patients were prostate cancer free at age 80 years (see Figure 2a ). In particular 116 men (4.8%) were diagnosed with PCa (median follow-up time equal to 2.85 years) and the remaining 2299 men (95.2%) contributed to the study as right-censored data with a median follow-up time equal to 9.76 years.

As previously mentioned, a  $PSA \geq 3$  ng/ml is a key threshold in the screening program. Figure 2b shows the Kaplan-Meier estimate of the survival function for individuals whose first PSA measurement is below or above this limit. A rapid decrease of the survival function for individuals with the first  $PSA \geq 3$  ng/ml is observed and a clinically relevant difference exists between the two categories.

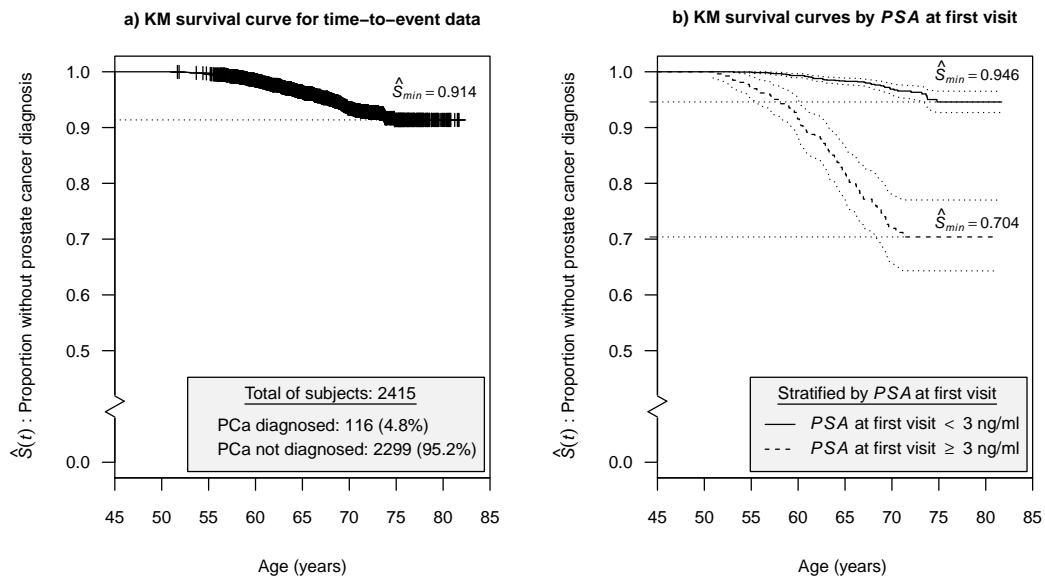


Figure 2. Plot of the Kaplan-Meier (KM) estimate of the survival function of time to PCa diagnosis for the participants in the screening arm of the Spanish ERSPC study a) overall, and b) stratified (and 95% confidence intervals) by the value (below or above 3 ng/ml) of the first PSA measurement. To facilitate the interpretation of the results, time zero has been labeled as 45 years in the  $x$ -axis.

### 3. A joint model with shared random effects for the Spanish ERSPC study

Our main goal was to assess the effect of PSA profiles over time on the hazard of time to PCa diagnosis. Our approach consisted of including random effects in the longitudinal covariates and jointly modeling the longitudinal and survival components. This joint procedure allows a detailed specification of every submodel [7, 9, 26, 34, 38] as well as achieving a reduction in the standard error of the estimates [21].

In what follows, we propose a shared random effects joint model for the screening arm of the ERSPC study. We start with some common notation and the basic longitudinal and survival submodels. The rest of Section 3 provides the estimation procedure, computational details and interpretation from frequentist and Bayesian perspectives.

#### 3.1 Model proposal and estimation procedure

For the  $i$ -th subject,  $i = 1, \dots, n$ , denoted by  $m_i(t)$  the true *LLPSA* value at time  $t$  and by  $M_i(t) = \{m_i(s), 0 \leq s \leq t\}$  the whole longitudinal history of the true marker levels up to

time  $t$ . The observable data for the  $i$ -th subject consists of the value of the observed longitudinal *LLPSA* profile at time point  $t_{ij}$ , that is  $y_{ij} = LLPSA_{ij} = \log(1 + \log(1 + PSA_{ij}))$ . Let  $y_i(t)$  be the hypothetical observed value of *LLPSA* at time  $t$  which is assumed to deviate from the true value  $m_i(t)$  by a certain error measurement  $\varepsilon_i(t)$  in such a way that  $y_i(t) = m_i(t) + \varepsilon_i(t)$  holds.

The joint model that we propose binds a linear mixed-effects model for the longitudinal part and a relative risk survival model for the time to PCa diagnosis, which incorporates the true historical subject profile of the longitudinal process [26]. The survival submodel is based on the Weibull distribution due to its flexibility for representing different types of risks. The joint model is specified as

$$\begin{cases} (y_{ij} | \theta_i) = m_i(t_{ij}) + \varepsilon_{ij} = \beta_0 + b_{i0} + (\beta_1 + b_{i1})t_{ij} + \varepsilon_{ij} \\ h_i(t | \mathcal{M}_i(t), w_{1i}, \theta_i) = h_0(t | \sigma_t) \exp\{\gamma_0 + \gamma_1 w_{1i} + \alpha m_i(t)\}, \end{cases} \quad (1)$$

