Supplemental Online Content

Finelli A, Beer TM, Chowdhury S, et al. Comparison of joint and landmark modeling for predicting cancer progression in men with castration-resistant prostate cancer: a secondary post hoc analysis of the PREVAIL randomized clinical trial. *JAMA Netw Open*. 2021;4(6):e2112426. doi:10.1001/jamanetworkopen.2021.12426

This supplemental material has been provided by the authors to give readers additional information about their work.

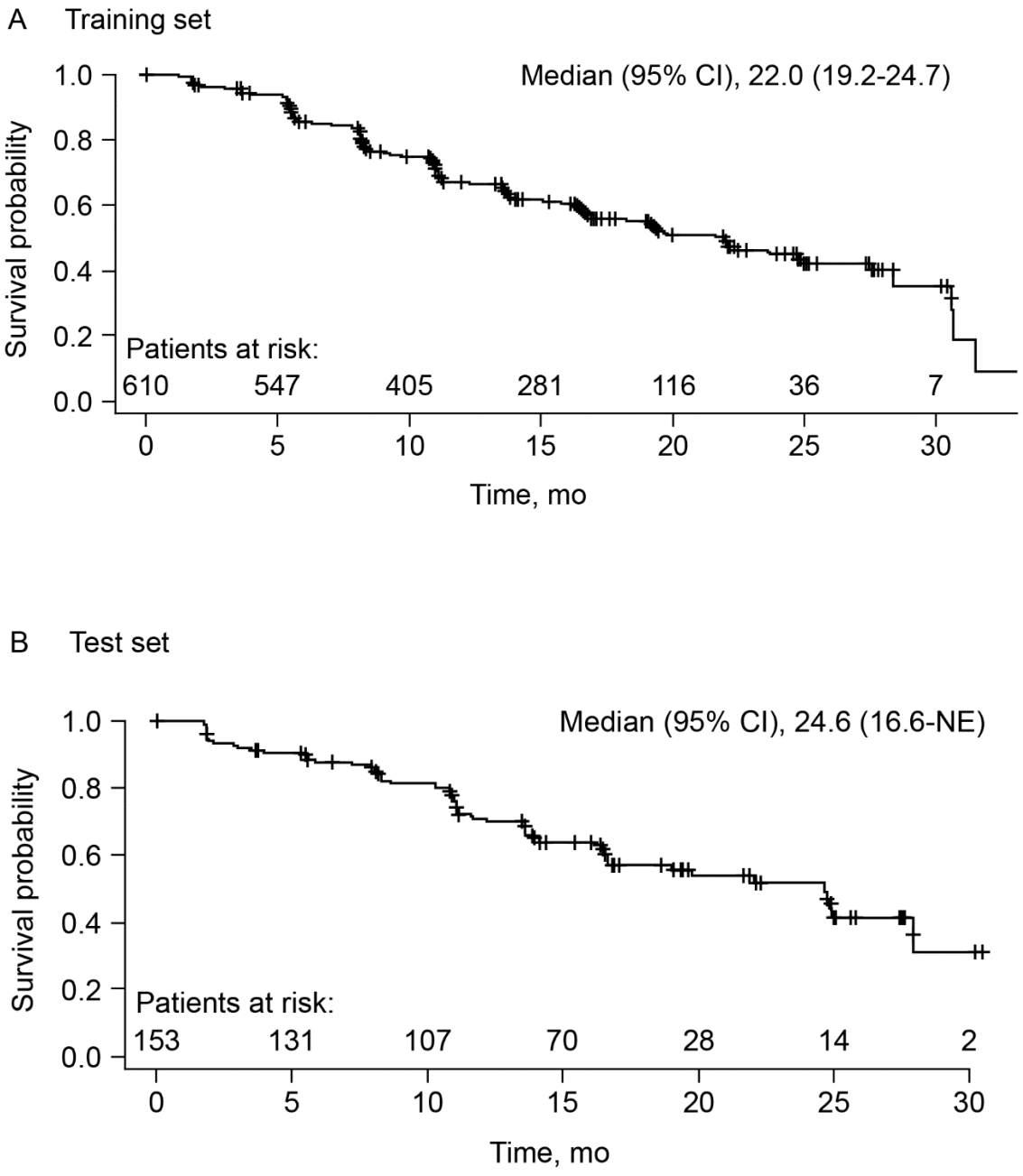
eAppendix 1. Definition of Radiographic Progression-Free Survival in the PREVAIL Study
rPFS was defined as time from randomization to first objective evidence of radiographic disease progression or death due to any cause within 168 days after treatment discontinuation, whichever occurred first. Radiographic disease progression included confirmed bone disease progression and soft tissue disease progression. Radiographic progression was confirmed by central independent radiology review until ≥ 410 rPFS events were confirmed as required for primary PFS analysis. After this, radiographic progression was confirmed by local radiology review.

eTable 1. Baseline Characteristics in the PREVAIL Phase 3 Trial

Category/statistic	Enzalutamide + no visceral disease (N = 773)
Age at study entry, years	
n	773
Mean (SD)	71.2 (8.54)
Median (min-max)	71.0 (44.0-93.0)
BMI at study entry (kg/m^2)	
n	771
Mean (SD)	28.41 (4.612)
Median (min-max)	27.60 (18.3-46.8)
ECOG score	
n	773
0	525 (67.9%)
1	248 (32.1%)
Total Gleason score	
n	743
≥ 8	381 (51.3%)
Pain Score assessed by BPF-SF item 3	
n	761
0–1	509 (66.9%)
2–3	239 (31.4%)
> 3	13 (1.7%)
Baseline serum PSA (ng/mL)	
n	773
Mean (SD)	137.92 (280.053)
Median (min-max)	51.20 (0.8-3182.0)
Baseline lactate dehydrogenase (U/L)	
n	772
Mean (SD)	204.65 (110.320)
Median (min-max)	184.00 (52.0-1861.0)
Baseline alkaline phosphatase (U/L)	
n	773
Mean (SD)	153.44 (282.570)
Median (min-max)	93.00 (34.0-4485.0)
Baseline hemoglobin (g/L)	
n	773
Mean (SD)	129.84 (12.541)
Median (min-max)	130.00 (84.0-168.0)

Abbreviations: BMI, body mass index; BPI-SF, Brief Pain Inventory-Short Form; ECOG, Eastern Cooperative Oncology Group; PSA, prostate-specific antigen; SD, standard deviation.

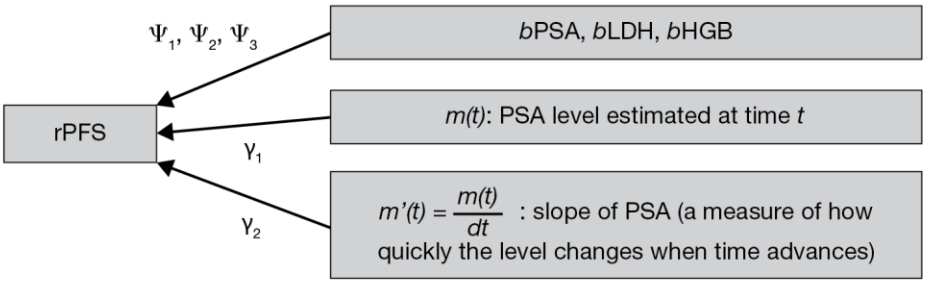
eFigure 1. Survival Probability Over Time of Patients in the PREVAIL Study



Abbreviations: CI, confidence interval; NE, not estimable.

eFigure 2. Directed Graph of Selected Joint and Landmark Models and Underlying Model Assumptions[a]

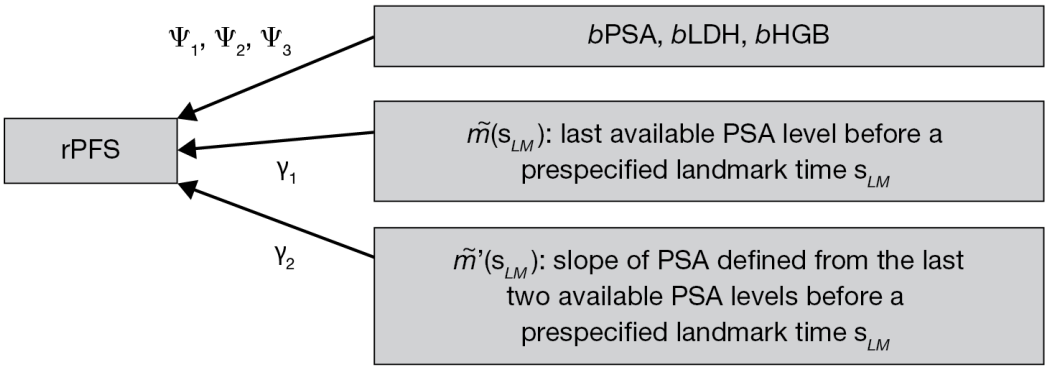
Joint Models



Hazard function h(t) of rPFS:
$h\{t|b\text{PSA}, b\text{LDH}, b\text{HGB}, \boldsymbol{M}(t)\} = h_0(t)\exp\{\Psi_1 b\text{PSA} + \Psi_2 b\text{LDH} + \Psi_3 b\text{HGB} + \gamma_1 m(t) + \gamma_2 m'(t)\}$,
where $\boldsymbol{M}(t)$ is history of PSA profile up to time t, and $h_0(t)$ represents a baseline hazard function.

