

What's cooking in the dynamic prediction pot

Mirko Signorelli

 mirkosignorelli.github.io

 [in/signorelli](#)

Mathematical Institute
Leiden University

September 18, 2025

Dipartimento di Scienze Statistiche "Paolo Fortunati", UniBo



Universiteit
Leiden

Dynamic prediction

Strict vs relaxed data landmarking

Penalized Regression Calibration

The R package `pencal`

Benchmarking

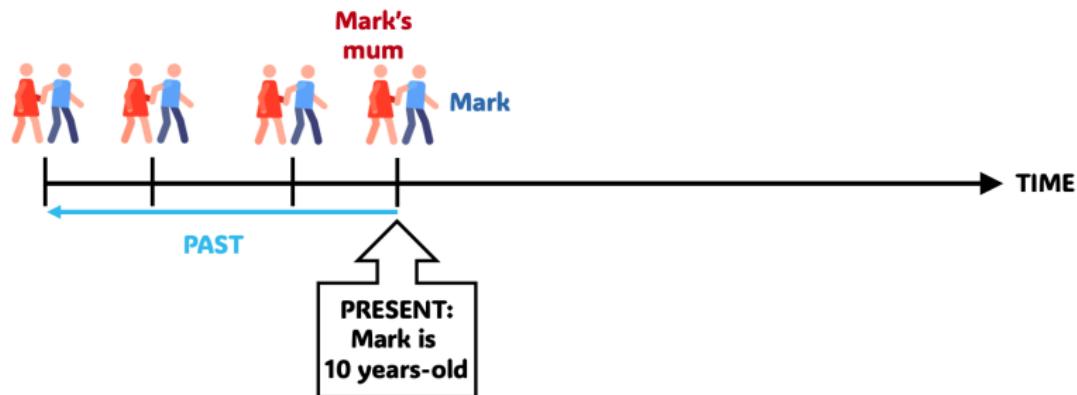
What's cooking in the dynamic prediction pot

References

Appendix

Dynamic prediction 101

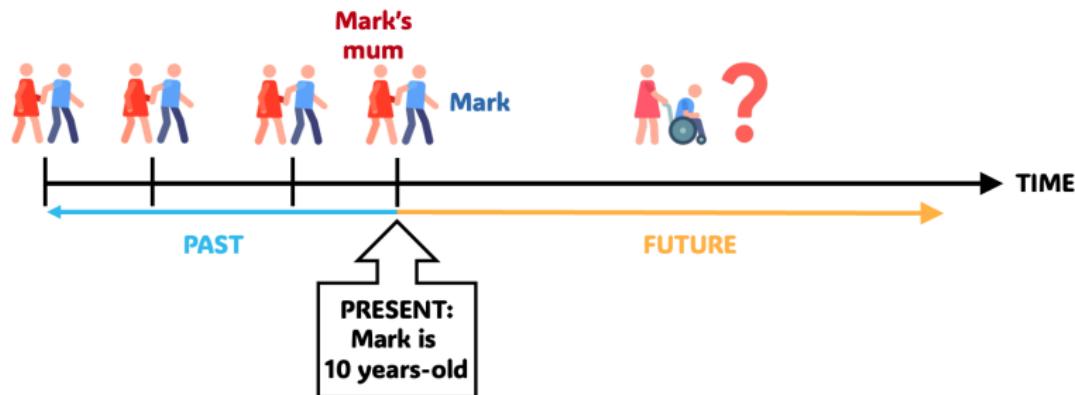
- ▶ Situation at time t :



- ▶ Mark suffers from Duchenne muscular dystrophy (DMD)
- ▶ Usually: **loss of ambulation** during adolescence

Dynamic prediction 101

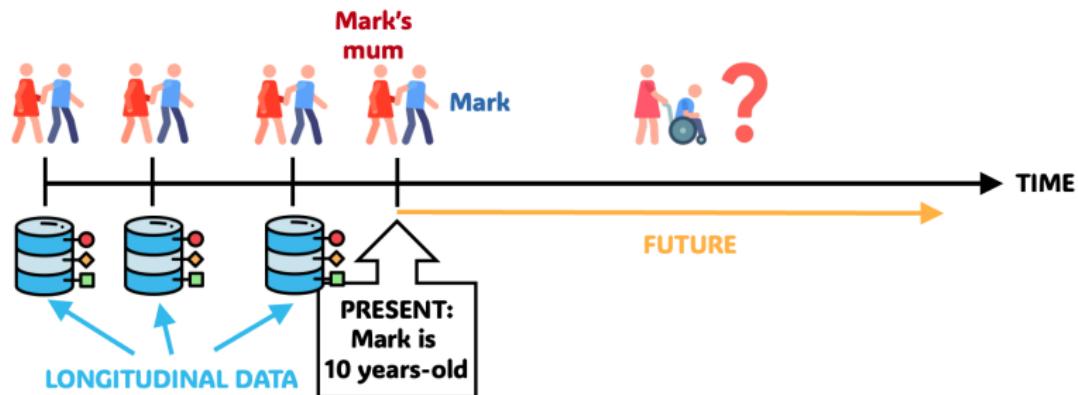
- ▶ Situation at time t :



- ▶ Mum: $P(\text{wheelchair within 2 years})?$

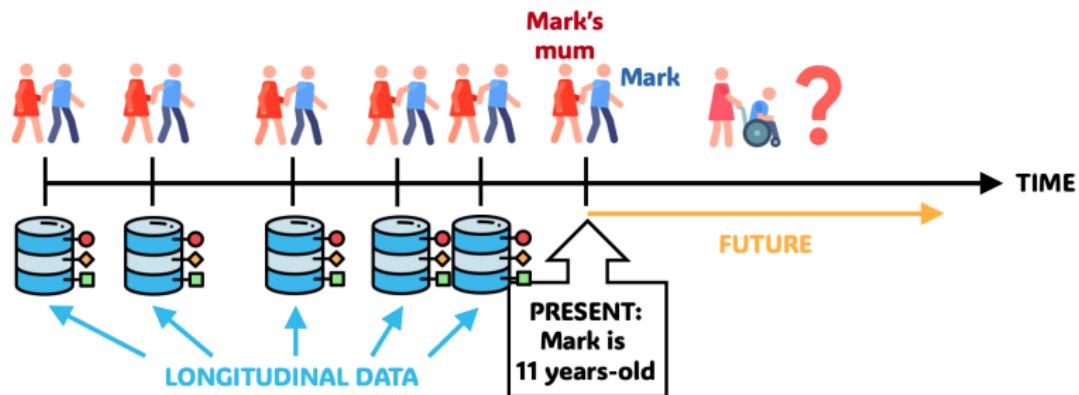
Dynamic prediction 101

- ▶ Situation at time t :



Dynamic prediction 101

- One year later ($t + 1$):



Goals and examples

► Goals:

1. predict future survival $S(t|\ell_1) = P(T > t | T > \ell_1)$, $t > \ell_1$, using repeated measurements over $[0, \ell_1]$
2. update predictions when newer information becomes available,
i.e. update $S(t|\ell_2)$ given repeated measurements over $[0, \ell_2]$,
 $t > \ell_2 > \ell_1$

Goals and examples

- ▶ Goals:
 1. predict future survival $S(t|\ell_1) = P(T > t | T > \ell_1)$, $t > \ell_1$, using repeated measurements over $[0, \ell_1]$
 2. update predictions when newer information becomes available, i.e. update $S(t|\ell_2)$ given repeated measurements over $[0, \ell_2]$, $t > \ell_2 > \ell_1$
- ▶ Example 1: ADNI & ROSMAP studies
 - ▶ outcome: time to dementia / AD diagnosis
 - ▶ hundreds of longitudinal predictors available, $n \in [1500, 5000]$
- ▶ Example 2: Mark-MD study
 - ▶ outcome: time to loss of ambulation
 - ▶ 240 longitudinal predictors (proteins), $n = 93 \Rightarrow$ high-dimensional!