where  $\theta_i = (\beta_0, \beta_1, b_{i0}, b_{i1}, \sigma_{b_0}, \sigma_{b_1}, \rho_{b_0 b_1}, \sigma, \sigma_t, \gamma_0, \gamma_1, \alpha)^T$  is the full vector of uncertainties associated with individual  $i$ . For the longitudinal submodel,  $(\beta_0, \beta_1)^T$  and  $(b_{i0}, b_{i1})^T$  are the fixed and the random effects for the intercept and the slope term, respectively, and  $\varepsilon_{ij}$  the error term for the  $i$ -th subject at the  $j$ -th measurement with  $(\varepsilon_{ij} | \sigma) \sim \mathcal{N}(0, \sigma^2)$ ; furthermore we assume that both random effects given  $\sigma_{b_0}, \sigma_{b_1}$  and their correlation coefficient,  $\rho_{b_0 b_1}$ , follow a bivariate normal distribution with a mean of zero and an unstructured covariance matrix. For the survival submodel, the baseline hazard is given by  $h_0(t | \sigma_t) = \sigma_t t^{\sigma_t - 1}$ , where  $\sigma_t$  is the shape parameter;  $w_1 = a_1 \times y_1$  is a covariate which describes the interaction between the age,  $a_1$ , and the *LLPSA* measurement,  $y_1$ , at study entry (first visit), respectively. Its inclusion in the model is justified by the association between dynamic changes in PSA over time and PCa diagnosis ([12]), presumably as the result of benign enlargement of the prostate gland;  $(\gamma_0, \gamma_1)^T$  are the regression coefficients for the baseline hazard and the  $w_1$  interaction. Finally,  $\alpha$  stands for the association parameter assessing the relationship between the longitudinal and the survival submodels.

Since preliminary analysis for joint model (1) showed a large correlation between  $b_{i0}$  and  $b_{i1}$  we discarded the slope random effect and restrict posterior analyses, both frequentist and Bayesian, interpretations and conclusions to the following model

$$h_i(t | \mathcal{M}_i(t), w_{1i}, \theta_i) = h_0(t | \sigma_t) \exp\{\gamma_0 + \gamma_1 w_{1i} + \alpha(\beta_0 + b_{i0} + \beta_1 t)\}, \quad (2)$$

where  $(b_{i0} | \sigma_{b_0}) \sim \mathcal{N}(0, \sigma_{b_0}^2)$  and the error term given  $\sigma$  in the longitudinal submodel follows, as stated above, a  $\mathcal{N}(0, \sigma^2)$  law.

### 3.2 Frequentist approach. Results and interpretation

The joint model expression given in (2) was fitted through the `jointModel` function, implemented in the R package JM [23]. In particular, the fitting of the initial linear mixed submodel was performed using the `nlme` package [20], while the `survival` package [32] was used for estimation of the preliminary survival submodel. Models were validated using the `residuals` function. JM addresses the key feature of dealing with the shared random effects by numerical integration and by using a pseudo-adaptive Gauss-Hermite quadrature algorithm [25]. In order to select the appropriate model within the family of joint models given in expression (1), we used the Wald test, the AIC criterium and  $p$ -values  $< 0.05$ .

The longitudinal submodel in (2) consisted of a linear fixed-effects part with age as a time covariate, and a random effects part which included only a random intercept. The frequentist estimation resulted in a marginal regression line with positive estimated intercept,  $\hat{\beta}_0 = 0.369$ , and

slope,  $\hat{\beta}_1 = 0.014$ . The variability associated with the random intercept was high,  $\hat{\sigma}_{b_0}^2 = 0.050$ , indicating the great heterogeneity of the population with regard to the *LLPSA* measurements at time zero. Estimates for the parameters in model (2) are given in Table 2.

In the survival submodel, the estimated coefficient associated with the  $w_1$  interaction was  $\hat{\gamma}_1 = -0.068$  indicating a protective effect. Finally, the joint model found a strong association between the *LLPSA* profiles and risk of PCa, with an estimated value equal to 7.207 ((6.041, 8.372) 95% confidence interval).

Table 2. Joint model estimates for frequentist analysis of PCa diagnosis and longitudinal PSA values. The survival submodel follows a Weibull distribution and includes the  $a_1 \times y_1$  interaction as a covariate, whereas the longitudinal submodel is a linear mixed effects model with random intercept.

Parameters	Value	95% Conf. Int.
Longitudinal Submodel		
$\beta_0$	0.369	(0.351, 0.388)
$\beta_1$	0.014	(0.013, 0.015)
$\sigma^2$	0.011	(0.011, 0.012)
$\sigma_{b_0}^2$	0.050	(0.044, 0.057)
Survival Submodel		
$\gamma_0$	-13.706	(-15.299, -12.113)
$\gamma_1$	-0.068	(-0.107, -0.028)
$\sigma_t$	1.887	(1.459, 2.315)
Association		
$\alpha$	7.207	(6.041, 8.372)

The estimated relative risk ( $\widehat{HR}$ ) associated with a one year increase in the age at entry in the interaction term  $w_1$  was  $\exp(\hat{\gamma}_1) = 0.93$ , which indicates a 7% decrease in the risk of PCa. Given that high PSA values are related to PCa incidence this might seem an unexpected result. But, an elevated level of PSA can also result from benign enlargement of the prostate or from inflammation, conditions that are relatively rare in the young but common in older men [36]. Therefore, the protective effect of the  $w_1$  interaction may indicate that an increase in PSA is associated with a larger increase in risk for a younger man than for an older man.

The joint model reflects a strong positive association between PSA values and risk of PCa, and allows estimation of the impact of different increments of PSA at specific PSA values. Table 3 presents the estimated HRs that correspond to different increases in PSA for different PSA values. For instance, considering the estimated parameter  $\alpha = 7.207$  in model (2), a unit increase in PSA represents a 4.7, 1.9 or 1.2-fold increase in the hazard rate of PCa, for reference “true PSA” values 1, 3 or 10, respectively. Therefore, the same changes in PSA have more impact on PCa risk for low PSA values. However, for particular percentual increases of PSA, the HRs did not vary much across different PSA values. Increases of 10% or 20% in PSA levels result in HRs around 1.2 or 1.4, respectively.

Table 3. Estimated hazard ratio ( $\widehat{HR}$ ) of PCa for different increases of PSA at specific PSA reference levels, under the frequentist model.

		PSA reference level (ng/ml)						
		1	2	3	5	10	15	20
$\widehat{HR}$ for an	1 unit	4.70	2.52	1.90	1.47	1.20	1.12	1.09
increase of	10%	1.23	1.24	1.24	1.23	1.20	1.18	1.17
PSA of	20%	1.48	1.52	1.51	1.47	1.41	1.38	1.36

Figure 3 illustrates a contour plot of the estimated HR for a  $\Delta$ PSA variation as a function of PSA. For instance, we can see that an increase of 2.0 ng/ml in PSA would represent a higher impact for a subject with 1.0 ng/ml of PSA ( $\widehat{HR} = 11.9$ ) than for a subject whose PSA value was 8.0 ng/ml ( $\widehat{HR} = 1.6$ ).