Longitudinal biomarker:
'JM2bs+LME' model assumed the value of PSA changes linearly with respect to its observation time.
'JM2bs+MEM-LQ' model incorporated both linear and quadratic time effects.

Landmark Models



Hazard function h(t) of rPFS in LMcurrent model:
$h\{t|\tilde{s}_{LM}, m(s_{LM}), b\text{PSA}, b\text{LDH}, b\text{HGB}\} = h_0(t)\exp\{\Psi_1 b\text{PSA} + \Psi_2 b\text{LDH} + \Psi_3 b\text{HGB} + \tilde{\gamma}_1 m(s_{LM})\}$,
where $h_0(t)$ represents a baseline hazard function.

Hazard function h(t) of rPFS in LMslope model:
$h\{t|\tilde{s}_{LM}, m'(s_{LM}), b\text{PSA}, b\text{LDH}, b\text{HGB}\} = h_0(t)\exp\{\Psi_1 b\text{PSA} + \Psi_2 b\text{LDH} + \Psi_3 b\text{HGB} + \tilde{\gamma}_1 m(s_{LM}) + \tilde{\gamma}_2 m'(s_{LM})\}$, $h_0(t)$ represents a baseline hazard function.

Abbreviations: *b*, baseline log transformed; HGB, hemoglobin; LDH, lactate dehydrogenase; PSA, prostate-specific antigen.

[a] For simplicity of presentation, subject index i is suppressed in all listed models.

Baseline Covariates

To select baseline covariates for the model comparisons, the following list of candidate predictors of radiographic time to progression was analyzed by fitting a simple Cox proportional hazard model to each:

- Age
- Body mass index
- Eastern Cooperative Oncology Group score
- Total Gleason score
- Pain score (Brief Pain Inventory-Short Form, item 3)
- Baseline log-transformed (*b*) albumin level
- *b* alkaline phosphatase level
- *b* prostate-specific antigen
- *b* lactate dehydrogenase
- *b* hemoglobin

A summary of fitted results is presented in **eTable 2**.

eTable 2. Coefficient Estimates × 10 (SEs × 10) Obtained by Fitting a Simple Cox Proportional Hazard Models Without Considering the Effect of Repeatedly Measured PSA Profile

	Coefficient estimate (SE)	
Covariates	**Basic Cox regression 1 (N = 720)**	**Basic Cox regression 2 (N = 763)**
Age	-0.11 (0.07)	
BMI	-0.11 (0.13)	
ECOG score	-0.88 (1.31)	
Total Gleason	-0.10 (0.48)	
Pain score	-0.08 (0.47)	
*b*PSA	1.61 (0.46)[a]	1.67 (0.41)[a]
*b*ALB	-9.82 (6.96)	
*b*ALP	-0.05 (0.94)	
*b*LDH	4.56 (2.05)[a]	5.18 (1.79)[a]
*b*HGB	-20.74 (6.74)[a]	-19.02 (5.56)[a]
Log likelihood	-1760.99	-1940.82
AIC	3541.97	3887.64

Abbreviations: ALB, albumin; AIC, Akaike information criterion, ALP, alkaline phosphatase; *b*, baseline log transformed; BMI, body mass index; ECOG, Eastern Cooperative Oncology Group; HGB, hemoglobin; LDH, lactate dehydrogenase; PSA, prostate-specific antigen; SE, standard error.

[a] Covariates that are significant under $\alpha = 0.05$ significance level.

eAppendix 2. Details of the Statistical Models Used and How These Were Applied

Longitudinal Biomarker

Throughout the study, we assumed that the longitudinal prostate-specific antigen (PSA) profile (on log scale) is a time-dependent risk factor, denoted by $Y_i(t)$. The subject-specific variation of PSA profile is posited as

$$Y_i(t) = m_i(t) + \epsilon_i(t), \cdots (1),$$

where $m_i(t)$ is the true, unobserved maker process at time t, and $\epsilon_i(t)$ is the random error with mean zero and unknown variance σ^2. The true marker processes of patient i and i' ($i \neq i'$) are assumed independent in this model. An important advantage of this framework is that it can easily accommodate various types of time-varying effects; for example, we can postulate that the marker values change linearly with respect to its observation times:

$$m_i(t) = (\alpha_0 + a_{0i}) + (\alpha_1 + a_{1i})t, \cdots (2),$$

where α_0 and α_1 denote the regression coefficients associated with fixed intercept and slope for time t, a_{0i}, and a_{1i} denote the coefficients associated with the random effects for patient i.

As common in the literature, it is further assumed that the random effects parameters have $[a_{0i}, a_{1i}]^T \sim N(0, \Sigma)$ and are independent of $\epsilon_i(\cdot)$. The individual PSA trajectories in **eFigure 3** (top left) show a trend of PSA decreasing to nadir and then rebound later in the follow-up period, indicating that the trend for PSA over time may not be linear. The linearity assumption for the longitudinal biomarker may be impractical in this case. Therefore, the addition of a quadratic term was used to better characterize the changes in PSA over time. Such effects can be modeled using a mixed effects model with linear and quadratic time effects:

$$m_i(t) = (\alpha_0 + a_{0i}) + (\alpha_1 + a_{1i})t + (\alpha_2 + a_{2i})t^2, \cdots (3),$$

where α_2 is associated with the fixed effect of quadratic trend over time, a_{2i} denotes the coefficients associated with the random effect for the quadratic trend. In this framework, the random effects are assumed $[a_{0i}, a_{1i}, a_{2i}]^T \sim N(0, \Sigma)$ and capture between-patient dependencies. Estimated model parameters for mixed-effect models (2) and (3), without considering the missing data mechanism, are summarized below:

Coefficient Estimates Obtained by Fitting Mixed-Effects Models of Longitudinal PSA Profile

		Coefficient estimate (SE)	
Longitudinal parameters		**Linear mixed-effects model**	**Quadratic mixed-effects model**
α_0:	Fixed intercept	23.35 (0.74)	32.25 (0.64)
α_1:	Fixed linear time effect	-0.45 (0.05)	-4.28 (0.14)
α_2:	Fixed quadratic time effect		0.27 (0.01)
Σ_{11}:	Variance of a_{0i}	36.61	27.24
Σ_{12}:	Covariance of a_{0i} and a_{1i}	-0.62	0.04
Σ_{13}:	Covariance of a_{0i} and a_{2i}		1.11
Σ_{22}:	Variance of a_{1i}	0.10	0.98
Σ_{23}:	Covariance of a_{2i} and a_{3i}		-8.67
Σ_{33}:	Variance of a_{2i}		4.0×10^{-3}
σ^2:	Error variance	10.68	4.53
Log-likelihood		-9408.84	-8172.72
AIC		18829.68	16365.43

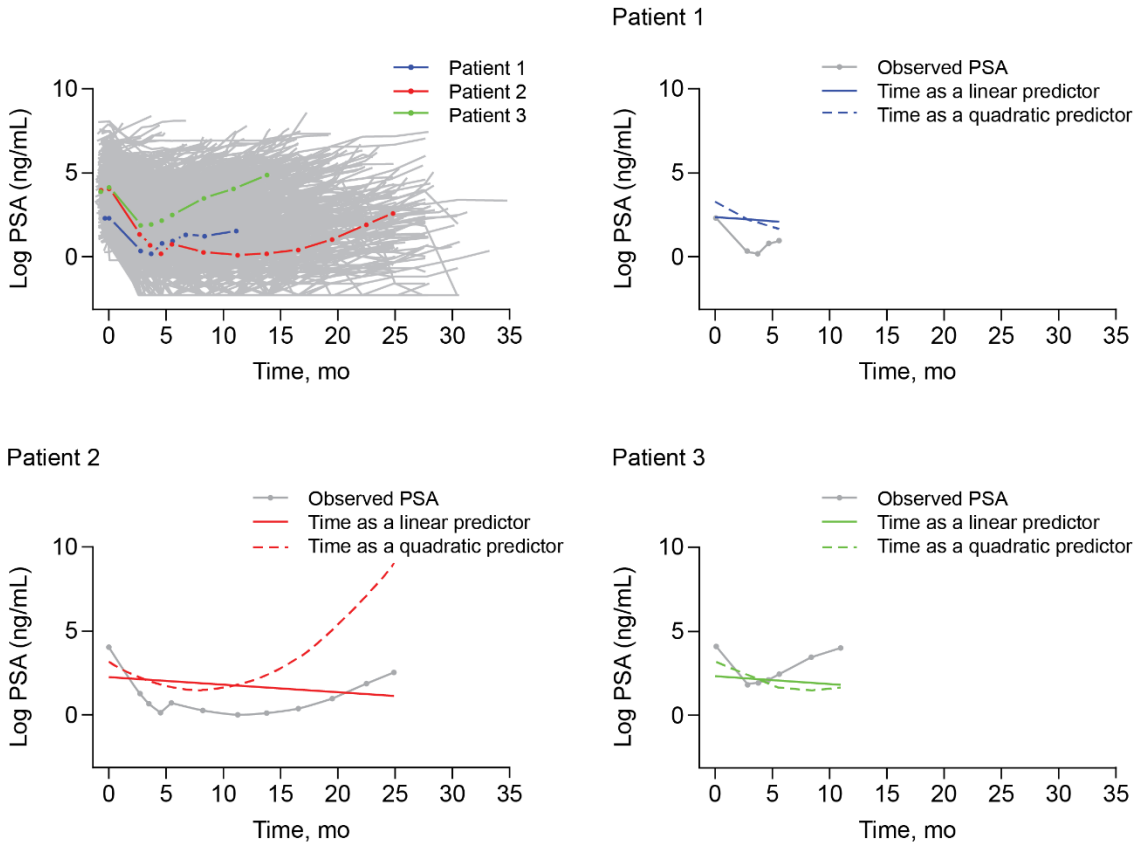
Abbreviations: AIC, Akaike information criterion; PSA, prostate-specific antigen; SE, standard error.
Coefficients and SEs are × 10.