Traditional approaches (1/2): LOCF landmarking

- ▶ Last observation carried forward (LOCF) landmarking (Van Houwelingen, 2007):
 1. choose a **landmark time** ℓ
 2. build a **model** for $S(t|\ell) = P(T > t | T > \ell)$ using **last observation** taken up until ℓ for each (longitudinal) covariate. Typically a Cox model where

$$h(t|x_i) = h_0(t)e^{x_i^T \beta}$$

Traditional approaches (1/2): LOCF landmarking

- ▶ Last observation carried forward (LOCF) landmarking (Van Houwelingen, 2007):
 1. choose a **landmark time** ℓ
 2. build a **model** for $S(t|\ell) = P(T > t | T > \ell)$ using **last observation** taken up until ℓ for each (longitudinal) covariate. Typically a Cox model where
$$h(t|x_i) = h_0(t)e^{x_i^T \beta}$$
- ▶ PROs: easy to implement & explain to clinicians
- ▶ CONs: doesn't truly use longitudinal data (only last observation) + last observation up to ℓ may be rather outdated

Traditional approaches (2/2): joint modelling

- ▶ Joint modelling (Tsiatis & Davidian, 2004):
 - ▶ specify a **shared random effects model** for longitudinal covariates and survival outcome
 - ▶ informally:

$$\begin{cases} Y_i(t) = f_1(\textcolor{blue}{u}_i, \dots) \\ h_i(t) = f_2(\textcolor{blue}{u}_i, \dots) \end{cases}$$

Traditional approaches (2/2): joint modelling

- ▶ Joint modelling (Tsiatis & Davidian, 2004):
 - ▶ specify a **shared random effects model** for longitudinal covariates and survival outcome
 - ▶ informally:
- ▶ PROs: uses longitudinal data & mathematically elegant
- ▶ CONS:
 1. joint estimation is **computationally intensive** (hours/days) \Rightarrow only a handful (3-10) of longitudinal covariates
 2. frequent **estimation / convergence errors** \Rightarrow problem for model estimation and validation of predictive performance (e.g., via cross-validation)

\Rightarrow what to do when you have **many longitudinal predictors?**

Some recent developments

- ▶ Many recent methodological proposals (Devaux et al., 2023; Gomon et al., 2024; Li & Luo, 2019; Lin et al., 2021; Signorelli et al., 2021; Signorelli, 2024)

Some recent developments

- ▶ Many recent methodological proposals (Devaux et al., 2023; Gomon et al., 2024; Li & Luo, 2019; Lin et al., 2021; Signorelli et al., 2021; Signorelli, 2024)
- ▶ Two-step modelling (or “model-based landmarking”):
 1. model trajectories of longitudinal predictors
 2. use summaries of longitudinal predictors to predict $S(t)$ or $S(t|\ell)$, $t \geq \ell$
- ▶ Essentially a hybrid strategy between LOCF landmarking, and joint modelling → attempt to combine the best of both worlds

Methods overview

- Model-based landmarking methods in a nutshell:

		Longitudinal covariates	
		Multivariate Functional PCA	Mixed Effects Models
Survival outcome	Cox PH model	MFPCCox (Li and Luo, 2019, Gomon et al., 2024)	Penalized Regression Calibration (Signorelli et al., 2021, Signorelli 2024)
	Random survival forest	Functional Random Survival Forest (Lin et al., 2021)	DynForest (Devaux et al., 2023)

Methods overview

- Model-based landmarking methods in a nutshell:

		Longitudinal covariates	
		Multivariate Functional PCA	Mixed Effects Models
Survival outcome	Cox PH model	MFPCCox (Li and Luo, 2019, Gomon et al., 2024)	Penalized Regression Calibration (Signorelli et al., 2021, Signorelli 2024)
	Random survival forest	Functional Random Survival Forest (Lin et al., 2021)	DynForest (Devaux et al., 2023)

- Inconsistent: **should you use longitudinal data measured at $t > \ell$ to train a model that predicts $S(t|\ell)$?**

Dynamic prediction

Strict vs relaxed data landmarking

Penalized Regression Calibration

The R package `pencal`

Benchmarking

What's cooking in the dynamic prediction pot

References

Appendix

Which longitudinal data to use for model training?

- ▶ LOCF landmarking $\rightarrow S(t|\ell) \rightarrow$ cut longitudinal data off at ℓ
- ▶ Joint modelling $\rightarrow S(t) \rightarrow$ use all longitudinal data (no landmark)

Which longitudinal data to use for model training?

- ▶ LOCF landmarking $\rightarrow S(t|\ell) \rightarrow$ cut longitudinal data off at ℓ
- ▶ Joint modelling $\rightarrow S(t) \rightarrow$ use all longitudinal data (no landmark)
- ▶ Two-step methods: **inconsistent & unclear!** \rightarrow we looked into this in Gomon et al. (2024)

Original Research Article



Dynamic prediction of survival using multivariate functional principal component analysis: A strict landmarking approach

Statistical Methods in Medical Research
2024, Vol. 33(2) 256–272

© The Author(s) 2024



Article reuse guidelines:
sagepub.com/journals-permissions
DOI: [10.1177/09622802231224631](https://doi.org/10.1177/09622802231224631)
journals.sagepub.com/home/smm



Daniel Gomon¹ , Hein Putter² , Marta Fiocco^{1,2}
and Mirko Signorelli¹

Strict vs relaxed data landmarking

- What we compared in Gomon et al. (2024):

262

Statistical Methods in Medical Research 33(2)

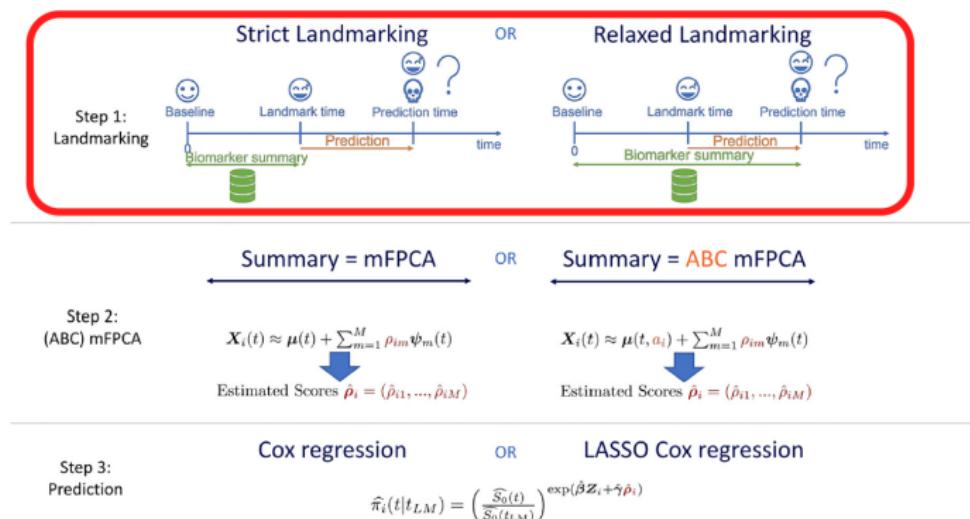
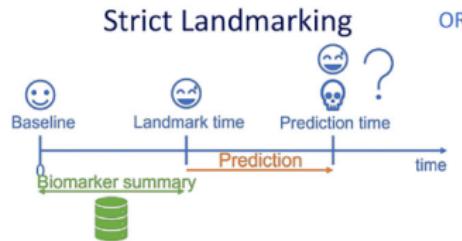


Figure 1. Graphical summary of the methods proposed in Section 2. See also Section 2.6.

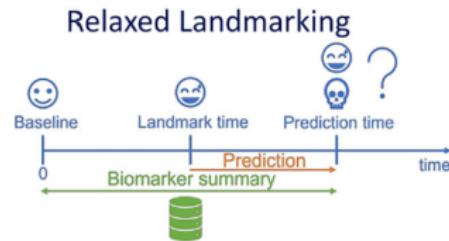
- Original MFPC Cox paper (Li & Luo, 2019) used relaxed data landmarking

Opinion poll

► Pick your fighter:

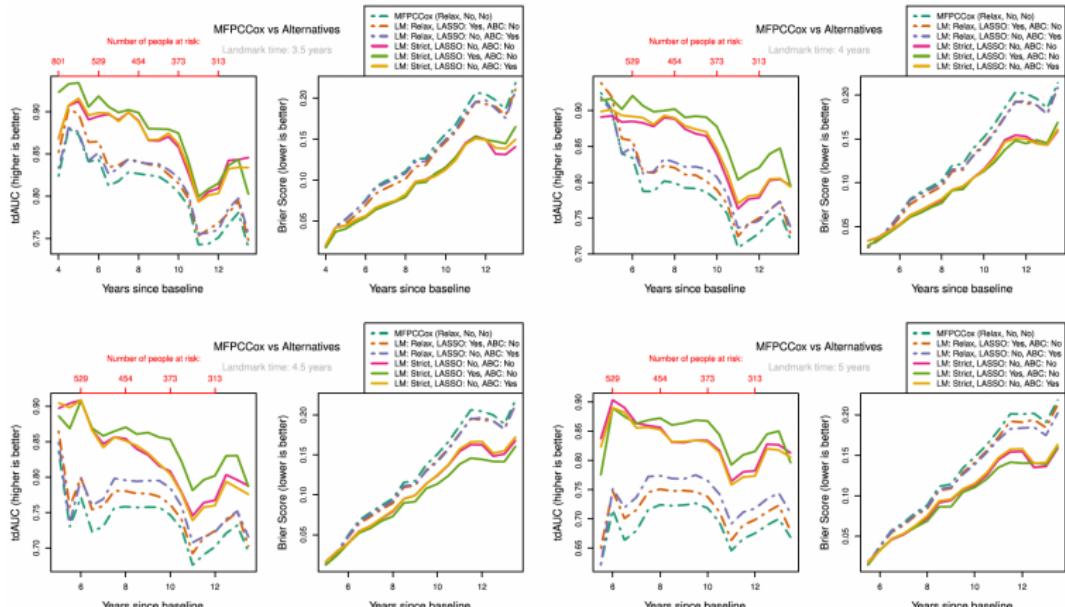


OR



Results on the ADNI dataset (Gomon et al., 2024)