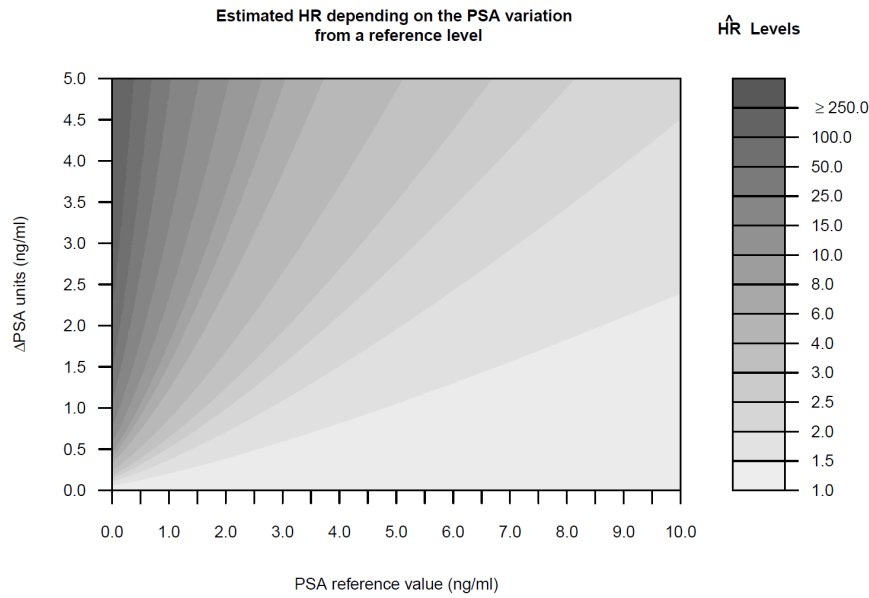


Figure 3. Contour plot of the estimated HR for a  $\Delta$ PSA variation as a function of PSA.

### 3.3 The proposed joint model revisited. A Bayesian approach

We assume the sampling model in (2). From the Bayesian approach, the joint model is not entirely specified until the elicitation of a joint prior distribution for all the uncertainties in the model,  $\theta = \cup_{i=1}^n \theta_i$ . We assume that all the random components in  $\mathbf{b}_0 = (\mathbf{b}_{10}, \dots, \mathbf{b}_{n0})^T$  are i.i.d.  $\mathcal{N}(0, \sigma_{b_0}^2)$  given  $\sigma_{b_0}^2$ , and consider prior independency for the rest of parameters and hyperparameters in the joint model. Accordingly, the common prior distribution will be the product of the corresponding marginal prior distributions

$$\pi(\theta) = \pi(\sigma_t) \pi(\gamma_0) \pi(\gamma_1) \pi(\alpha) \pi(\beta_0) \pi(\beta_1) \pi(\mathbf{b}_0 | \sigma_{b_0}) \pi(\sigma_{b_0}) \pi(\sigma),$$

with noninformative Normal priors for parameters  $\beta_i$ 's,  $\gamma_i$ 's and  $\alpha$ , and noninformative Gamma distributions for the standard deviations associated with the measurement error, the random intercept, and the shape parameter of the baseline hazard in the survival submodel.

Note that in the Bayesian approach, the shared joint model induces conditional independence between the longitudinal and the survival processes given not only the random effects and covariates but also the *fixed* effects and the rest of parameters as a result of its stochastic role. Bayesian inference expresses all the uncertainties in the problem in probabilistic terms, with the posterior distribution being the most relevant element and, usually, the starting point of all relevant inferences. Because of the complexity of our model, the posterior distribution,  $\pi(\theta | \mathcal{D}_n)$ , where  $\mathcal{D}_n$  stands for the data, is not analytical. We have approximated it through Markov Chain Monte Carlo methods (MCMC). In particular, we performed three MCMC chains with  $K = 150\,000$  iterations each, of which 50 000 were used for the burn-in period. In addition, they were thinned by storing every 50th iteration in order to decrease autocorrelation in the sample. Computations were carried out using the statistical software WinBUGS [18] and the R-package JMBayes [27].

Table 4. Posterior summaries for the parameters and hyperparameters of the PSA joint model.

Parameters	Mean	95% Cred. Int.
Longitudinal submodel		
$\beta_0$	0.362	(0.342, 0.382)
$\beta_1$	0.014	(0.013, 0.016)
$\sigma^2$	0.011	(0.010, 0.012)
$\sigma_{b_0}^2$	0.070	(0.066, 0.074)
Survival submodel		
$\gamma_0$	-12.972	(-14.485, -11.650)
$\gamma_1$	-0.064	(-0.101, -0.027)
$\sigma_t$	1.708	(1.362, 2.125)
Association		
$\alpha$	6.919	(5.858, 8.026)

Table 4 shows the approximate marginal posterior mean and 95% credible interval of each one of the parameters and hyperparameters of the model. Posterior summaries provided similar numerical results to those obtained from the frequentist methodology, mainly as a result of the use of noninformative priors. The strength of the association between the longitudinal and survival outcomes is estimated through the marginal distribution of the association parameter  $\alpha$ . The posterior expectation  $E(\alpha | \mathcal{D}_n) = 6.919$ , and posterior probability  $P(5.858 < \alpha < 8.026 | \mathcal{D}_n) = 0.95$  provide strong support in favor of joint modeling and against a potential marginal survival analysis. Random effects are also assessed through their marginal posterior distribution. Figure 4 on the left shows the posterior distribution of the random intercept for three different subjects in the study, in particular subjects 170, 707, and 1006, selected with regard to their low, medium and high level of PSA measurements, respectively. These posteriors clearly indicate the heterogeneity of the subject-specific behaviour of the PSA at the time of the subject’s entry into the study.