Under the assumption of model (1), we are interested in predicting rPFS for a new patient i_0 from the same population, given that patient i_0 has survived up to time s. To illustrate this idea, let $\boldsymbol{M}_{i_0}(s) = \{m_{i_0}(s^*): 0 \leq s^* \leq s\}$ be the true image of longitudinal PSA profile (on log scale) for this new patient and $\boldsymbol{Y}_{i_0} = \left[Y_{i_0}(t_{i_0 1}), Y_{i_0}(t_{i_0 2}), \ldots, Y_{i_0}(s)\right]^T$ be the PSA measurements (on log scale) recorded up to the survival time s. Using landmark and joint models, we want to compare predicted probabilities of rPFS up to a certain horizon, $s + \omega$ $(s > 0, \omega > 0)$, for this new patient given the complete marker information available on the patient accumulated up to time s:

$$p_{i_0}(s + \omega | s) = Pr\{T_{i_0} \geq s + \omega | T_{i_0} > s, \boldsymbol{\mathcal{D}}_n, \boldsymbol{X}_{i_0}, \boldsymbol{M}_{i_0}(s)\}. \cdots (4),$$

where ω is a clinically relevant window for the horizon. In actual analyses, this study assessed predictive performance of the models at several different follow-up times $s = 4, 7, 10$ months using the window of $\omega = 5$ months.

eFigure 3. Estimated PSA Trajectories Using Mixed-Effects Models



Abbreviations: PSA, prostate-specific antigen.

Top left figure shows the PSA profile (on log scale) collected along the patients' follow-up times from the PREVAIL Phase 3 trial data. When patients' PSA measurements were recorded, the raw values of 0 were assigned to 0.1 ng/mL. The low limits corresponding to -2.3 in the PSA trajectories are the consequence of this (ie, $log(0.1) = -2.3$). Top right and bottom figures show the estimated PSA profile (on log scale) for the three randomly selected patients obtained by fitting model 2 and 3, without considering missing data mechanism.

Joint Model Analysis

In the framework of joint models, two steps are required to associate the true unobserved longitudinal PSA profile to the hazard for radiographic progression: (i) a specification of a model for the true, unobserved marker process $m_i(\cdot)$ and (ii) a specification of a model for the survival outcome. Therefore, a relative risk model that explicitly accounts for the evolution of the true marker process can be written as:

$$h_i\{t|\boldsymbol{X}_i,\ \boldsymbol{M}_i(t)\} = \frac{\log_{dt\to 0} Pr\{t \le T_i < t + dt|T_i \ge t,\ \boldsymbol{M}_i(t), \boldsymbol{X}_i\}}{dt}$$
$$= h_0(t)\exp\{\psi_1 x_{i1} + \cdots + \psi_p x_{ip} + \gamma m_i(t)\} \cdots (5)$$

where $\boldsymbol{M}_i(t) = \{m_i(s^*): 0 \le s^* < t\}$ is the history of the true unobserved longitudinal process, $\boldsymbol{X}_i$ is a vector of baseline covariates (eg, *b*PSA, *b*LDH, *b*HGB) with a corresponding regression coefficient ψ_k $(k = 1, \ldots, p)$, and γ quantifies the strength of the association between the marker $m_i(t)$ and the risk for a radiographic progression. To complete the specification of (5), the choice for the baseline hazard function $h_0(t)$ has to be determined. As a standard option used in survival analysis, the baseline hazard function may be assumed to be of a specific parametric form such as Weibull or Gamma distributed. Alternatively, the baseline hazard function can be modeled flexibly by using step functions or spline-based approaches.[1] This study adopted the latter approach. It is worth noting that using the completely unspecified baseline hazard function may be considered in order to avoid the impact of misspecifying the baseline risk function and therefore leading to possibly underestimated standard errors.[2]

Note that $m_i(\cdot)$ is different from $Y_i(\cdot)$, with the latter contaminated with measurement errors, $\epsilon_i(\cdot)$. Recovering the true PSA profile is available with estimating $m_i(t)$ using an appropriate model assumption. This study employed mixed-effects models with linear and quadratic time effects, as presented in (2) and (3). Under the assumptions of (2) and (3), the longitudinal and survival components are linked through a joint distribution of the longitudinal and time-to-event outcomes $(\boldsymbol{Y}_i,\ T_i^*,\ \delta_i)$. In early development,[3-5] the joint density of $(\boldsymbol{Y}_i,\ T_i^*,\ \delta_i)$ is factored into a product of the densities of $\boldsymbol{M}_i$ and $T_i|\boldsymbol{M}_i$:

$$p(\boldsymbol{M}_i, T_i^*, \delta_i) = \int p(\boldsymbol{M}_i|\tilde{a}_i)\{h(T_i^*|\tilde{a}_i)^{\delta_i} S(T_i^*|\tilde{a}_i)\}p(\tilde{a}_i)d\tilde{a}_i, \cdots (6),$$

where $\tilde{a}_i$ is a vector of random effects, $p(\cdot)$ is a density function, and $S(\cdot)$ is a survival function, and parameters were estimated by maximizing the log likelihood corresponding to the joint distribution of the longitudinal and time-to-event outcomes $(\boldsymbol{Y}_i,\ T_i^*,\ \delta_i)$. In the literature, estimation of the model parameters can be based either on maximum likelihood[2,6] or a Bayesian approach using Markov Chain Monte Carlo (MCMC) algorithms.[7,8] We focused here on the former approach.

Landmark Analysis

Another way to incorporate patients' marker history in survival analysis is using the landmarking analysis proposed by van Houwelingen.[9] The landmarking approach applies a simple Cox proportional hazards model to the data of patients still alive at the time of interest, or *landmark time* s_{LM} (eg, 4-month or 7-month from randomization). Since this approach directly employs the standard Cox proportional hazards model, estimation can be implemented based on a partial log-likelihood method. The primary clinical interest is to: (i) update the prognosis in a dynamic way, taking into account available marker information and (ii) predict survival up to a certain horizon, $s_{LM} + \omega$ (ω: a clinically relevant window for prediction).

For $s_{LM} \le t \le s_{LM} + \omega$, the simplest landmark analysis can be performed by fitting a model defined as

$$h_i\{t|s_{LM}, \tilde{m}_i(s_{LM}), \boldsymbol{X}_i\} = h_0(t|s_{LM})exp\{\psi_1 x_{i1} + \cdots + \psi_p x_{ip} + \gamma_{LM}\tilde{m}_i(s_{LM})\},\ \cdots (7),$$

where the baseline hazard function $h_0(t|s_{LM})$ is assumed completely unspecified and depends on the value of landmark time, γ_{LM} is the unknown regression coefficients associated with the marker value at the landmark time s_{LM}, and $\tilde{m}_i(s_{LM})$ is the last available PSA measurement (on log-scale) recorded prior to the landmark time s_{LM}. The choice of model, landmark time, and horizon may change based on clinical considerations. In actual analyses, this study considered several different follow-up times $s_{LM} \in [4, 7, 10]$ in months, and a window for prediction was set equal to $\omega = 5$ months. An implicit assumption in (7) is that the value of $\tilde{m}_i(\cdot)$ is known at any time point. In the presence of censored data induced by loss to follow-up, not every patient has marker values at the selected landmark points, and thus $\tilde{m}_i(\cdot)$ must be imputed based on a reasonable model of $m_i(\cdot)$. This study used the last observation carried forward (LOCF) approximation considered by van Houwelingen[9] and van Houwelingen and Putter.[10]

PREVAIL Post Hoc Analysis

Competitive Models

The choice of models considered in this study was inspired by Rizopoulos et al.,[11] where the main focus in this paper was to build dynamic predictions of reoperation-free survival probabilities for patients who received subcoronary implantation and a root replacement, using echocardiographic measurements of aortic gradient as a longitudinal biomarker. **eTable 3** shows a complete list of joint and landmark models adopted in this study. To take into account the changes in marker level over time, the subject specific PSA trajectories are modeled using both linear mixed-effects model (LME) posited in (2) and the mixed effects model with linear/quadratic time effects (MEM-LQ) posited in (3).