- $n = 1625$; 5 baseline + 21 longitudinal covariates; 4 landmark times
- tdAUC: higher = better; Brier: lower = better
- Solid lines = strict data landmarking; dashed = relaxed data landmarking



Explanation

- ▶ Relaxed data landmarking often used in combination with two-step methods
 1. (naïve) idea: more data \Rightarrow better estimates and predictions
 2. problem 1: the earlier the event, the shorter the follow-up \Rightarrow selection bias that gets stronger over time ☺
 3. problem 2: using future longitudinal data to predict the future ☺

Explanation

- ▶ Relaxed data landmarking often used in combination with two-step methods
 1. (naïve) idea: more data \Rightarrow better estimates and predictions
 2. problem 1: the earlier the event, the shorter the follow-up \Rightarrow selection bias that gets stronger over time ☺
 3. problem 2: using future longitudinal data to predict the future ☺
- ▶ Strict data landmarking avoids this selection bias by conditioning on common landmark time
 1. specify conditional model for $S(t|\ell)$, $t > \ell$ using as predictors only data up until ℓ
 2. better predictions even if it uses less data (result consistent across methods, datasets, landmarks and performance measures!) ☺

Dynamic prediction

Strict vs relaxed data landmarking

Penalized Regression Calibration

The R package `pencal`

Benchmarking

What's cooking in the dynamic prediction pot

References

Appendix

Refresher

		Longitudinal covariates	
Survival outcome	Cox PH model	Multivariate Functional PCA	Mixed Effects Models
	Random survival forest	MFPCCox (Li and Luo, 2019, Gomon et al., 2024)	Penalized Regression Calibration (Signorelli et al., 2021, Signorelli 2024)
		Functional Random Survival Forest (Lin et al., 2021)	DynForest (Devaux et al., 2023)

Refresher

Survival outcome	Cox PH model	Longitudinal covariates	
		Multivariate Functional PCA	Mixed Effects Models
		MFPCCox (Li and Luo, 2019, Gomon et al., 2024)	Penalized Regression Calibration (Signorelli et al., 2021, Signorelli 2024)
	Random survival forest	Functional Random Survival Forest (Lin et al., 2021)	DynForest (Devaux et al., 2023)

Refresher

		Longitudinal covariates	
Survival outcome	Cox PH model	Multivariate Functional PCA	Mixed Effects Models
	Random survival forest	Functional Random Survival Forest (Lin et al., 2021)	Penalized Regression Calibration (Signorelli et al., 2021, Signorelli 2024)

Penalized Regression Calibration

Penalized Regression Calibration (Signorelli et al., 2021; Signorelli, 2024):

Received: 25 January 2021 | Revised: 10 August 2021 | Accepted: 10 August 2021
DOI: 10.1002/stm.9178

RESEARCH ARTICLE

Statistics
in Medicine WILEY

Penalized regression calibration: A method for the prediction of survival outcomes using complex longitudinal and high-dimensional data

Mirko Signorelli¹ | Pietro Spitali² | Cristina Al-Khalili Szgyarto³ | The MARK-MD Consortium | Roula Tsonaka⁴

¹Mathematical Institute, Leiden University, Leiden, The Netherlands
²Department of Human Genetics, Leiden University Medical Center, Leiden, The Netherlands

Longitudinal and high-dimensional measurements have become increasingly common in biomedical research. However, methods to predict survival outcomes using covariates that are both longitudinal and high-dimensional are

CONTRIBUTED RESEARCH ARTICLE



The R Journal

pencal: an R Package for the Dynamic Prediction of Survival with Many Longitudinal Predictors

by Mirko Signorelli

Abstract In survival analysis, longitudinal information on the health status of a patient can be used to dynamically update the predicted probability that a patient will experience an event of interest. Traditional approaches to dynamic prediction such as joint models become computationally unfeasible with more than a handful of longitudinal covariates, warranting the development of methods that can handle a larger number of longitudinal covariates. We introduce the R package `pencal`, which implements the Penalized Regression Calibration (PRC) approach to estimate survival curves using longitudinal covariates as predictors of survival. `pencal` uses mixed-effects models to summarize the trajectories of the longitudinal covariates up to a prespecified landmark time, and a penalized Cox model to predict survival based on both baseline covariates and summary measures of the longitudinal covariates. This article illustrates the structure of the R package, provides a step by step example showing how to estimate PRC, compute dynamic predictions of survival and validate performance, and shows how parallelization can be used to significantly reduce computing time.

Penalized Regression Calibration

Penalized Regression Calibration (Signorelli et al., 2021; Signorelli, 2024):

1. build a **conditional** model for $S(t|\ell)$ (\Rightarrow use strict data landmarking!)

Received: 25 January 2021 | Revised: 10 August 2021 | Accepted: 10 August 2021
DOI: 10.1002/stm.9178

RESEARCH ARTICLE

Statistics
in Medicine WILEY

Penalized regression calibration: A method for the prediction of survival outcomes using complex longitudinal and high-dimensional data

Mirko Signorelli¹ | Pietro Spitali² | Cristina Al-Khalili Szigyarto³ | The MARK-MD Consortium | Roula Tsounaka⁴

¹Mathematical Institute, Leiden University, Leiden, The Netherlands
²Department of Human Genetics, Leiden University Medical Center, Leiden, The Netherlands

Longitudinal and high-dimensional measurements have become increasingly common in biomedical research. However, methods to predict survival outcomes using covariates that are both longitudinal and high-dimensional are

CONTRIBUTED RESEARCH ARTICLE



The R Journal

pencal: an R Package for the Dynamic Prediction of Survival with Many Longitudinal Predictors

by Mirko Signorelli

Abstract In survival analysis, longitudinal information on the health status of a patient can be used to dynamically update the predicted probability that a patient will experience an event of interest. Traditional approaches to dynamic prediction such as joint models become computationally unfeasible with more than a handful of longitudinal covariates, warranting the development of methods that can handle a larger number of longitudinal covariates. We introduce the R package `pencal`, which implements the `Penalized Regression Calibration` (PRC) method to estimate survival curves using longitudinal covariates as predictors of survival. `pencal` uses mixed-effects models to summarize the trajectories of the longitudinal covariates up to a prespecified landmark time, and a penalized Cox model to predict survival based on both baseline covariates and summary measures of the longitudinal covariates. This article illustrates the structure of the R package, provides a step by step example showing how to estimate PRC, compute dynamic predictions of survival and validate performance, and shows how parallelization can be used to significantly reduce computing time.

Penalized Regression Calibration

Penalized Regression Calibration (Signorelli et al., 2021; Signorelli, 2024):

1. build a **conditional** model for $S(t|\ell)$ (\Rightarrow use strict data landmarking!)
2. model longitudinal predictors over $[0, \ell]$ using **mixed-effects models**

Received: 25 January 2021 | Revised: 10 August 2021 | Accepted: 10 August 2021
DOI: 10.1002/stm.9178

RESEARCH ARTICLE

Statistics
in Medicine WILEY

Penalized regression calibration: A method for the prediction of survival outcomes using complex longitudinal and high-dimensional data

Mirko Signorelli¹ | Pietro Spitali² | Cristina Al-Khalili Szgyarto³ | The MARK-MD Consortium | Roula Tsounaka⁴

¹Mathematical Institute, Leiden University, Leiden, The Netherlands
²Department of Human Genetics, Leiden University Medical Center, Leiden, The Netherlands

Longitudinal and high-dimensional measurements have become increasingly common in biomedical research. However, methods to predict survival outcomes using covariates that are both longitudinal and high-dimensional are

CONTRIBUTED RESEARCH ARTICLE



The R Journal

pencal: an R Package for the Dynamic Prediction of Survival with Many Longitudinal Predictors

by Mirko Signorelli

Abstract In survival analysis, longitudinal information on the health status of a patient can be used to dynamically update the predicted probability that a patient will experience an event of interest. Traditional approaches to dynamic prediction such as joint models become computationally unfeasible with more than a handful of longitudinal covariates, warranting the development of methods that can handle a larger number of longitudinal covariates. We introduce the R package `pencal`, which implements the Penalized Regression Calibration (PRC) method to estimate survival probabilities using longitudinal covariates as predictors of survival. `pencal` uses mixed-effects models to summarize the trajectories of the longitudinal covariates up to a prespecified landmark time, and a penalized Cox model to predict survival based on both baseline covariates and summary measures of the longitudinal covariates. This article illustrates the structure of the R package, provides a step by step example showing how to estimate PRC, compute dynamic predictions of survival and validate performance, and shows how parallelization can be used to significantly reduce computing time.