We now examine the linear mixed-effect longitudinal submodel and focus on  $(m_i(t) | \beta_0, b_{i0}, \beta_1) = \beta_0 + b_{i0} + \beta_1 t$ , the true *LLPSA* value of subject  $i$  at age  $t$  years over 45. Note that the particular expression for  $m_i(t)$  indicates that the progression of the true *LLPSA* is equal for all the patients, with the exception of their specific *LLPSA* level at entry into the screening program. The estimation of the joint model can be easily spread to compute the posterior distribution  $\pi(m_i(t) | \mathcal{D}_n)$ , with posterior expectation

$$E(m_i(t) | \mathcal{D}_n) = \int m_i(t | \beta_0, b_{i0}, \beta_1) \pi(\beta_0, b_{i0}, \beta_1 | \mathcal{D}_n) d(\beta_0, b_{i0}, \beta_1), \quad (3)$$

where  $\pi(\beta_0, b_{i0}, \beta_1 | \mathcal{D}_n)$  is the marginal posterior distribution of  $(\beta_0, b_{i0}, \beta_1)$  for which we have an approximated MCMC sample,  $\{\beta_0^{(k)}, b_{i0}^{(k)}, \beta_1^{(k)}, k = 1, \dots, K\}$ , from the common posterior distribution  $\pi(\theta | \mathcal{D}_n)$ . A Monte Carlo approximation of  $E(m_i(t) | \mathcal{D}_n)$  is easily derived as

$$\frac{1}{K} \sum_{k=1}^K (m_i(t) | \beta_0^{(k)}, b_{i0}^{(k)}, \beta_1^{(k)}) = \frac{1}{K} \sum_{k=1}^K (\beta_0^{(k)} + b_{i0}^{(k)} + \beta_1^{(k)} t), \quad \forall t \geq 0. \quad (4)$$

Figure 4 on the right shows the posterior expected trajectories for the true *LLPSA* level with age for subjects 170, 707, and 1006 in the study. Note that all of them are parallel as a consequence of the assumed random intercept and the common slope in the model.

As we have already mentioned, the relative risk or hazard ratio (HR) corresponding to an increase of one year in age at the first visit,  $a_1$ , depends on the subject’s level of *LLPSA*,  $y_1$ , in the form  $\exp(\gamma_1 y_1)$ . The posterior distribution of this hazard ratio can be computed from the MCMC approximate random sample from the marginal posterior distribution,  $\pi(\gamma_1 | \mathcal{D}_n)$ . Figure 5 on the left displays the posterior mean and a credible 95% interval of this HR with



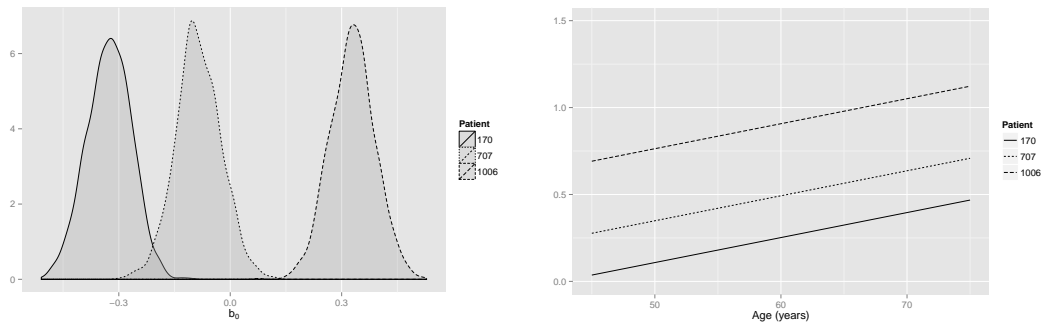


Figure 4. (left) Posterior distribution of the random intercept for subjects 170, 707, and 1006 in the study, and (right) posterior expectation of the true *LLPSA* level versus age for the same subjects as in (left).

regard to different values of *LLPSA* at the time of the first visit. For instance, comparing two subjects with identical *PSA* longitudinal profiles  $PSA = 11$  at entry (*i.e.*  $y_1 = 1.25$ ), but with one subject entering the study one year older, an expected reduction of about 7.5% in risk can be seen. In addition, the relative risk corresponding to an increase of  $k$  *LLPSA* units at entry is  $\exp(k \gamma_1 a_1)$ . Figure 5 on the right shows the posterior mean with age at first visit for values  $k = 0.1, 0.3$  and  $0.5$ . As expected, all the posterior mean values decrease with  $a_1$  and  $k$ , supporting the idea that an increment in the value of *LLPSA* at entry is less serious, in terms of risk, in older men than in younger people.

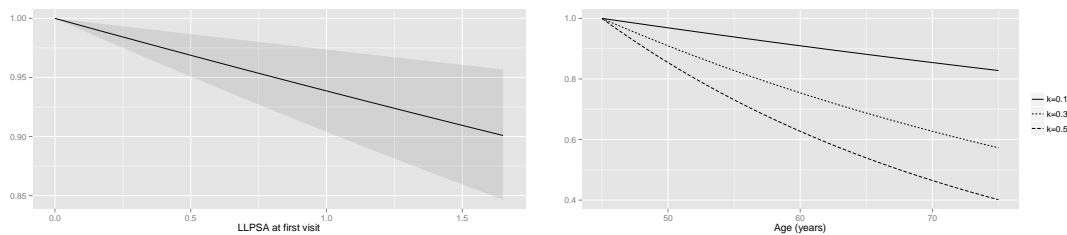


Figure 5. (left) Mean and 95% credible interval of the posterior distribution of the HR, corresponding to an increment of one unit in age at the first visit as a function of *LLPSA* at the first visit values, and (right) Mean of the posterior distribution of the HR corresponding to an increment of  $k = 0.1, 0.3, 0.5$  units in *LLPSA* at the first visit as a function of age at the first visit.

#### 4. Dynamic disease-free probabilities

The flexibility provided by the joint modeling approach is in line with a growing trend towards personalized Medicine [8, 22, 24]. From this perspective, it is important that screening programs incorporate statistical tools that can take into account the observed evolution of the health status of patients and facilitate estimation of disease-free probabilities at each visit. These estimates are of interest not only for individuals that follow the screening program but also for subjects from the population of which the sample is considered representative. In particular, the real challenge is being able to estimate these probabilities not only at each measurement time points, but also at a generic time  $u > t$  given *PSA* history tracking and disease-free time up to  $t$ .