For joint modeling, flexible types of baseline hazard functions are considered in order to avoid using possibly restrictive parametric assumptions. For the models labeled JM1pwc-JM3pwc, the baseline hazard function $h_0(t)$ is assumed piecewise constant, ie, $h_0(t) = \sum_{q=1}^{7} \xi_q I(\nu_{q-1} \leq \nu_q)$, where $0 = \nu_0 < \cdots < \nu_7$ is a split of time domain, with ν_7 being larger than the largest observed time. For the models labeled JM1bs-JM3bs, the log baseline hazard function is approximated by B-spline basis functions, ie, $log h_0(t) = exp\{\kappa_0 + \sum_{d=1}^{9} \kappa_d B_d(t, q)\}$ where $\kappa^T = [\kappa_0, \kappa_1, \ \ldots, \ \kappa_9]$ are the spline basis coefficients associated with cubic B-spline basis functions $B_d(\cdot,\cdot)$, and q denotes the degree of the B-spline basis functions.

The interpretation of model parameters will not be straightforward when the time-dependent slope or a summary of the whole trajectory up to the same time t is additionally included. For example, the former parameterization can capture situations where two patients show similar PSA levels at a certain time point, but they may differ in the rate of change of the PSA profiles. To compare the predictive performance of joint modeling with the landmark approach, we fitted landmark models with similar types of marker effects. These models are labeled LMcurrent, LMslope, and LMarea in **eTable 3**. Since our focus is more on dynamic predictions of survival probabilities during the early phase after the study entry, we assumed $s_{LM} \in \{4,7,10\}$ in months are medically relevant landmark points, and predictions were carried out using the window of $\omega = 5$ months. Note that the landmarking analysis requires extrapolating the last marker value in the case where the marker value is not available at the selected landmark points. Thus, corrected marker values $\tilde{m}(s_{LM})$ are derived using the last available PSA value of each patient before month s_{LM}. Similarly, corrected marker slopes $\tilde{m}'(s_{LM})$ are derived using the last two available measurements before month s_{LM}. For the same landmark points and the same prediction window, we also fitted the landmark datasets constructed using a relative marker change from the baseline $\frac{\tilde{m}(s_{LM}) - m(s_0)}{m(s_0)}$, where $m(s_0)$ is the baseline PSA and using a relative marker change from the previous marker assessment $\frac{\tilde{m}(s_{LM}) - \tilde{m}(s_{LM-1})}{\tilde{m}(s_{LM-1})}$, where the difference is calculated using the last two available measurements before s_{LM}, with $s_{LM-1} < s_{LM}$. These models are labeled LMbase and LMpre in **eTable 3**. Unlike any other models listed in the table, maker values were not log transformation in order to provide easy interpretation between the event time of interest and the relative changes of PSA profile.

eTable 3. Joint and Landmark Models Used for PREVAIL Post Hoc Analysis, Motivated by Rizopoulos et al[11]

Model ($t > s_{LM}$)	Hazard	Label
$h\{t\|b\text{PSA},\ b\text{LDH},\ b\text{HGB}, \boldsymbol{M}(t)\}$[a]	$h_0(t)exp\{\psi_1 b\text{PSA} + \psi_2 b\text{LDH} + \psi_3 b\text{HGB} + \gamma_1 m(t)\}$ $h_0(t)exp\{\psi_1 b\text{PSA} + \psi_2 b\text{LDH} + \psi_3 b\text{HGB} + \gamma_1 m(t) + \gamma_2 m'(t)\}$ $h_0(t)exp\left\{\psi_1 b\text{PSA} + \psi_2 b\text{LDH} + \psi_3 b\text{HGB} + \gamma_1 m(t) + \gamma_3 \int_0^t m(s)ds\right\}$	JM1pwc JM2pwc JM3pwc
$h\{t\|b\text{PSA},\ b\text{LDH},\ b\text{HGB}, \boldsymbol{M}(t)\}$[b]	$h_0(t)exp\{\psi_1 b\text{PSA} + \psi_2 b\text{LDH} + \psi_3 b\text{HGB} + \gamma_1 m(t)\}$ $h_0(t)exp\{\psi_1 b\text{PSA} + \psi_2 b\text{LDH} + \psi_3 b\text{HGB} + \gamma_1 m(t) + \gamma_2 m'(t)\}$ $h_0(t)exp\left\{\psi_1 b\text{PSA} + \psi_2 b\text{LDH} + \psi_3 b\text{HGB} + \gamma_1 m(t) + \gamma_3 \int_0^t m(s)ds\right\}$	JM1bs JM2bs JM3bs
$h\{t\|s_{LM}, \tilde{m}(s_{LM}), b\text{PSA},\ b\text{LDH},\ b\text{HGB}\}$[c]	$h_0(t)exp\{\psi_1 b\text{PSA} + \psi_2 b\text{LDH} + \psi_3 b\text{HGB} + \gamma_1 \tilde{m}(s_{LM})\}$ $h_0(t)exp\{\psi_1 b\text{PSA} + \psi_2 b\text{LDH} + \psi_3 b\text{HGB} + \gamma_1 \tilde{m}(s_{LM}) + \gamma_2 \tilde{m}'(s_{LM})\}$ $h_0(t)exp\left\{\psi_1 b\text{PSA} + \psi_2 b\text{LDH} + \psi_3 b\text{HGB} + \gamma_1 \tilde{m}(s_{LM}) + \gamma_3 \sum_{s=0}^{s_{LM}} \tilde{m}\ (s)\Delta s\right\}$	LMcurrent LMslope LMarea
$h\{t\|s_{LM}, \tilde{m}(s_{LM}), m(s_0), b\text{PSA},\ b\text{LDH},\ b\text{HGB}\}$ (s_0: baseline)[d]	$h_0(t)exp\left\{\psi_1 b\text{PSA} + \psi_2 b\text{LDH} + \psi_3 b\text{HGB} + \gamma_1 \frac{\tilde{m}(s_{LM}) - m(s_0)}{m(s_0)}\right\}$	LMbase
$h\{t\|s_{LM}, \tilde{m}(s_{LM}), \tilde{m}(s_{LM-1}), b\text{PSA},\ b\text{LDH},\ b\text{HGB}\}$ (s_{LM-1}: previous assessment point)[e]	$h_0(t)exp\left\{\psi_1 b\text{PSA} + \psi_2 b\text{LDH} + \psi_3 b\text{HGB} + \gamma_1 \frac{\tilde{m}(s_{LM}) - \tilde{m}(s_{LM-1})}{\tilde{m}(s_{LM-1})}\right\}$	LMpre

Abbreviations: *b*, baseline log transformed; JM, joint model; HGB, hemoglobin; LDH, lactate dehydrogenase; LMarea, landmark model constructed using area under the step function defined from the PSA measurement; LMbase, landmark model constructed using a relative marker change from the baseline; LMcurrent, landmark model with current PSA level; LMpre, landmark model using a relative marker change from the previous marker assessment; LMslope, landmark model with additional PSA change; PSA, prostate-specific antigen.

For simplicity of presentation, subject index i is suppressed in all listed models. Models indicated with 1-3 incorporated the log-transformed PSA marker values for the analysis. Models indicated with 4-5 did not apply the log transformation to the PSA marker values for easy interpretation.

[a] $h_0(t) = \sum_{q=1}^{7} \xi_q I(v_{q-1} \le v_q)$ where $0 = v_0 < \cdots < v_7$ denotes a split of time scale; $m'(t) = \frac{m(t)}{dt}$.

[b] $logh_0(t) = exp\{\kappa_0 + \sum_{d=1}^{9} \kappa_d B_d(t, q)\}$, where $\kappa^T = [\kappa_0, \kappa_1, \ \ldots, \ \kappa_9]$ are the spline basis coefficients associated with cubic B-spline basis functions $B_d(\cdot,\cdot)$, and q denotes the degree of the B0-spline basis functions.

[c] $\tilde{m}(s_{LM})$ = last available PSA value of each patient before month s_{LM}; $\tilde{m}'(s_{LM})$ = slope defined from the last two available measurements; $\sum_{s=0}^{s_{LM}} \tilde{m}\ (s)\Delta s$ = area under the step function defined from the observed PSA measurements up to month s_{LM}.

[d] $\tilde{m}(s_{LM})$ = last available PSA value of each patient before month s_{LM}; $\frac{\tilde{m}(s_{LM}) - m(s_0)}{m(s_0)}$ = a relative change of PSA from the baseline.