Penalized Regression Calibration

Penalized Regression Calibration (Signorelli et al., 2021; Signorelli, 2024):

1. build a **conditional** model for $S(t|\ell)$ (\Rightarrow use strict data landmarking!)
2. model longitudinal predictors over $[0, \ell]$ using **mixed-effects models**
3. use summaries from mixed models to predict $S(t|\ell)$

Received: 25 January 2021 | Revised: 10 August 2021 | Accepted: 10 August 2021
DOI: 10.1002/stm.9178

RESEARCH ARTICLE

Statistics
in Medicine WILEY

Penalized regression calibration: A method for the prediction of survival outcomes using complex longitudinal and high-dimensional data

Mirko Signorelli¹ | Pietro Spitali² | Cristina Al-Khalili Szgyarto³ | The MARK-MD Consortium | Roula Tsounaka⁴

¹Mathematical Institute, Leiden University, Leiden, The Netherlands
²Department of Human Genetics, Leiden University Medical Center, Leiden, The Netherlands

Longitudinal and high-dimensional measurements have become increasingly common in biomedical research. However, methods to predict survival outcomes using covariates that are both longitudinal and high-dimensional are

CONTRIBUTED RESEARCH ARTICLE



The R Journal

pencal: an R Package for the Dynamic Prediction of Survival with Many Longitudinal Predictors

by Mirko Signorelli

Abstract In survival analysis, longitudinal information on the health status of a patient can be used to dynamically update the predicted probability that a patient will experience an event of interest. Traditional approaches to dynamic prediction such as joint models become computationally unfeasible with more than a handful of longitudinal covariates, warranting the development of methods that can handle a larger number of longitudinal covariates. We introduce the R package `pencal`, which implements the Penalized Regression Calibration (PRC) approach to estimate survival curves using longitudinal covariates as predictors of survival. `pencal` uses mixed-effects models to summarize the trajectories of the longitudinal covariates up to a prespecified landmark time, and a penalized Cox model to predict survival based on both baseline covariates and summary measures of the longitudinal covariates. This article illustrates the structure of the R package, provides a step by step example showing how to estimate PRC, compute dynamic predictions of survival and validate performance, and shows how parallelization can be used to significantly reduce computing time.

Notation

- ▶ $i = 1, \dots, n$ subjects, m_i repeated measurements per subject
 - ▶ k **baseline** covariates: $x_i = (x_{1i}, \dots, x_{ki})^T$
 - ▶ p **longitudinal** covariates $y_{ij} = (y_{1ij}, \dots, y_{pij})^T$, measured on subject i at times t_{i1}, \dots, t_{i,m_i} ($j = 1, \dots, m_i$)
 - ▶ (t_i, δ_i) **survival** outcome

Notation

- ▶ $i = 1, \dots, n$ subjects, m_i repeated measurements per subject
 - ▶ k **baseline** covariates: $x_i = (x_{1i}, \dots, x_{ki})^T$
 - ▶ p **longitudinal** covariates $y_{ij} = (y_{1ij}, \dots, y_{pij})^T$, measured on subject i at times t_{i1}, \dots, t_{i,m_i} ($j = 1, \dots, m_i$)
 - ▶ (t_i, δ_i) **survival** outcome
- ▶ **Landmark** ℓ
 - ▶ $\mathcal{I}(\ell)$ subjects still at risk at the landmark
 - ▶ $\mathcal{Y}_i(\ell)$ i 's longitudinal covariates up until the landmark

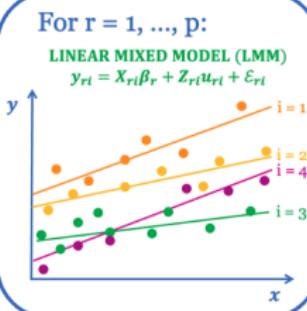
Notation

- ▶ $i = 1, \dots, n$ subjects, m_i repeated measurements per subject
 - ▶ k **baseline** covariates: $x_i = (x_{1i}, \dots, x_{ki})^T$
 - ▶ p **longitudinal** covariates $y_{ij} = (y_{1ij}, \dots, y_{pij})^T$, measured on subject i at times t_{i1}, \dots, t_{i,m_i} ($j = 1, \dots, m_i$)
 - ▶ (t_i, δ_i) **survival** outcome
- ▶ **Landmark** ℓ
 - ▶ $\mathcal{I}(\ell)$ subjects still at risk at the landmark
 - ▶ $\mathcal{Y}_i(\ell)$ i 's longitudinal covariates up until the landmark
- ▶ Goal: predict

$$S_i(t|\ell, x_i, \mathcal{Y}_i(\ell)) = P(T_i > t | T_i > \ell, x_i, \mathcal{Y}_i(\ell)), \quad i \in \mathcal{I}(\ell)$$

Modelling steps

Step 1



Step 2

Compute the predicted random effects

$$\hat{u}_{rt} = \hat{W}_{ri}(y_{ri} - X_{ri}\hat{\beta}_r)$$

where

$$\hat{W}_{ri} = \hat{D}_r Z_{ri} \hat{V}_{ri}^{-1}$$

Step 3

Estimate

$$h(t|\eta_i) = h_0(t)e^{\eta_i},$$

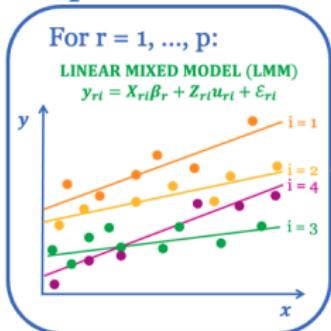
where

$$\eta_i = x_i^T \tau + \hat{u}_i^T \gamma,$$

using **penalized maximum likelihood** (ridge / lasso / elasticnet)

Step 1: model the longitudinal data

Step 1



Step 2

Compute the predicted random effects

$$\hat{u}_{ri} = \hat{W}_{ri}(y_{ri} - X_{ri}\hat{\beta}_r)$$

where

$$\hat{W}_{ri} = \hat{D}_r Z_{ri} \hat{V}_{ri}^{-1}$$

Step 3

Estimate

$$h(t|\eta_i) = h_0(t)e^{\eta_i},$$

where

$$\eta_i = x_i^T \tau + \hat{u}_i^T \gamma,$$

using **penalized maximum likelihood** (ridge / lasso / elasticnet)

Step 1: model the longitudinal data

The evolution of the longitudinal predictors over $[0, \ell]$ is modelled using

1. linear mixed models (LMM)

$$y_{si} = W_{si}\beta_s + Z_{si}u_{si} + \varepsilon_{si}, \quad i \in \mathcal{I}(\ell),$$

where $u_{si} \sim N(0, D_s)$ is a vector of random effects, $\varepsilon_{si} \sim N(0, \sigma_s^2 I_{m_i})$

2. linear multivariate latent process mixed models (MLPMM, Proust-Lima et al. (2013)) → see Signorelli et al. (2021) for specifics

Example

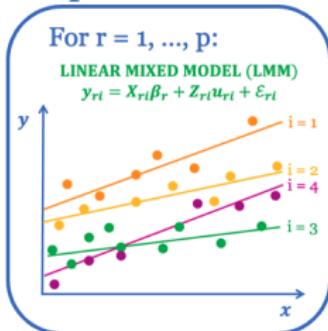
- ▶ LMM with random intercept and random slope for age:

$$y_{sij} = \beta_{s0} + u_{s0i} + (\beta_1 + u_{s1i})a_{ij} + \varepsilon_{sij},$$

$$u_{si} \sim N_2(0, D_s), \quad \varepsilon_{si} \sim N(0, \sigma_s^2 I_{m_i})$$

Step 2: summarize the longitudinal predictors

Step 1



Step 2

Compute the predicted random effects

$$\hat{u}_{ri} = \hat{W}_{ri}(y_{ri} - X_{ri}\hat{\beta}_r)$$

where

$$\hat{W}_{ri} = \hat{D}_r Z_{ri} \hat{V}_{ri}^{-1}$$

Step 3

Estimate

$$h(t|\eta_i) = h_0(t)e^{\eta_i},$$

where

$$\eta_i = x_i^T \tau + \hat{u}_i^T \gamma,$$

using penalized maximum likelihood (ridge / lasso / elasticnet)

Step 2: summarize the longitudinal predictors

- ▶ Compute summaries of longitudinal predictors
- ▶ This requires computation of predicted random effects

Step 2: summarize the longitudinal predictors

- ▶ Compute summaries of longitudinal predictors
- ▶ This requires computation of **predicted random effects**
- ▶ Can be done **analytically** (for LMM & linear MLPMM).