In that scenario, the goal consists of estimating the probability of being diagnosis-free for a subject  $i$  with covariate  $w_{1i}$  after a certain set of longitudinal measurements  $\mathcal{Y}_i$  which assures disease-free status up to the time  $t$  of the last visit. Bayesian reasoning approaches this issue through the posterior distribution  $\pi(\Pr(T_i \geq u \mid T_i > t, \mathcal{Y}_i, w_{1i}, \theta) \mid \mathcal{D}_n, \mathcal{Y}_i)$  with posterior expectation

$$\Pr(T_i \geq u \mid T_i > t, \mathcal{Y}_i, w_{1i}, \mathcal{D}_n) = \int \Pr(T_i \geq u \mid T_i > t, \mathcal{Y}_i, w_{1i}, \boldsymbol{\theta}) \pi(\boldsymbol{\theta} \mid \mathcal{D}_n, \mathcal{Y}_i) d\boldsymbol{\theta}, \quad u > t, \tag{5}$$

where  $\pi(\boldsymbol{\theta} \mid \mathcal{D}_n, \mathcal{Y}_i)$  is the current posterior distribution of  $\boldsymbol{\theta}$ . Some discussion about the meaning of this posterior distribution with regard to individuals participating or not in the study becomes necessary. If there is interest in any of the specific individuals who have participated in the study, the posterior distribution  $\pi(\boldsymbol{\theta} \mid \mathcal{D}_n, \mathcal{Y}_i)$  does not need to be updated with  $\mathcal{Y}_i$  because it already contains all the relevant information about that individual, and consequently,  $\pi(\boldsymbol{\theta} \mid \mathcal{D}_n, \mathcal{Y}_i) = \pi(\boldsymbol{\theta} \mid \mathcal{D}_n)$ . If there is interest in estimating the conditional free-disease probability of a new subject who has not participated in the study in a sequential manner, the relevant posterior distributions,  $\pi(\boldsymbol{\theta} \mid \mathcal{D}_n, \mathcal{Y}_i)$  need to be computed. The probability  $\Pr(T_i \geq u \mid T_i > t, \mathcal{Y}_i, w_{1i}, \boldsymbol{\theta})$  in the integrand of (5) depends on the current survival time,  $T_i > t$ , but not on the particular longitudinal trajectory,  $\mathcal{Y}_i$ , of the follow-up of the patient. This fact is a result of the shared joint effects model, which assumes independence between the survival and the longitudinal process, given the common random effects, covariates, and parameters.

Figure 6 shows the posterior probability of (5) together with a 95% credible interval for subjects 170, 707, and 1006 in the study at the end of their follow-up. We were focusing on people in the study, and consequently these posteriors were computed from the common posterior distribution  $\pi(\boldsymbol{\theta} \mid \mathcal{D}_n)$  in Section 3.3. The graphics also include their observed trajectory and the posterior mean of their true *LLPSA* level as it was computed in (4) and displayed in Figure 4. The optimistic information expected for subjects 170 and 707, despite the differences in their subsequent estimated random intercept, should be noted, and the more worrying situation for subject 1006, with an estimation of the true *LLPSA* markedly higher than for the two previous subjects.

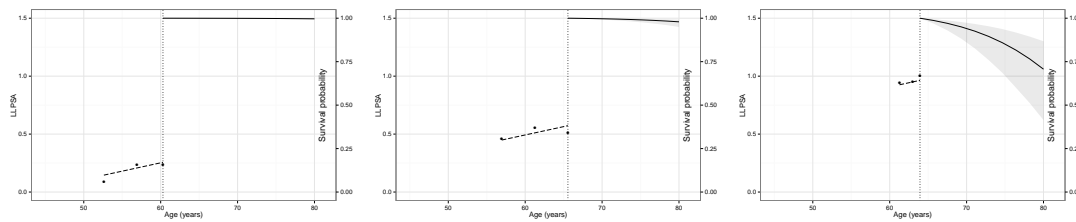


Figure 6. Dynamic posterior survival probabilities for subjects 170 (on the left), 707 (in the middle), and 1006 (on the right) in the study. Vertical lines correspond to the age of the last PSA measurement. The solid curve on the right of the vertical line is the posterior median of the target posterior distribution, and the shaded area is a 95% credible interval for this posterior median. Points on the left are the available longitudinal measurements and the line represents the posterior mean of their true *LLPSA* level.

As previously mentioned, inference for individuals not in the study requires the sequential updating of the posterior probability of  $\boldsymbol{\theta}$  on the basis of the information provided by the subsequent longitudinal measurements. We would expect a minimal change in the estimation of the common parameters but not of the subject-specific random effects. The update of posterior distributions for which we only have an approximate sample is a complex process which is beyond the scope of this paper. [24] develops a very clever proposal for approximating  $\pi(\boldsymbol{\theta} \mid \mathcal{D}_n, \mathcal{Y}_i)$  and updating the specific random effects associated with the individual but not the common parameters. In particular, he uses empirical Bayesian estimation for the random effects and an asymptotic normal distribution, based on the maximum likelihood estimation, for the common population parameters.

Figure 7 shows the estimated disease-free probabilities for a generic individual not in the study, which have been sequentially updated as the longitudinal follow-up progresses. They were

obtained using the method proposed by Rizopoulos, using the R package JM. A more pronounced increasing trend in the longitudinal trajectory of the PSA is associated with a lower PCa-free survival probability, which highlights the relevance of an accurate follow-up in personalized decision making. Specifically, according to Figure 7, Table 5 shows the dynamic estimates for disease-free PCa probabilities at 65, 70, 75 and 80 years old and at 10 years after a subject’s last visit.