[e] $\tilde{m}(s_{LM})$ = last available PSA value of each patient before month s_{LM}; $\frac{\tilde{m}(s_{LM}) - \tilde{m}(s_{LM-1})}{\tilde{m}(s_{LM-1})}$ = a relative change of PSA profile from the previous PSA assessment with $s_{LM-1} < s_{LM}$.

eTable 4. Reference Models Used in the PREVAIL Post Hoc Analyses Model Evaluation

Model ($t > s_{LM}$)	Hazard	Label
$h\{t \mid b\text{PSA},\ b\text{LDH},\ b\text{HGB}, \boldsymbol{M}(t)\}$[a]	$h_0(t)exp\{\psi_1 b\text{PSA} + \psi_2 b\text{LDH} + \psi_3 b\text{HGB} + \gamma_1 m(t) + \gamma_2 m'(t)\}$	JM2bs
$h\{t \mid s_{LM}, \tilde{m}(s_{LM}), b\text{PSA},\ b\text{LDH},\ b\text{HGB}\}$[b]	$h_0(t)exp\{\psi_1 b\text{PSA} + \psi_2 b\text{LDH} + \psi_3 b\text{HGB} + \gamma_1 \tilde{m}(s_{LM})\}$ $h_0(t)exp\{\psi_1 b\text{PSA} + \psi_2 b\text{LDH} + \psi_3 b\text{HGB} + \gamma_1 \tilde{m}(s_{LM}) + \gamma_2 \tilde{m}'(s_{LM})\}$	LMcurrent LMslope

Abbreviations: *b*, baseline log-transformed; HGB, hemoglobin; JM2bs, joint model; LDH, lactate dehydrogenase; LMcurrent, landmark model with current PSA level; LME, linear mixed effect; LMslope, landmark model with additional PSA change; MEM-LQ, nonlinear mixed effect; PSA, prostate-specific antigen.

[a] $log h_0(t) = exp\{\kappa_0 + \sum_{d=1}^{9} \kappa_d B_d(t, q)\}$, where $\kappa^T = [\kappa_0, \kappa_1, \ \ldots, \ \kappa_9]$ are the spline basis coefficients associated with cubic B-spline basis functions $B_d(\cdot,\cdot)$, and q denotes the degree of the B0-spline basis functions[§] $\tilde{m}(s_{LM})$ = last available.

[b] PSA value of each patient before month s_{LM}; $\tilde{m}'(s_{LM})$ = slope defined from the last two available measurements.

Model Evaluation

Predictive performance of the joint models and landmark models was assessed in terms of discrimination and calibration abilities, using a prediction window of 5 months. There is a substantial amount of literature on the appropriate measure of predictive accuracy in the setting of dynamic prediction and in the potential presence of competing risks; for detailed reviews, refer to Zheng and Heagerty,[12] van Houwelingen,[9] Blanche et al.,[13] and Rizopoulos et al.,[4,11] and the references therein.

Intuitively, the model is well calibrated if the true survival probabilities do not differ from the predicted ones. A common approach for formalizing this concept is based on the expected prediction errors derived by using a quadratic loss function:

$$PE(t|s) = E[\{N_i(t) - p_i(t|s)\}^2], \dots (8),$$

where $N_i(t) = I(T_i > t)$ is the status of rPFS at time t, and $p_i(t|s) = Pr\{T_i \geq t | T_i > s, \mathrm{D}_n, \boldsymbol{X}_i, \boldsymbol{M}_i(s)\}$ is the conditional survival probability given the longitudinal marker information up to time s. Another diagnostic measure for the predictive accuracy is based on the area under the receiver operating characteristic (ROC) curve, which evaluates how the model accurately discriminates the ranking of patients' survival probability in concordance with the ranking of their disease status within a selected period. To illustrate this idea, consider any random pair of patients i and j (i) who have survived up to time t and (ii) who have provided a series of marker values up to clinically meaningful time $t + \omega$. Then, the discriminative capability of the model can be assessed by estimating the AUC curve defined as

$$AUC(t, w) = Pr\big[p_i(t + w|t) < p_j(t + w|t) | \{T_i \in (t, t + w)\} \cap \{T_j > t + w\}\big]. \dots (9)$$

In a situation where patient i experiences a radiographic progression within the interval $(t,\ t + w]$, and patient j does not experience the event at the same period, one would expect to get a higher AUC value to demonstrate a successful discriminative capability. In general, AUC equaling 1 indicates maximum discrimination, whereas AUC equaling 0.5 indicates no discriminative ability.

In our post hoc analysis, a significant issue in assessing model performance is that the same data cannot be used to fit the model and to assess its predictive accuracy at the same time. This would result in a model that perfectly fits the data, but predictions might fail when new independent data are utilized. Such an issue is often called overfitting. A popular strategy for bypassing this issue is to split the data into two parts, one for model development based on a training set and the other one for assessment of the predictive performance based on a validation set. To employ a more systematic way for assessing the model, we applied 5-fold cross validation with 50 repetitions. Specifically, for each of the 50 repetitions, the data set were split into five subsets of similar size, and then the model was fitted five times using the four subsets, and the remaining subset was used to calculate the accuracy measures defined in (8) and (9). The results from the repeated 5-fold cross validation are displayed in **eTable 5** and **eTable 6**.

eTable 5. Cross-Validated PE (t + 5 | t) Obtained From the Selected Joint Models and Landmark Models for the Follow-Up Times 2, 4, 7, and 10 Months[a]

Follow-up time, mo	JM2pwc+LME	JM2pwc+MEM-LQ	JM2bs+LME	JM2bs+MEM-LQ
2	0.096 (0.005)	0.098 (0.024)	0.096 (0.005)	0.095 (0.019)
4	0.129 (0.006)	0.130 (0.027)	0.128 (0.006)	0.129 (0.021)
7	0.149 (0.006)	0.154 (0.024)	0.150 (0.006)	0.150 (0.007)
10	0.147 (0.007)	0.144 (0.024)	0.150 (0.006)	0.138 (0.005)
Follow-up time, mo	**LMcurrent**	**LMslope**	**LMbase**	**LMpre**
2	0.106 (0.008)	0.107 (0.008)	NA	NA
4	0.125 (0.006)	0.123 (0.007)	0.127 (0.007)	0.133 (0.005)
7	0.146 (0.005)	0.144 (0.005)	0.148 (0.007)	0.159 (0.005)
10	0.140 (0.008)	0.132 (0.008)	0.147 (0.009)	0.143 (0.006)

Abbreviations: LM current, landmark model with current PSA level; LM Slope, landmark model with additional PSA change; LMbase, landmark model constructed using a relative marker change from the baseline; LMpre, landmark model using a relative marker change from the previous marker assessment; PSA, prostate-specific antigen; SE, standard error.

The JM2bs+MEM-LQ model generally produced smaller predictions errors (indicated by shading) compared with the other joint models. For the landmark models, the LMslope model provided smaller prediction errors (indicated by shading) for the follow up times of month 4, 7, and 10 compared with other landmark approaches.

[a] The values in the parenthesis are the estimated SEs across the repetitions.

eTable 6. Cross-Validated AUC Obtained From the Selected Joint Models and Landmark Models for the Follow-Up Times 2, 4, 7, and 10 Months[a]

Follow-up time	JM2pwc+LME	JM2pwc+MEM-LQ	JM2bs+LME	JM2bs+MEM-LQ
2 months	0.618 (0.022)	0.600 (0.035)	0.626 (0.022)	0.608 (0.025)
4 months	0.732 (0.016)	0.746 (0.035)	0.730 (0.017)	0.751 (0.022)
7 months	0.714 (0.015)	0.733 (0.035)	0.706 (0.016)	0.739 (0.014)
10 months	0.659 (0.021)	0.773 (0.032)	0.653 (0.022)	0.780 (0.016)

Follow-up time	LMcurrent	LMslope	LMbase	LMpre
2 months	0.614 (0.028)	0.620 (0.028)	NA	NA
4 months	0.752 (0.020)	0.769 (0.021)	0.737 (0.020)	0.706 (0.018)
7 months	0.751 (0.015)	0.770 (0.014)	0.718 (0.022)	0.662 (0.017)
10 months	0.743 (0.018)	0.800 (0.018)	0.653 (0.030)	0.737 (0.018)

Abbreviations: AUC: area under the curve; JM, joint model; LMarea, landmark model constructed using area under the step function defined from the PSA measurement; LMcurrent, landmark model with current PSA level; LME, linear mixed effect; LMbase, landmark model constructed using a relative marker change from the baseline; LMpre, landmark model using a relative marker change from the previous marker assessment; LMslope, landmark model with additional PSA change; MEM-LQ, nonlinear mixed effect; PSA, prostate-specific antigen; SE, standard error.

The JM2bs+MEM-LQ model yields relatively high AUCs (indicated by shading), ranging from 0.739 to 0.780, for the follow-up times of month 4, 7, and 10. For the landmark models, the LMslope model gives higher AUCs (indicated by shading), ranging from 0.769 to 0.800, for the follow-up times of month 4, 7, and 10 compared with other landmark approaches.

[a] The values in the parenthesis are the estimated SEs across the repetitions.