1. LMM:

$$\hat{u}_{si} = E(u_{si} | Y_{si} = y_{si}) = \hat{D}_s Z_{si}^T \hat{V}_{si}^{-1} (y_{si} - W_{si} \hat{\beta}_s),$$

where $\hat{V}_{si} = Z_{si} \hat{D}_s Z_{si}^T + \hat{\sigma}_s^2 I$

2. MLPMM:

For the MLPMM, we can compute the predicted random effects $\hat{u}_{si} = (\hat{u}_{s0i}, \hat{u}_{s1i})$ and $\hat{b}_{si} = (\hat{b}_{1si}, \dots, \hat{b}_{rsi})$ by adapting the formulas provided by Ebrahimpour et al¹⁵ to the MLPMM of Equation (1) as follows:

$$(\hat{u}_{si}, \hat{b}_{si}) = E(u_{si}, b_{si} | Y_{si} = y_{si}) = \begin{bmatrix} Z_i \hat{\Sigma}_{u_i} \\ \hat{\Sigma}_{b_i} I_r \otimes \mathbf{1}_{m_r, 1} \end{bmatrix} \hat{\Sigma}_{y_{si}}^{-1} \hat{y}_{si}, \quad (4)$$

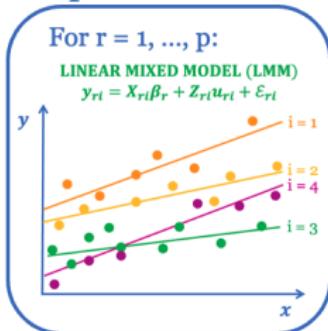
where $y_{si} = (y_{1si}, \dots, y_{1sim_i}, \dots, y_{rsi}, \dots, y_{rsm_i})^T$, \hat{y}_{si} is the equivalent of y_{si} with $y_{qij} = y_{qij} - \hat{\beta}_{q0i} - \hat{\beta}_{qi1} a_{ij}$ as entries,

Z_i is the random-effects design matrix associated to y_{si} in (1), $\Sigma_{b_i} = \begin{bmatrix} \sigma_{b1i}^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_{br_i i}^2 \end{bmatrix}$, $\Sigma_{u_i} = \begin{bmatrix} \sigma_{u1i}^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_{ur_i i}^2 \end{bmatrix}$, $\Sigma_{u_i} =$

$\begin{bmatrix} \sigma_{u0i}^2 & \sigma_{u0i, u1i} \\ \sigma_{u0i, u1i} & \sigma_{u1i}^2 \end{bmatrix}$, and $\Sigma_{y_{si}} = Z_i \Sigma_{u_i} Z_i^T + I_n \otimes \Sigma_{e_i} I_{m_i} + I_{r_i} \otimes \Sigma_{b_i} \mathbf{1}_{m_r, m_i}$, where I denotes identity matrices and $\mathbf{1}_{a,b}$ "all-ones" matrices (ie, matrices whose entries are all equal to 1) of dimension $a \times b$.

Step 3: predict the survival outcome

Step 1



Step 2

Compute the predicted random effects

$$\hat{u}_{ri} = \hat{W}_{ri}(y_{ri} - X_{ri}\hat{\beta}_r)$$

where

$$\hat{W}_{ri} = \hat{D}_r Z_{ri} \hat{V}_{ri}^{-1}$$

Step 3

Estimate

$$h(t|\eta_i) = h_0(t)e^{\eta_i},$$

where

$$\eta_i = x_i^T \tau + \hat{u}_i^T \gamma,$$

using **penalized maximum likelihood** (ridge / lasso / elasticnet)

Step 3: predict the survival outcome

- ▶ Cox model linking x_i and \hat{u}_i to survival outcome:

$$h(t|x_i, \hat{u}_i) = h_0(t_i) \exp(x_i^T \gamma + \hat{u}_i^T \delta)$$

- ▶ Estimated using penalized likelihood estimation to deal with large number of predictors (ridge / elasticnet)
- ▶ Computation of $\hat{S}(t|\ell, x_i, \hat{u}_i)$ proceeds as usual in Cox models

Dynamic prediction

Strict vs relaxed data landmarking

Penalized Regression Calibration

The R package `pencal`

Benchmarking

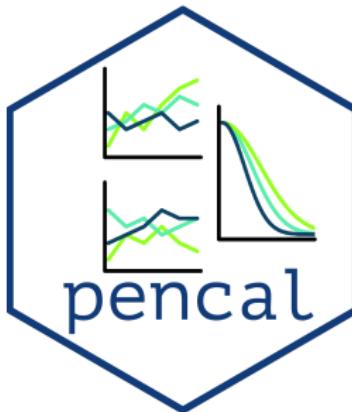
What's cooking in the dynamic prediction pot

References

Appendix

The R package: pencal

- ▶ Penalized Regression Calibration implemented in the R package `pencal`



- ▶ Available on  CRAN

Input data

- ▶ Example using the PBC2 dataset, with $\ell = 2$
- ▶ Longitudinal predictors: a data frame in long format

```
ldata[1:6, 1:7]
```

```
##   id      age  fuptime serBilir serChol albumin alkaline
## 1  2 56.44782 0.0000000    1.1     302    4.14    7395
## 2  2 56.94612 0.4983025    0.8      NA    3.60    2107
## 3  2 57.44716 0.9993429    1.0      NA    3.55    1711
## 4  3 70.07447 0.0000000    1.4     176    3.48     516
## 5  3 70.55635 0.4818749    1.1      NA    3.29     353
## 6  3 71.07108 0.9966050    1.5     233    3.57     218
```

Input data

- ▶ Example using the PBC2 dataset, with $\ell = 2$
- ▶ Longitudinal predictors: a data frame in long format

```
ldata[1:6, 1:7]
```

```
##   id      age  fuptime serBilir serChol albumin alkaline
## 1  2 56.44782 0.0000000     1.1     302    4.14    7395
## 2  2 56.94612 0.4983025     0.8      NA    3.60    2107
## 3  2 57.44716 0.9993429     1.0      NA    3.55    1711
## 4  3 70.07447 0.0000000     1.4     176    3.48     516
## 5  3 70.55635 0.4818749     1.1      NA    3.29     353
## 6  3 71.07108 0.9966050     1.5     233    3.57     218
```

- ▶ Baseline covariates + survival outcome: a data frame in wide format

```
head(sdata, 4)
```

```
##   id      time event baselineAge      sex treatment
## 1  2 14.152338     0    56.44782 female D-penicil
## 2  3  2.770781     1    70.07447  male D-penicil
## 3  4  5.270507     1    54.74209 female D-penicil
## 4  5  4.120578     0    38.10645 female placebo
```

Step 1: fit_lmms

```
long_covs = c('logSerBil', 'logSerChol', 'albumin',
             'logAlk', 'logSGOT', 'platelets',
             'logProthr')

step1 = fit_lmms(y.names = long_covs,
                  fixefs = ~ age, ranefs = ~ age | id,
                  long.data = ldata, surv.data = sdata,
                  t.from.base = fuptime)
```

- `ranefs` argument follows the notation used by the `nlme` package (random intercept: `~ 1 | id`, random slope: `~ x | id, ...`)

Extracting output from the fitted LMMs

```
summary(step1, yname = 'logSerBil', what = 'betas') |> round(6)
```

```
## (Intercept)      age  
##   0.518320  -0.001045
```

```
summary(step1, yname = 'logSerBil', what = 'tTable') |> round(4)
```

	Value	Std.Error	DF	t-value	p-value
## (Intercept)	0.5183	0.2788	566	1.8590	0.0636
## age	-0.0010	0.0055	566	-0.1884	0.8506

```
summary(step1, yname = 'logSerBil', what = 'variances')
```