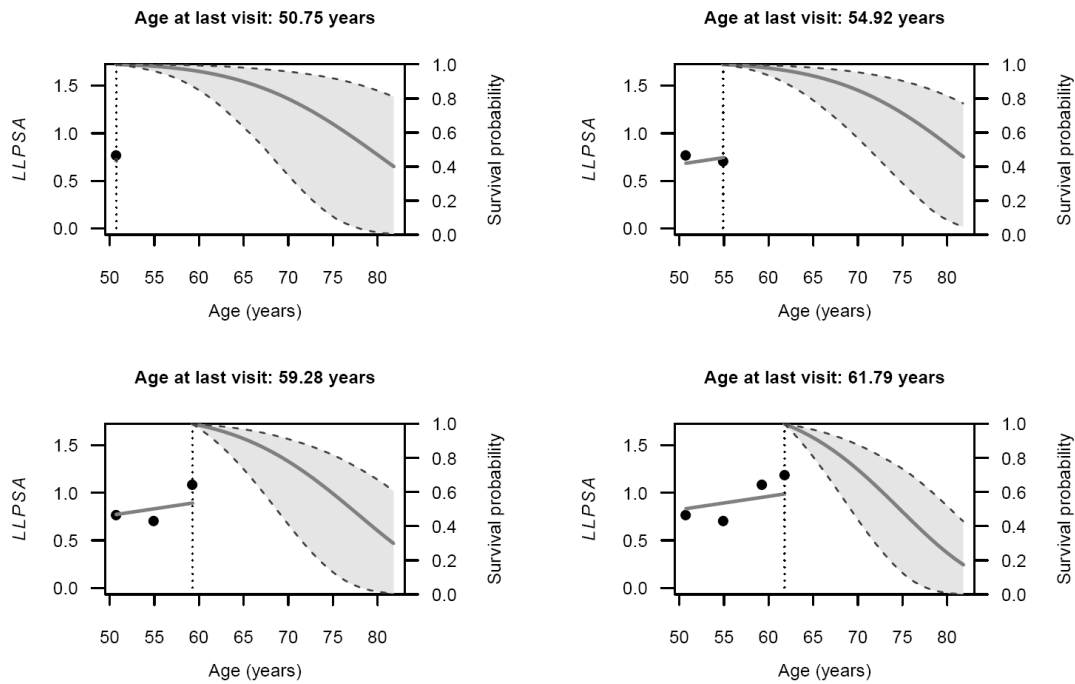


Figure 7. Successive *LLPSA* longitudinal trajectories and dynamic PCa free survival probabilities (with 95% pointwise credible intervals at each time point) for a patient after their first to fourth successive PSA measurements, based on the Spanish branch of the ERSPC study.

Table 5. Dynamic estimates for the PCa-free probabilities at 65, 70, 75 and 80 years old and at 10 years after their last visit for a patient after their first to fourth successive PSA measurements, based on the Spanish branch of the ERSPC study.

Age	PSA	<i>LLPSA</i>	PCa-free probability at				
			age 65	age 70	age 75	age 80	10 years after last visit
50.75	2.20	0.772	0.928	0.848	0.704	0.502	0.967
54.92	1.82	0.711	0.945	0.874	0.755	0.567	0.945
59.28	6.24	1.092	0.932	0.816	0.635	0.395	0.836
61.79	8.85	1.190	0.927	0.752	0.501	0.240	0.668

### 5. Discussion

In this study, using data from the Spanish arm of the ERSPC study, we have estimated a joint model of longitudinal measurements of PSA and time to diagnosis of PCa. Our model overcomes the limitations of simpler statistical tools because it accounts for a) the effect of PSA as an endogenous time-dependent covariate measured with error and b) the non-random dropout that results when an individual is diagnosed with PCa [26, 34]. In particular, we used both the frequentist and the Bayesian approaches, and built a shared-parameter joint model that combined some baseline covariates at study entry and longitudinal PSA values to predict PCa incidence.

The selected joint model included a linear trajectory with random intercept for the  $\log(1 + \log(1 + PSA))$  values in the longitudinal submodel. The survival submodel included the interaction of *age* and *LLPSA* at study entry together with the *true* longitudinal marker. Our model is consistent with the literature and clinical knowledge. On the one hand, the observed PSA is highly associated with the risk of being diagnosed with PCa. On the other hand, there is a protective effect of the  $\text{age} \times \text{PSA}$  term, consistent with the age-varying effect of PSA on PCa risk [36]. Our contribution consisted of using joint modeling techniques to 1) obtain an efficient estimate of the the impact of PSA trajectories on time to PCa diagnosis; and 2) refine the PCa free survival estimates by using predictions based on the whole true history of PSA evolution for each subject. Our work can be considered a first step that should be followed by a more comprehensive modeling process accounting for additional predictive factors measured over time.

Our results show that both the frequentist and Bayesian paradigms produced very close parameter estimates and subsequent 95% confidence and credible intervals. This agreement is not an unexpected fact because a) the same sampling model was used for both approaches; b) a very slightly informative prior distribution was used for the Bayesian analysis, which provided virtually no additional information in the estimation process; and c) the sample size was large enough to be the dominant and determining source of information in the study. Bayesian methods often involve slower and computationally intensive tools for approximating the posterior distribution. However, once it is obtained they provide a natural framework for dealing with random effects and predictions. In particular, in our study we have applied Bayesian inference to assess individual true *LLPSA* levels and obtain dynamic predictions of the survival probabilities. Given that screening for prostate cancer may benefit many but also may result in adverse events in some, it is essential to develop tools that allow interventions to be more effectively targeted to individuals at high risk and reduced in those at low risk, thus minimizing harms.

Research on clinical decision-making, based on dynamic PSA measures, has been more frequent in the clinical setting than in early detection. Prediction tools of recurrence based on longitudinal PSA values have been proposed to guide decisions for patients treated for PCa. [22] and [38], with different joint modeling approaches, showed that accounting for longitudinal PSA measures was a powerful tool for individual prediction. [31], following the joint modeling work of [22], implemented real-time predictions of PCa recurrence on a web-based calculator. In the early detection area, several studies have assessed the predictive value of dynamic PSA markers like PSA velocity, PSA doubling time or PSA percentual change. In general, these PSA change markers added little to the predictive accuracy of PSA levels alone or in combination with other variables [12, 35]. However, and consistent with our results, the studies that used joint modeling approaches showed that the whole PSA history had a relevant role for PCa risk assessment [3, 13, 15].