Key Model Evaluation Results

- In **eTable 5**, along the follow-up times of months 4, 7, and 10, JM2bs+MEM-LQ model generally produced smaller predictions errors compared with the other joint models. For the landmark models, the LMslope model provided smaller prediction errors along the follow-up times of months 4, 7, and 10 compared with other landmark approaches
- **eTable 6** shows that the proposed dynamic prediction does not have good discriminative ability, irrespective of the choice of models when the follow-up time is set equal to month 2. This might be due to the fact that most patients did not experience radiographic progression by month 2, and thus it is more difficult to determine the ranking of patients' survival probability in this circumstance. At the same time, this might be due to the consequence of utilizing minimized marker information by selecting the prediction time right after the treatment was initiated
- It is evident from the results that the JM2bs+MEM-LQ model yields relatively high AUCs, ranging from 0.739 to 0.780, along the follow-up times of month 4, 7, and 10. For the landmark models, the LMslope model gives higher AUCs, ranging from 0.769 to 0.800, along the follow-up times of months 4, 7, and 10 compared with other landmark approaches
- Landmarking and joint modeling in this particular data set suggests that the dynamic prediction based on the JM2bs+MEM-LQ and LMslope models are uniformly better in terms of calibration and discrimination than any other modeling approaches considered in this study

Data Preprocessing and Estimated Model Parameters

Patient data in the PREVAIL trial requires data preprocessing to ensure accurate and clinically meaningful risk assessments. For all models listed in **eTable 3**, PSA measurements taken before randomization and/or after the occurrence of radiographic progressions were not used in the analysis. In all cases listed in the table, we excluded patients with missing baseline covariates (ie, *b*PSA, *b*LDH, *b*HGB) from the analysis set. In the case of LMbase and LMpre models, if patients' PSA measurements are only available at baseline, these patients were excluded from the analysis set, as it is not possible to calculate the relative changes. The parameter estimates and the estimated standard errors for joint modeling are presented in **eTable 7**. For the LMcurrent, LMslope, and LMarea models, the estimated results are presented in **eTable 8**. For the LMbase and LMpre models, the estimated results are shown in **eTable 9**.

Key Data Preprocessing and Estimated Model Parameters Results

- In joint modeling, the strength of the association between $m_i(t)$ (ie, log PSA profile) and the risk for the event was estimated differently with respect to how the underlying PSA trajectories were modeled (MEM-LQ vs LME). For example, a unit increase in the level of log PSA profile yielded a 1.48-fold increase and a 1.82-fold increase in a patient's risk for the event when the data was fitted with JM2bs + MEM-LQ and JM2bs + LME, respectively. Furthermore, the slope parameters were quantified differently with respect to the posited mixed effects models as well (**eTable 7**)
- In joint modeling, the estimated model parameters were affected by the underlying assumptions for the baseline hazard function, but with only a negligible difference (**eTable 7**)
- In the landmarking approach, the strength of the association between $\widetilde{m}(s_{LM})$ and the risk for the event kept quite similar irrespective of the choice of the landmark times, particularly when the LMcurrent model was considered (**eTable 8**)
- In the landmarking analysis, a stronger association was detected between the time-dependent slope and the hazard for the radiographic progression when s_{LM}=10 was selected compared to when s_{LM} =4 or 7 was used (**eTable 8**)
- The slope parameters estimated by fitting the joint models were obviously larger than the slope parameters estimated by fitting the landmark models regardless of the choice of the landmark times (**eTable 7 and eTable 8**)
- Both in joint modeling and in landmarking, the hazard for the radiographic progression was associated with both the current value and the slope of the log PSA trajectories. Furthermore, models that incorporated the time-dependent slopes provided larger likelihood and smaller AIC values than their counterparts (**eTable 7 and eTable 8**)
- The hazard for the radiographic progression was associated with the level of relative change of PSA from the baseline irrespective of the choice of the landmark times (**eTable 9**)

eTable 7. Coefficient Estimates × 10 (SEs × 10) of JM1pwc–JM3pwc and JM1bs–JM3bs[a]

		Coefficient estimate (SE)				
Covariate	**JM1pwc+LME**	**JM2pwc+LME**	**JM3pwc+LME**	**JM1pwc+ MEM-LQ**	**JM2pwc+ MEM-LQ**	**JM3pwc+ MEM-LQ**
$m(t)$[b]	3.77 (0.42)[c]	5.31 (0.48)[c]	3.45 (0.60)[c]	5.56 (0.40)[c]	3.86 (0.45)[c]	6.54 (0.50)[c]
$m'(t)$[b]		-68.52 (9.08)[c]			38.68 (4.27)[c]	
$\int_0^t m(s)\,ds$[b]			0.04 (0.05)			-0.13 (0.04)[c]
$b\mathrm{PSA}$[b]	-1.42 (0.53)[c]	-3.02 (0.58)[c]	-1.53 (0.53)[c]	-3.03 (0.52)[c]	-2.00 (0.58)[c]	-2.65 (0.53)[c]
$b\mathrm{LDH}$[b]	4.33 (1.85)[c]	5.53 (2.04)[c]	4.11 (1.88)[c]	2.77 (1.86)	-0.04 (2.09)	2.46 (1.85)
$b\mathrm{HGB}$[b]	-17.77 (5.53)[c]	-19.40 (5.73)	-19.83 (5.50)[c]	-13.57 (5.55)	-21.04 (6.29)[c]	-11.58 (5.60)[c]
Longitudinal parameter						
α_0: fixed intercept	24.98 (0.82)[c]	25.71 (0.73)[c]	25.06 (0.81)[c]	28.03 (0.48)[c]	27.83 (0.47)[c]	27.98 (0.47)[c]
α_1: fixed slope	-0.50 (0.04)[c]	-0.56 (0.04)[c]	-0.50 (0.04)[c]	-5.09 (0.10)[c]	-5.03 (0.09)[c]	-5.08 (0.09)[c]
α_2: fixed quadratic				0.32 (0.00)[c]	0.32 (0.00)[c]	0.32 (0.00)[c]
Log likelihood	-10815.78	-10789.08	-10815.47	-9830.92	-9769.76	-9825.85
AIC	21665.56	21614.17	21666.94	19703.84	19583.52	19695.70

		Coefficient estimate (SE)				
Covariate	**JMbs+LME**	**JM2bs+LME**	**JM3bs+LME**	**JM1bs+MEM-LQ**	**JM2bs+MEM-LQ**	**JM3bs+MEM-LQ**
$m(t)$[b]	3.81 (0.42)[c]	6.08 (0.55)[c]	4.30 (0.65)[c]	5.61 (0.43)[c]	3.95 (0.46)[c]	7.44 (0.54)[c]
$m'(t)$[b]		-84.21 (10.84)[c]			37.48 (5.04)[c]	
$\int_0^t m(s)\,ds$[b]			-0.05 (0.05)			-0.23 (0.05)[c]

bPSA[b]	-1.36 (0.53)[c]	-3.53 (0.64)[c]	-1.28 (0.53)[c]	-2.99 (0.55)[c]	-1.86 (0.58)[c]	-2.48 (0.54)[c]
bLDH[b]	5.40 (1.80)[c]	7.95 (2.15)[c]	5.96 (1.77)[c]	3.23 (1.85)	1.60 (2.02)	2.40 (1.85)
bHGB[b]	-12.38 (5.60)[c]	-11.11 (5.97)	-7.46 (5.67)	-9.47 (5.59)	-10.73 (6.31)	-6.93 (5.70)
Longitudinal parameter						
α_0: fixed intercept	25.01 (0.82)[c]	25.81 (0.77)[c]	24.93 (0.83)[c]	28.05 (0.48)[c]	27.82 (0.47)[c]	28.01 (0.47)[c]
α_1: fixed slope	-0.50 (0.04)[c]	-0.58 (0.04)[c]	-0.50 (0.04)[c]	-5.10 (0.10)[c]	-5.02 (0.09)[c]	-5.08 (0.09)[c]
α_2: fixed quadratic				0.32 (0.00)[c]	0.32 (0.00)[c]	0.32 (0.00)[c]
Log likelihood	-10815.21	-10778.96	-10818.70	-9828.46	-9781.94	-9816.17
AIC	21668.42	21597.92	21677.40	19702.93	19611.89	19680.35

Abbreviations: AIC, Akaike information criterion; *b*, baseline log transformed; HGB, hemoglobin; JM, joint model; LDH, lactate dehydrogenase; PSA, prostate-specific antigen; SE, standard error.

Data preprocessing yielded N = 763 patients.

The shaded areas indicate models with smaller AIC value and larger log likelihood.

Coefficients and SEs are × 10.

[a] The log PSA profile is incorporated using the LME (left) and MEM-LQ (right).

[b] Variables are log transformed before fitting the models.