	Variance	StdDev	Corr
## (Intercept)	7.332118e-01	0.856277849	(Intr)
## age	4.731627e-05	0.006878682	0.103
## Residual	1.437622e-01	0.379159888	

Step 2: `summarize_lmms`

```
step2 = summarize_lmms(step1)
```

Step 2: sample output

```
summary(step2)
```

```
## Number of predicted random effect variables: 14  
## Sample size: 278
```

```
round(step2$ranef.orig[1:4, 1:6], 6)
```

```
##   logSerBil_b_int logSerBil_b_age logSerChol_b_int  
## 2      -0.382988     -0.001661     -0.071154  
## 3     -0.117107     -0.000584     -0.598453  
## 4      0.168600      0.000922     -0.370434  
## 5      0.380035      0.001170     -0.291031  
##   logSerChol_b_age albumin_b_int albumin_b_age  
## 2      0.000660      0.179725      3e-06  
## 3      0.004916      0.018124      1e-06  
## 4      0.003468     -0.529776     -7e-06  
## 5      0.002886     -0.148329      8e-06
```

Step 3: fit_prclmm

```
step3 = fit_prclmm(step2, surv.data = sdata,
                    baseline.covs = ~ baselineAge + sex + treatment,
                    penalty = 'ridge', standardize = T)
```

Step 3: fitted model

```
summary(step3)
```

```
## Fitted model: PRC-LMM
## Penalty function used: ridge
## Tuning parameters:
##      lambda alpha
## 1 0.2126761      0
## Sample size: 278
## Number of events: 107
## Bootstrap optimism correction: not computed
## Penalized likelihood estimates (rounded to 4 digits):
##   baselineAge sexfemale treatmentD-penicil logSerBil_b_int
## 1      0.0476    -0.2872        -0.0157      0.4341
##   logSerBil_b_age logSerChol_b_int logSerChol_b_age
## 1      111.3935      0.0986     -10.5311
##   albumin_b_int albumin_b_age logAlk_b_int logAlk_b_age
## 1     -1.1361    23070.92      0.0874     -12.5617
##   logSGOT_b_int logSGOT_b_age platelets_b_int
## 1      0.238       272.246     -0.0011
##   platelets_b_age logProthr_b_int logProthr_b_age
## 1     -0.2046      2.8114     -573.3093
```

Prediction of survival probabilities

```
Shat = survpred_prclmm(step1, step2, step3, times = 3:5)
```

- ▶ This will compute $\hat{S}(t|2)$, $t = 3, 4, 5$:

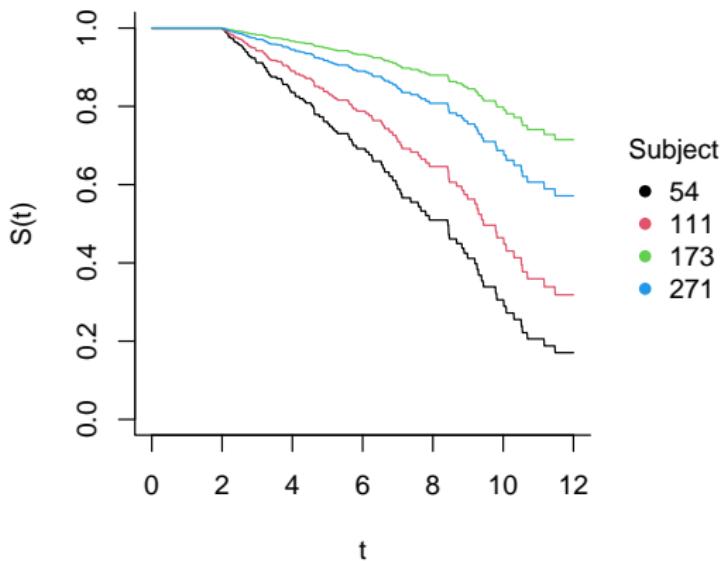
```
head(Shat$predicted_survival, 4) |> dfrround(3)
```

```
##   id S(3) S(4) S(5)
## 2  2 0.940 0.887 0.833
## 3  3 0.856 0.739 0.632
## 4  4 0.814 0.671 0.545
## 5  5 0.946 0.898 0.849
```

- ▶ Prediction for **new** subjects? Possible through additional arguments `new.longdata` and `new.basecovs`

Visualizing predictions

```
survplot_prc(step1, step2, step3,  
             ids = c(54, 111, 173, 271), tmax = 12)
```



Dynamic prediction

Strict vs relaxed data landmarking

Penalized Regression Calibration

The R package `pencal`

Benchmarking

What's cooking in the dynamic prediction pot

References

Appendix

Data

- We considered data from three longitudinal studies:

ROSMAP



- Event: Alzheimer's Disease diagnosis
- n = 3293
- 5 baseline covariates
- 30 longitudinal covariates
- Follow-up: [1, 29] years

ADNI



- Event: diagnosis of dementia
- n = 1643
- 5 baseline covariates
- 21 longitudinal covariates
- Follow-up: [0, 15.5] years

PBC2



- Event: death (primary biliary cirrhosis trial)
- n = 312
- 3 baseline covariates
- 8 longitudinal covariates
- Follow-up: [0, 14] years

- Performance evaluated at multiple landmark times
- Performance measures: tdAUC & Brier score
- 10-fold cross-validation, repeated 10 times

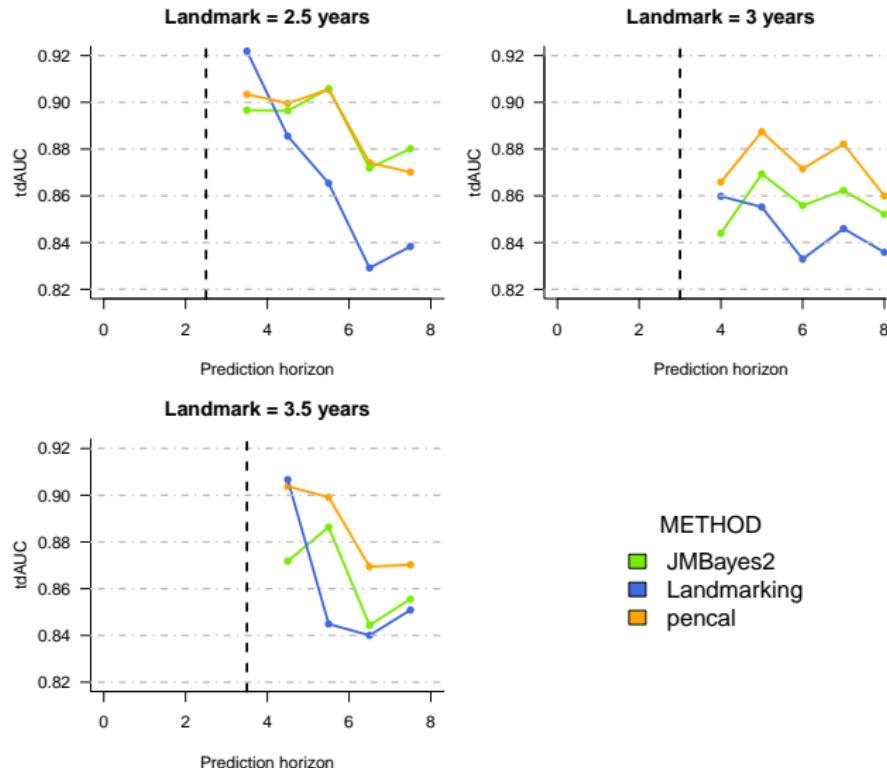
Methods

- ▶ Methods included in comparison:

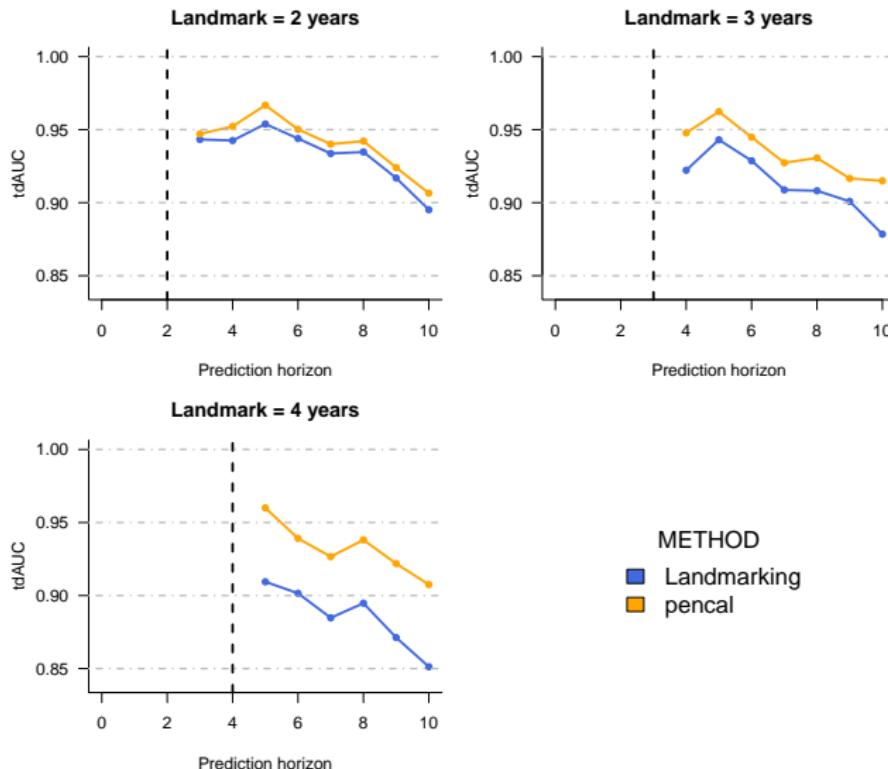
Method	PBC2	ADNI	ROSMAP
Landmarking	✓	✓	✓
Joint model	✓	✗	✗
PRC	✓	✓	✓

- ▶ Landmarking → `survival::coxph()`
- ▶ Joint modelling → `JMbayes2`
 - ▶ Frequent convergence problems
 - ▶ Couldn't fit model (with repeated-CV) to larger datasets
- ▶ PRC → `pencal`
 - ▶ Log-transformed skewed longitudinal variables for analysis with LMM

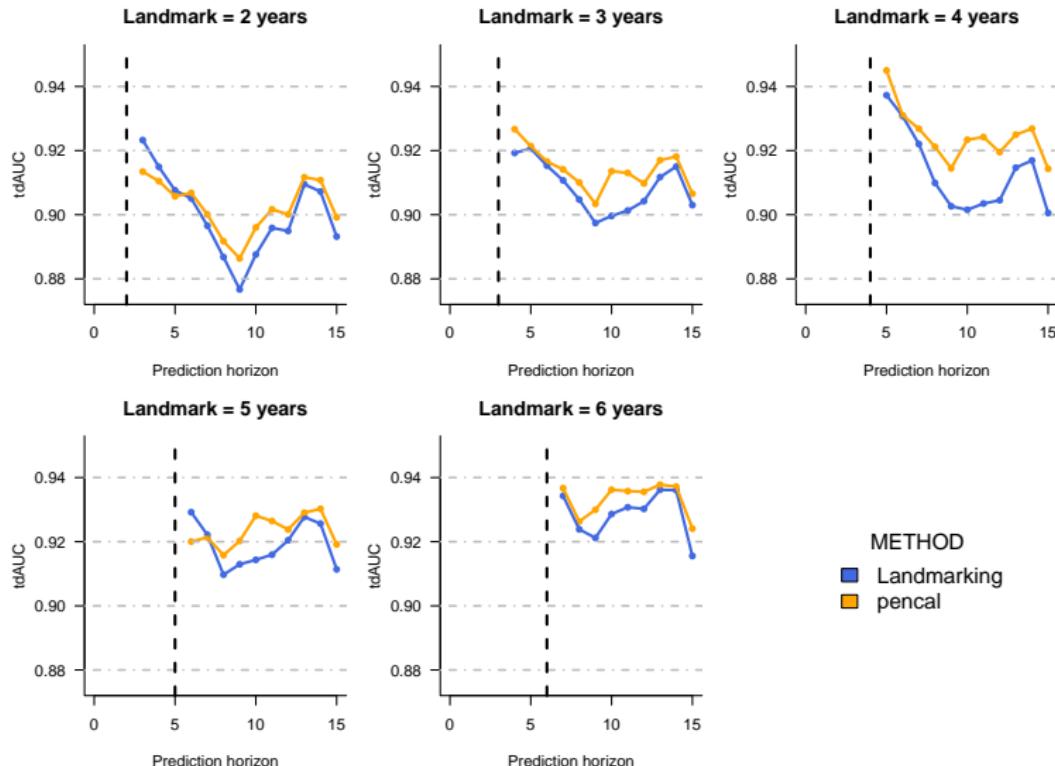
PBC2 dataset ($n = 312, p = 8$)



ADNI dataset ($n = 1643, p = 21$)



ROSMAP dataset ($n = 3293, p = 30$)



Bonus content

- Not presented here: a comparison of the 4 multistep methods

Survival outcome	Cox PH model	Longitudinal covariates	
		Multivariate Functional PCA	Mixed Effects Models
		MFPCCox (Li and Luo, 2019, Gomon et al., 2024)	Penalized Regression Calibration (Signorelli et al., 2021, Signorelli 2024)
	Random survival forest	Functional Random Survival Forest (Lin et al., 2021)	DynForest (Devaux et al., 2023)

Bonus content

- ▶ Not presented here: a comparison of the 4 multistep methods
- ▶ Motivation: all recently proposed (2019-2024), very little experience on how they perform on real-world data
- ▶ Preprint (Signorelli & Retif, 2025+):

The screenshot shows a red-themed arXiv preprint page. At the top, the arXiv logo is followed by the path > stat > arXiv:2403.14336. To the right is a search bar with placeholder 'Search...', a dropdown menu for 'All fields', and a 'Search' button. Below the header, the title 'Statistics > Methodology' is displayed, along with the submission and revision dates: [Submitted on 21 Mar 2024 (v1), last revised 17 Apr 2025 (this version, v2)]. The main title of the paper is 'Benchmarking multi-step methods for the dynamic prediction of survival with numerous longitudinal predictors'. The authors listed are Mirko Signorelli and Sophie Retif. The abstract discusses the development of multi-step methods for dynamic survival prediction using longitudinal covariates, mentioning mixed-effects models, multivariate functional principal component analysis, Cox models, and random survival forests. It notes that these methods are still quite new and lack knowledge about their applicability, limitations, and predictive performance. The paper has been viewed 1,000 times. On the right side of the page, there is a sidebar titled 'Access Paper:' with links to 'View PDF', 'HTML (experimental)', 'TeX Source', and 'Other Formats', along with a 'view license' link. Below this is 'Current browse context: stat.ME' with links to 'prev' and 'next', and 'new | recent | 2024-03'. There is also a 'Change to browse by:' section with links to 'stat' and 'stat.AP'. Further down are sections for 'References & Citations' (with links to NASA ADS, Google Scholar, and Semantic Scholar) and 'Export BibTeX Citation'. A 'Bookmark' button is at the bottom.

arXiv > stat > arXiv:2403.14336

Search... All fields Help | Advanced Search

Statistics > Methodology

[Submitted on 21 Mar 2024 (v1), last revised 17 Apr 2025 (this version, v2)]

Benchmarking multi-step methods for the dynamic prediction of survival with numerous longitudinal predictors

Mirko Signorelli, Sophie Retif

In recent years, the growing availability of biomedical datasets featuring numerous longitudinal covariates has motivated the development of several multi-step methods for the dynamic prediction of time-to-event ("survival") outcomes. These methods employ either mixed-effects models or multivariate functional principal component analysis to model and summarize the longitudinal covariates' evolution over time. Then, they use Cox models or random survival forests to predict survival probabilities, using as covariates both baseline variables and the summaries of the longitudinal variables obtained in the previous modelling step. Because these multi-step methods are still quite new, to date little is known about their applicability, limitations, and predictive performance when applied to real-world data. To gain a better understanding of these aspects, we performed a

Access Paper:

[View PDF](#)
[HTML \(experimental\)](#)
[TeX Source](#)
[Other Formats](#)
[view license](#)

Current browse context:
stat.ME
< prev | next >
new | recent | 2024-03

Change to browse by:
stat
stat.AP

References & Citations
[NASA ADS](#)
[Google Scholar](#)
[Semantic Scholar](#)

[Export BibTeX Citation](#)

Bookmark

Dynamic prediction

Strict vs relaxed data landmarking

Penalized Regression Calibration

The R package `pencal`

Benchmarking

What's cooking in the dynamic prediction pot

References

Appendix



Why such a weird title?

Why such a weird title?

- ▶ My relationship with dynamic prediction (DP)
 1. Started working on DP problems around 2019



Why such a weird title?

- ▶ My relationship with dynamic prediction (DP)
 1. Started working on DP problems around 2019
 2. 2019-2024: 4 projects and publications on DP, **but: still a lot of work to do!**



Why such a weird title?