Screening for prostate cancer has originated substantial debate in the clinical and public health communities. The true benefit of screening for PCa remains uncertain and may result in harms, such as false positive results, overdiagnosis and overtreatment, for many men that never would become symptomatic [11]. The PSA test alone, regardless of the cut-off value assigned, is not suitable for a screening test due to high false-negative and false-positive results [10]. Risk calculators that include PSA and other variables have shown better predictive ability with respect to PSA levels alone [29]. An improvement of predictive models is necessary to better inform individual patient decision making and health policy decisions on PCa screening.

Although joint modeling seems a very suitable approach for our data, this study has several limitations. First, the number of observations per subject was small, only 14% of them had three or more PSA measurements. Having more measurements per individual would have prevented some potential bias in the random effect estimate and would have produced more precise estimates of the association parameter. In addition, more PSA measurements per subject probably would have resulted in a richer longitudinal model, better capturing the heterogeneity of the sample and better feeding the survival submodel. Second, the fact that subjects entered the study

at different ages may have reduced the accuracy of the estimated association parameter. And third, in contrast with the multivariate risk calculators, our model included only the repeated PSA measurements over time, accounting for the age and PSA interaction at study entry. Our dataset, which was not a clinical database, did not include variables like prostate volume or family history that could have improved the predictive ability of the model. Other variables like the result of digital rectal exam and transurethral ultrasound were not measured systematically but only when a biopsy was performed. The recent approach of [28] on multivariate longitudinal joint models should be considered in future works on individual risk assessment.

In conclusion, the combination of frequentist and Bayesian approaches, for shared random effects joint models, enriched the analysis and interpretation of our data. Moreover, the complementarity of both approaches helped to establish a bridge between both ways of thinking. This study gives support to the age-decreasing effect of PSA on PCa risk. In addition, it corroborates a strong association between longitudinal PSA values and diagnosis of PCa. Future works should explore whether joint models with additional longitudinal biomarkers and richer data can improve the performance of the existing risk calculators.

### Acknowledgements

This paper has been partially supported by research grants MTM2012-38067-C02-01 and MTM2010-19528 from the Spanish Ministry of Economy and Competitiveness and the Spanish Ministry of Education and Science, respectively. Authors are grateful to the TRUEJM group, specifically to professor Dimitris Rizopoulos from the Erasmus Medical Center, for the fruitful discussions on joint modeling issues. We are also particularly indebted to Dr. Marcos Luján from the Hospital Universitario Infanta Cristina for providing the Spanish ERSPC database and for his generous collaboration on data interpretation and review of the manuscript. We thank JP Glutting for reviewing the English and editing the manuscript.

### Conflict of Interest

*The authors have declared no conflict of interest.*

### References

- [1] G.L. Andriole, E.D. Crawford, R.L. Grubb, S.S. Buys, D. Chia, T.R. Church, M.N. Fouad, E.P. Gelmann, P.A. Kvale, D.J. Reding, J.L. Weissfeld, L.A. Yokochi, B. O'Brien, J.D. Clapp, J.M. Rathmell, T.L. Riley, R.B. Hayes, B.S. Kramer, G. Izmirlian, A.B. Miller, P.F. Pinsky, P.C. Prorok, J.K. Gohagan, C.D. Berg, and P.P. Team, *Mortality results from a randomized prostate-cancer screening trial*, *The New England Journal of Medicine* 360 (2009), pp. 1310–1319.
- [2] M. Bayarri and J. Berger, *The Interplay of Bayesian and Frequentist Analysis*, *Statistical Science* 19 (2004), pp. 58–80.
- [3] L.J. Brandt, S.L. Sheng, C.H. Morrell, G.N. Verbeke, E. Lesaffre, and H.B. Carter, *Screening for prostate cancer by using random-effects models*, *Journal of the Royal Statistical Society: Series A* 166 (2003), pp. 51–62.
- [4] H.B. Carter, P.C. Albertsen, M.J. Barry, R. Etzioni, S.J. Freedland, K.L. Greene, L. Holmberg, P. Kantoff, B.R. Konety, M.H. Murad, D.F. Penson, and A.L. Zietman, *Early detection of prostate cancer: Aua guideline*, *American Urological Association* 190 (2013), pp. 419–426.
- [5] G. Casella and R. Berger, *Reconciling Bayesian and Frequentist Evidence in the One-Sided Testing Problem*, *Journal of the American Statistical Association* 82 (1987), pp. 106–111.
- [6] R. Etzioni, A. Tsodikov, A. Mariotto, A. Szabo, S. Falcon, J. Wegelin, D. diTommaso, K. Karnofski, R. Gulati, D. Penson, and E. Feuer, *Quantifying the role of PSA screening in the US prostate cancer mortality decline*, *Cancer Causes and Control* 19 (2008), pp. 175–181.
- [7] C.L. Faucett and D.C. Thomas, *Simultaneously modelling censored survival data and repeatedly measured covariates: a gibbs sampling approach*, *Statistics in Medicine* 15 (1996), pp. 1663–1685.