[c] Covariates that are significant under $\alpha = 0.05$ significance level.

eTable 8. Coefficient Estimates × 10 (SEs × 10) of LMcurrent, LMslope, and LMarea With s_{LM}**= 4, 7, 10**

	Coefficient estimate (SE)								
	$s_{LM} = 4$ **(N = 679)**			$s_{LM} = 7$ **(N = 597)**			$s_{LM} = 10$ **(N = 512)**		
Covariate	**LMcurrent**	**LMslope**	**LMarea**	**LMcurrent**	**LMslope**	**LMarea**	**LMcurrent**	**LMslope**	**LMarea**
$\tilde{m}(s_{LM})$[a]	3.42 (0.43)[b]	2.90 (0.43)[b]	3.25 (0.60)[b]	3.52 (0.45)[b]	3.34 (0.45)[b]	5.59 (1.09)[b]	3.45 (0.50)[b]	2.54 (0.51)[b]	7.93 (1.29)[b]
$\tilde{m}'(s_{LM})$[a]		7.58 (1.78)[b]			4.92 (1.28)[b]			20.46 (3.82)[b]	
$\Sigma_{s=0}^{s_{LM}} \tilde{m}(s)\Delta s$[a]			0.11 (0.30)			-0.48 (0.23)[b]			-0.68 (0.18)[b]
bPSA[a]	-1.27 (0.59)[b]	-0.75 (0.59)	-1.58 (0.98)	-1.08 (0.63)	-0.76 (0.64)	0.36 (0.91)	-1.25 (0.72)	-0.16 (0.72)	1.29(0.95)
bLDH[a]	5.27 (2.02)[b]	3.76 (2.04)	5.24 (2.03)[b]	2.50 (2.52)	0.93 (2.55)	1.97 (2.52)	0.85 (3.54)	-3.24 (3.56)	-1.97 (3.54)
bHGB[a]	-19.56 (5.96)[b]	-20.88 (6.08)[b]	-19.46 (5.96)[b]	-21.35 (6.59)[b]	-20.60 (6.63)[b]	-21.69 (6.61)[b]	-22.33 (7.99)[b]	-19.07 (8.23)[b]	-21.09 (8.09)[b]
Log likelihood	-1584.05	-1575.62	-1583.98	-1231.51	-1225.85	-1229.44	-849.97	-836.87	-843.15
AIC	3176.11	3161.24	3177.96	2471.02	2461.71	2468.88	1707.94	1683.73	1696.3

Abbreviations: AIC, Akaike information criterion; *b*, baseline log transformed; HGB, hemoglobin; LDH, lactate dehydrogenase; LMarea, landmark model constructed using area under the step function defined from the PSA measurement; LMcurrent, landmark model with current PSA level; LMpre landmark model constructed using a relative marker change from the baseline; LMslope, landmark model with additional PSA change; PSA, prostate-specific antigen; SE, standard error.

The shaded areas indicate models with smaller AIC value and larger log likelihood.

Coefficients and SEs are × 10.

[a] Variables are log transformed before fitting the models.

[b] Covariates that are significant under $\alpha = 0.05$ significance level.

eTable 9. Coefficient Estimates × 10 (SEs × 10) of LMbase and LMpre With s_{LM}**= 4, 7, 10**

	Coefficient estimate (SE)					
	$s_{LM} = 4$ **(N = 673)**		$s_{LM} = 7$ **(N = 597)**		$s_{LM} = 10$ **(N = 512)**	
Covariate	**LMbase**	**LMpre**	**LMbase**	**LMpre**	**LMbase**	**LMpre**
$\frac{\widetilde{m}(s_{LM}) - m(s_0)}{m(s_0)}$	7.76 (0.87)[a]		5.09 (0.67)[a]		2.35 (0.49)[a]	
$\frac{\widetilde{m}(s_{LM}) - \widetilde{m}(s_{LM-1})}{\widetilde{m}(s_{LM-1})}$		7.01 (1.10)[a]		0.24 (0.43)		3.29 (0.53)[a]
$b\text{PSA}$[b]	2.21 (0.44)[a]	1.71 (0.45)[a]	2.22 (0.50)[a]	1.88 (0.50)[a]	1.74 (0.60)[a]	1.57 (0.59)[a]
$b\text{LDH}$[b]	3.76 (2.11)	2.70 (2.13)	1.98 (2.50)	2.92 (2.52)	1.52 (3.32)	-1.68 (3.37)
$b\text{HGB}$[b]	-19.63 (6.06)[a]	-20.87 (6.19)[a]	-22.51 (5.09)[a]	-21.92 (6.85)[a]	-23.55 (8.26)[a]	-14.53 (8.46)
Log likelihood	-1573.42	-1583.39	-1251.13	-1265.77	-872.04	-862.93
AIC	3154.84	3174.78	2510.26	2539.54	1752.08	1733.86

Abbreviations: AIC, Akaike information criterion; HGB, hemoglobin; LDH, lactate dehydrogenase; LMbase, landmark model constructed using a relative marker change from the baseline assessment; LMcurrent, landmark model with current PSA level; LMpre, landmark model constructed using a relative marker change from the baseline; PSA, prostate-specific antigen; SE, standard error.

The shaded areas indicate models with smaller AIC value and larger log likelihood.

Coefficients and SEs are × 10.

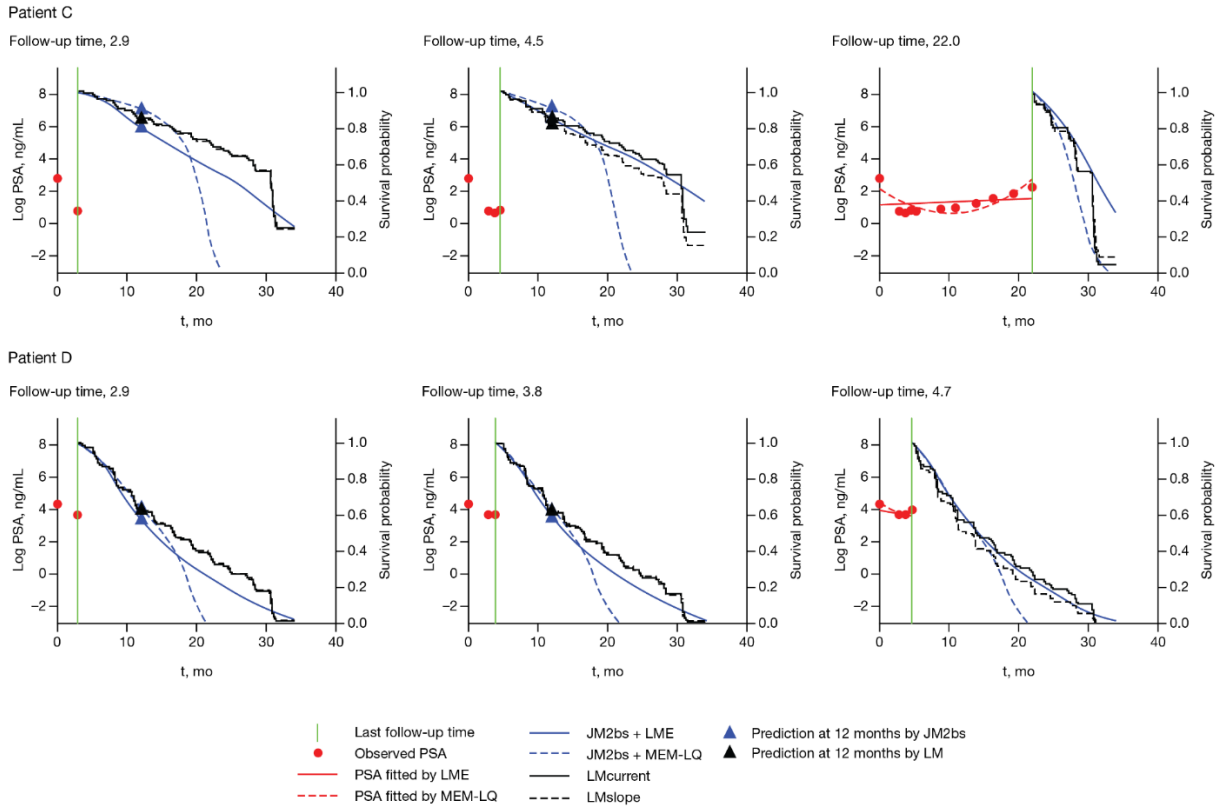
[a] Covariates that are significant under $\alpha = 0.05$ significance level.

[b] Variables are log transformed before fitting the models.

Predictive Performance of Longitudinal Submodels

In **eFigure 4**, prediction of conditional survival probability at specific follow-up times are compared using the estimation method of the selected models. The comparison is for Patient C and Patient D.

eFigure 4. Individual PSA Evolution and Dynamic Predictions of Conditional Survival Probability for Patient C vs Patient D



Abbreviations: JM2bs, joint model; LMcurrent, landmark model with current PSA level; LME, linear mixed effect; LMslope, landmark model with additional PSA change; MEM-LQ, nonlinear mixed effect; PSA, prostate-specific antigen; t, survival time.

Additional data showing comparison of LME and MEM-LQ fit by in-sample and out-of-sample mean squared prediction errors obtained based on the training and test set, with the standard error of squared prediction errors in parentheses are shown in **eTable 10.**

eTable 10. Comparison of LME and MEM-LQ Fit by In-Sample and Out-of-Sample MSPEs[a] Obtained Based on the Training and Test Set With the Standard Error of Squared PEs in Parentheses

Survival submodel	Longitudinal submodel	In-sample MSPE[b,d]	Out-of-sample MSPE[c,d]
JM1pwc	LME MEM-LQ	0.751 (0.911) 0.299 (0.411)	0.833 (1.037) 0.333 (0.446)
JM2pwc	LME MEM-LQ	0.737 (0.908) 0.312 (0.424)	0.814 (1.031) 0.348 (0.461)
JM3pwc	LME MEM-LQ	0.748 (0.911) 0.302 (0.413)	0.830 (1.037) 0.339 (0.451)
JM1bs	LME MEM-LQ	0.750 (0.911) 0.303 (0.414)	0.832 (1.037) 0.340 (0.453)
JM2bs	LME MEM-LQ	0.735 (0.907) 0.310 (0.421)	0.811 (1.028) 0.347 (0.460)
JM3bs	LME MEM-LQ	0.751 (0.911) 0.304 (0.414)	0.833 (1.037) 0.341 (0.454)

Abbreviations: LME, linear mixed effects; MEM-LQ, nonlinear mixed effects; MSPE, mean squared prediction error; PE, prediction error.

[a] MSPE is defined as $\frac{1}{n}\sum_{i=1}^{n}\frac{1}{J_i}\sum_{j=1}^{J_i}\{Y_i(t_j)-\hat{Y}_i(t_j)\}^2$, where $\hat{Y}_i(t_j)$ is the longitudinal marker process estimated for the i^{th} patient at the jth observation time t_j $(j=1,\cdots,J_i)$.

[b] In-sample MSPE measures the performance of model estimation based on the training set.

[c] Out-of-sample MSPE measures the performance of model prediction based on the test set.

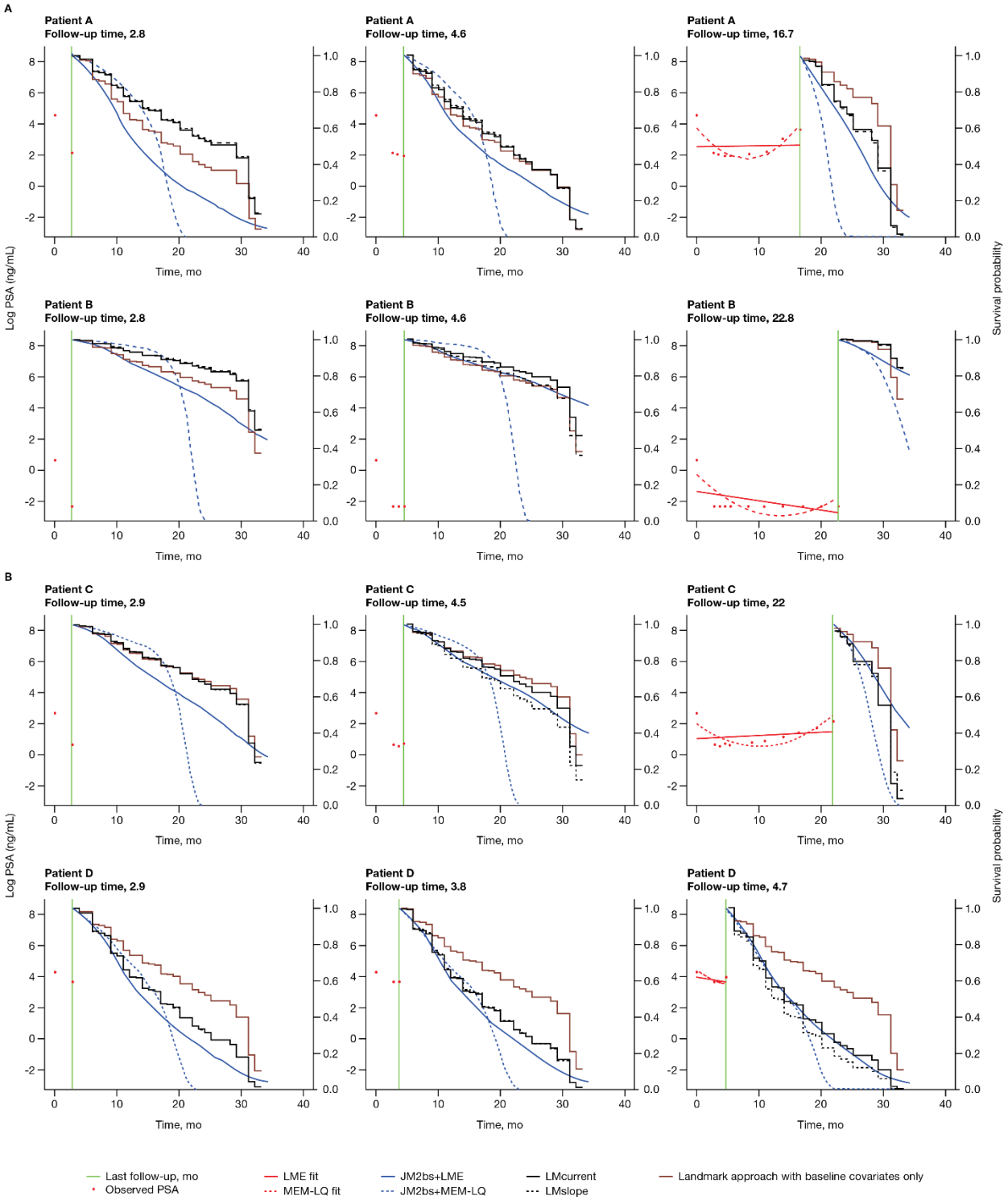
[d] Values in parenthesis are the standard error of the squared PEs: $\frac{1}{J_i}\sum_{j=1}^{J_i}\{Y_i(t_j)-\hat{Y}_i(t_j)\}^2$ $(i=1,\cdots,n)$.

Comparisons With Standard Cox Model

As presented in **eFigure 5**, we found a noticeable difference in predictions between the Cox PH models and other dynamic prediction models. The difference is due to the fact that the standard Cox PH model only incorporates the baseline covariates, and thus the predictions could not be updated although additional marker information was cumulated along the follow-up times.

The difference in the predictions was much more substantial when predictions were made at a later follow-up than at earlier times. This finding was consistent regardless of the patients' PSA kinetics or PSA response.

eFigure 5. Individual PSA Evolution and Dynamic Predictions of Conditional Survival Probability Based on the Standard Cox Model Compared With the 4 Selected Models: Patient A vs Patient B and Patient C vs Patient D



Abbreviations: JM2bs, joint model; LMcurrent, landmark model with current PSA level; LME, linear mixed effects; LMslope, landmark model with additional PSA change; MEM-LQ, nonlinear mixed effect; PSA, prostate-specific antigen; t, survival time.

eReferences

1. Ruppert D, Wand MP, Carroll RJ. *Semiparametric Regression.* Cambridge University Press; 2003.
2. Hsieh F, Tseng YK, Wang JL. Joint modeling of survival and longitudinal data: likelihood approach revisited. *Biometrics.* 2006;62(4):1037-1043. doi:10.111/j.1541-0420.2006.00570.x.
3. Tsiatis AA, Davidian M. Joint modeling of longitudinal and time-to-event data: An overview. *Stat Sin.* 2004;14:809-834.
4. Rizopoulos D. Dynamic predictions and prospective accuracy in joint models for longitudinal and time-to-event data. *Biometrics.* 2011;67(3):819-829. doi:10.1111/j.1541-0420.2010.01546.x.
5. Rizopoulos D. *Joint models for longitudinal and time-to-event data: with applications in R.* 1 ed. New York: CRC Press; 2012.
6. Henderson R, Diggle P, Dobson A. Joint modelling of longitudinal measurements and event time data. *Biostatistics.* 2000;1(4):465-480. doi:10.1093/biostatistics/1.4.465.
7. Brown ER, Ibrahim JG. Bayesian approaches to joint cure-rate and longitudinal models with applications to cancer vaccine trials. *Biometrics.* 2003;59(3):686-693. doi:10.1111/1541-0420.00079.
8. Chi YY, Ibrahim JG. Joint models for multivariate longitudinal and multivariate survival data. *Biometrics.* 2006;62(2):432-445. doi:10.1111/j.1541-0420.2005.00448.x.
9. van Houwelingen H. Dynamic prediction by landmarking in event history analysis. *Scand J Stat.* 2007;34(1):70-85. doi:10.1111/j.1467-9469.2006.00529.x.
10. van Houwelingen H, Putter H. *Dynamic prediction in clinical survival analysis.* CRC Press; 2011.
11. Rizopoulos D, Molenberghs G, Lesaffre E. Dynamic predictions with time-dependent covariates in survival analysis using joint modeling and landmarking. *Biom J.* 2017;59(6):1261-1276. doi:10.1002/bimj.201600238.
12. Zheng Y, Heagerty PJ. Partly conditional survival models for longitudinal data. *Biometrics.* 2005;61(2):379-391. doi:10.1111/j.1541-0420.2005.00323.x.
13. Blanche P, Proust-Lima C, Loubere L, Berr C, Dartigues JF, Jacqmin-Gadda H. Quantifying and comparing dynamic predictive accuracy of joint models for longitudinal marker and time-to-event in presence of censoring and competing risks. *Biometrics.* 2015;71(1):102-113. doi:10.1111/biom.12232.