- [8] F.G. Garre, A.H. Zwinderman, R.B. Geskus, and Y.W. Sijpkens, *A joint latent class changepoint model to improve the prediction of time to graft failure*, *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 171 (2008), pp. 299–308.
- [9] R. Henderson, P. Diggle, and A. Dobson, *Joint modelling of longitudinal measurements and event time data*, *Biostatistics* 1 (2000), pp. 465–480.
- [10] B. Holmstrom, M. Johansson, A. Bergh, U.H. Stenman, G. Hallmans, and P. Stattin, *Prostate specific antigen for early detection of prostate cancer: longitudinal study*, *BMJ* 339 (2009), p. b3537.
- [11] D. Ilic, M.M. Neuberger, M. Djulbegovic, and P. Dahm, *Screening for prostate cancer*, *Cochrane Database of Systematic Reviews* 1 (2013), p. CD004720.
- [12] B.A. Inman, J. Zhang, N.D. Shah, and B.T. Denton, *An examination of the dynamic changes in prostate-specific antigen occurring in a population-based cohort of men over time*, *BJU International* 110 (2012), pp. 375–381.
- [13] L.Y. Inoue, R. Etzioni, E.H. Slate, C. Morrell, and D.F. Penson, *Combining longitudinal studies of psa*, *Biostatistics* 5 (2004), pp. 483–500.
- [14] S. Lele, B. Dennis, and F. Lutscher, *Data cloning: easy maximum likelihood estimation for complex ecological models using Bayesian Markov chain Monte Carlo methods*, *Ecology Letters* 10 (2007), pp. 551–563.
- [15] H. Lin, B.W. Turnbull, C.E. McCulloch, and E.H. Slate, *Latent class models for joint analysis of longitudinal biomarker and event process data: application to longitudinal prostate-specific antigen readings and prostate cancer*, *Journal of the American Statistical Association* 97 (2002), pp. 53–65.
- [16] R. Little, *Calibrated Bayes: A Bayes/frequentist roadmap*, *The American Statistician* 60 (2006), pp. 213–223.
- [17] M. Lujan, A. Paez, A. Berenguer, and J.A. Rodriguez, *Mortality due to prostate cancer in the spanish arm of the european randomized study of screening for prostate cancer (erspc). results after a 15-year follow-up*, *Actas Urológicas Españolas* 36 (2012), pp. 403–409.
- [18] D.J. Lunn, A. Thomas, N. Best, and D. Spiegelhalter, *WinBUGS – A Bayesian modelling framework: Concepts, structure, and extensibility*, *Statistics and Computing* 10 (2000), pp. 325–337.
- [19] A. Neuhaus, T. Augustin, C. Heumann, and M. Daume, *A Review on Joint Models in Biometrical Research*, *Journal of Statistical Theory and Practice* 3 (2009), pp. 855–868.
- [20] J.C. Pinheiro and D.M. Bates, *Approximations to the log-likelihood function in the nonlinear mixed-effects model*, *Journal of Computational and Graphical Statistics* 4 (1995), pp. 12–35.
- [21] R. Prentice, *Covariate measurement errors and parameter estimation in a failure time regression model*, *Biometrika* 69 (1982), pp. 331–342.
- [22] C. Proust-Lima and J.M. Taylor, *Development and validation of a dynamic prognostic tool for prostate cancer recurrence using repeated measures of posttreatment psa: a joint modeling approach*, *Biostatistics* 10 (2009), pp. 535–549.
- [23] D. Rizopoulos, *JM: An R package for the joint modelling of longitudinal and time-to-event data*, *Journal of Statistical Software* 35 (2010), pp. 1–33, Available at <http://www.jstatsoft.org/v35/i09/>.
- [24] D. Rizopoulos, *Dynamic predictions and prospective accuracy in joint models for longitudinal and time-to-event data*, *Biometrics* 67 (2011), pp. 819–829.
- [25] D. Rizopoulos, *Fast fitting of joint models for longitudinal and event time data using a pseudo-adaptive gaussian quadrature rule*, *Computational Statistics and Data Analysis* 56 (2012), pp. 491–501.
- [26] D. Rizopoulos, *Joint models for longitudinal and time-to-event data with applications in R*, CRC Press, Boca Raton, FL, 2012.
- [27] D. Rizopoulos, *JMbayes: JMbayes: Shared parameter models for the joint modeling of longitudinal and time-to-event data using JAGS, WinBUGS, or OpenBUGS. R package 0.4-1* (2013), Available at <http://cran.r-project.org/package=JMbayes/>.
- [28] D. Rizopoulos and P. Ghosh, *A bayesian semiparametric multivariate joint model for multiple longitudinal outcomes and a time-to-event*, *Statistics in Medicine* 30 (2011), pp. 1366–1380.
- [29] F. Schroder and M.W. Kattan, *The comparability of models for predicting the risk of a positive prostate biopsy with prostate-specific antigen alone: a systematic review*, *European Urology* 54 (2008), pp. 274–290.
- [30] F.H. Schroder, J. Hugosson, M.J. Roobol, T.L. Tammela, S. Ciatto, V. Nelen, M. Kwiatkowski, M. Lujan, H. Lilja, M. Zappa, L.J. Denis, F. Recker, A. Paez, L. Maattanen, C.H. Bangma, G. Aus, S. Carlsson, A. Villers, X. Rebillard, T. van der Kwast, P.M. Kujala, B.G. Blijenberg, U.H. Stenman, A. Huber, K. Taari, M. Hakama, S.M. Moss, H.J. de Koning, A. Auvinen, and E. Investigators, *Prostate-cancer mortality at 11 years of follow-up*, *The New England Journal of Medicine* 366 (2012), pp. 981–990.
- [31] J.M. Taylor, Y. Park, D.P. Ankerst, C. Proust-Lima, S. Williams, L. Kestin, K. Bae, T. Pickles, and H. Sandler, *Real-time individual predictions of prostate cancer recurrence using joint models*, *Biometrics* 69 (2013), pp. 206–213.
- [32] T. Therneau and P. Grambsch, *Modeling Survival Data: Extending the Cox Model*, Springer, New York, 2000.
- [33] A. Tsiatis, V. DeGruttola, and M. Wulfsohn, *Modeling the relationship of survival to longitudinal data measured with error: applications to survival and CD4 counts in patients with AIDS*, *Journal of the American Statistical Association* 90 (1995), pp. 27–37.
- [34] A.A. Tsiatis and M. Davidian, *Joint modeling of longitudinal and time-to-event data: An overview*, *Statistica*

- Sinica 14 (2004), pp. 809–834.
- [35] A.J. Vickers, *Re: an empirical evaluation of guidelines on prostate-specific antigen velocity in prostate cancer detection*, Journal of the National Cancer Institute 103 (2011), pp. 1635–1636.
  - [36] A.J. Vickers, D. Ulmert, A.M. Serio, T. Bjork, P.T. Scardino, J.A. Eastham, G. Berglund, and H. Lilja, *The predictive value of prostate cancer biomarkers depends on age and time to diagnosis: towards a biologically-based screening strategy*, International Journal of Cancer 121 (2007), pp. 2212–2217.
  - [37] M.S. Wulfsohn and A.A. Tsiatis, *A joint model for survival and longitudinal data measured with error*, Biometrics 53 (1997), pp. 330–339.
  - [38] M. Yu, N. Law, J. Taylor, and H. Sandler, *Joint longitudinal-survival-cure models and their application to prostate cancer*, Statistica Sinica 14 (2004), pp. 835–862.