- ▶ My relationship with dynamic prediction (DP)
 1. Started working on DP problems around 2019
 2. 2019-2024: 4 projects and publications on DP, **but: still a lot of work to do!**
 3. 2024: **begged NWO for money** ("VIDI") to hire 2 PhD students & 1 postdoc → they granted the money ☺



Why such a weird title?

- ▶ My relationship with dynamic prediction (DP)
 1. Started working on DP problems around 2019
 2. 2019-2024: 4 projects and publications on DP, **but: still a lot of work to do!**
 3. 2024: **begged NWO for money** ("VIDI") to hire 2 PhD students & 1 postdoc → they granted the money ☺
 4. 2025-2030: **6 more years** working on DP methods & software



Why such a weird title?



- ▶ My relationship with dynamic prediction (DP)
 1. Started working on DP problems around 2019
 2. 2019-2024: 4 projects and publications on DP, **but: still a lot of work to do!**
 3. 2024: **begged NWO for money** ("VIDI") to hire 2 PhD students & 1 postdoc → they granted the money ☺
 4. 2025-2030: **6 more years** working on DP methods & software
- ▶ So, why this title?!?
 1. **2019 → 2030** ⇒ it feels like **I'm making a stew** with plenty of different ingredients and flavours

Why such a weird title?



- ▶ My relationship with dynamic prediction (DP)
 1. Started working on DP problems around 2019
 2. 2019-2024: 4 projects and publications on DP, **but: still a lot of work to do!**
 3. 2024: **begged NWO for money ("VIDI")** to hire 2 PhD students & 1 postdoc → they granted the money ☺
 4. 2025-2030: **6 more years** working on DP methods & software
- ▶ So, why this title?!?
 1. **2019 → 2030** ⇒ it feels like **I'm making a stew** with plenty of different ingredients and flavours
 2. Time to try to tell what I've been cooking in the DP pot!

What's cooking in the dynamic prediction pot

Limitations of PRC:

1. LMM and MPLMM: linear only
2. single survival outcome, right-censored



What's cooking in the dynamic prediction pot



Limitations of PRC:

1. LMM and MPLMM: linear only
2. single survival outcome, right-censored

Ongoing and future work:

- ▶ longitudinal data
 1. LMMs → GLMMs (with Xiang Li)
 2. more multivariate mixed modelling options

What's cooking in the dynamic prediction pot



Limitations of PRC:

1. LMM and MPLMM: linear only
2. single survival outcome, right-censored

Ongoing and future work:

- ▶ longitudinal data
 1. LMMs → GLMMs (with Xiang Li)
 2. more multivariate mixed modelling options
- ▶ survival outcome(s)
 1. cure models (with Marta Cipriani and prof. Marco Alfò)
 2. prediction intervals (with Lorenzo Carvisiglia and dr. Saverio Ranciati)
 3. competing risks
 4. interval censoring

Summary

- ▶ Multistep DP methods
 - 1. strike a balance between landmarking and joint modelling
 - 2. enable DP when you have numerous longitudinal predictors
 - 3. should be estimated using strict data landmarking (Gomon et al., 2024)

Summary

- ▶ Multistep DP methods
 - 1. strike a balance between landmarking and joint modelling
 - 2. enable DP when you have numerous longitudinal predictors
 - 3. should be estimated using strict data landmarking (Gomon et al., 2024)
- ▶ Penalized regression calibration (PRC, Signorelli et al. (2021)):
 - 1. step 1 & 2: mixed-effects models, step 3: penalized Cox model
 - 2. usually outperforms LOCF landmarking, joint models and the other multistep methods on ADNI, ROSMAP and PBC2 datasets (Signorelli & Retif (2025+))

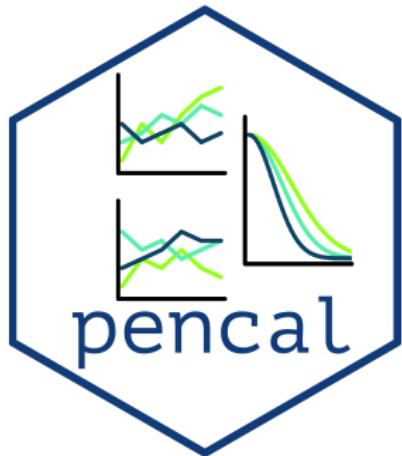
Summary

- ▶ Multistep DP methods
 - 1. strike a balance between landmarking and joint modelling
 - 2. enable DP when you have numerous longitudinal predictors
 - 3. should be estimated using strict data landmarking (Gomon et al., 2024)
- ▶ Penalized regression calibration (PRC, Signorelli et al. (2021)):
 - 1. step 1 & 2: mixed-effects models, step 3: penalized Cox model
 - 2. usually outperforms LOCF landmarking, joint models and the other multistep methods on ADNI, ROSMAP and PBC2 datasets (Signorelli & Retif (2025+))
- ▶ PRC implemented in R package `pencal` (Signorelli, 2024)
 - 1. user-friendly: automates model estimation & prediction
 - 2. easy parallelization: specify argument `n.cores > 1` inside `pencal`'s functions

Summary

- ▶ Multistep DP methods
 1. strike a balance between landmarking and joint modelling
 2. enable DP when you have numerous longitudinal predictors
 3. should be estimated using strict data landmarking (Gomon et al., 2024)
- ▶ Penalized regression calibration (PRC, Signorelli et al. (2021)):
 1. step 1 & 2: mixed-effects models, step 3: penalized Cox model
 2. usually outperforms LOCF landmarking, joint models and the other multistep methods on ADNI, ROSMAP and PBC2 datasets (Signorelli & Retif (2025+))
- ▶ PRC implemented in R package `pencal` (Signorelli, 2024)
 1. user-friendly: automates model estimation & prediction
 2. easy parallelization: specify argument `n.cores > 1` inside `pencal`'s functions
- ▶ 2025-2030: extensions in the making (stew in progress)

Thank you for listening!



[🏠 mirkosignorelli.github.io](https://mirkosignorelli.github.io)
[✉️ m.signorelli@math.leidenuniv.nl](mailto:m.signorelli@math.leidenuniv.nl)
[in/mirko-signorelli](https://www.linkedin.com/in/signorelli)

Dynamic prediction

Strict vs relaxed data landmarking

Penalized Regression Calibration

The R package `pencal`

Benchmarking

What's cooking in the dynamic prediction pot

References

Appendix

References I

- Devaux, A., Helmer, C., Genuer, R., & Proust-Lima, C. (2023). Random survival forests with multivariate longitudinal endogenous covariates. *Statistical Methods in Medical Research*, 32(12), 2331–2346.
- Gomon, D., Putter, H., Fiocco, M., & Signorelli, M. (2024). Dynamic prediction of survival using multivariate functional principal component analysis: A strict landmarking approach. *Statistical Methods in Medical Research*, 33(2), 256–272.
- Li, K., & Luo, S. (2019). Dynamic prediction of Alzheimer's disease progression using features of multiple longitudinal outcomes and time-to-event data. *Statistics in Medicine*, 38(24), 4804–4818.
- Lin, J., Li, K., & Luo, S. (2021). Functional survival forests for multivariate longitudinal outcomes: Dynamic prediction of Alzheimer's disease progression. *Statistical Methods in Medical Research*, 30(1), 99–111.
- Proust-Lima, C., Amieva, H., & Jacqmin-Gadda, H. (2013). Analysis of multivariate mixed longitudinal data: A flexible latent process approach. *British Journal of Mathematical and Statistical Psychology*, 66(3), 470–487.
- Signorelli, M. (2024). pencal: an R Package for the Dynamic Prediction of Survival with Many Longitudinal Predictors. *The R Journal*, 16(2), 134–153.
- Signorelli, M., & Retif, S. (2025+). *Benchmarking multi-step methods for the dynamic prediction of survival with numerous longitudinal predictors*. arXiv:2403.14336.

References II

- Signorelli, M., Spitali, P., Szigyarto, C., Consortium, M.-M., & Tsonaka, R. (2021). Penalized regression calibration: A method for the prediction of survival outcomes using complex longitudinal and high-dimensional data. *Statistics in Medicine*, 40(27), 6178–6196.
- Tsiatis, A. A., & Davidian, M. (2004). Joint modeling of longitudinal and time-to-event data: An overview. *Statistica Sinica*, 809–834.
- Van Houwelingen, H. C. (2007). Dynamic Prediction by Landmarking in Event History Analysis. *Scandinavian Journal of Statistics*, 34(1), 70–85.

Dynamic prediction

Strict vs relaxed data landmarking

Penalized Regression Calibration

The R package `pencal`

Benchmarking

What's cooking in the dynamic prediction pot

References

Appendix

Mark-MD application

SIGNORELLI ET AL.

Statistics
in Medicine - WILEY | 17

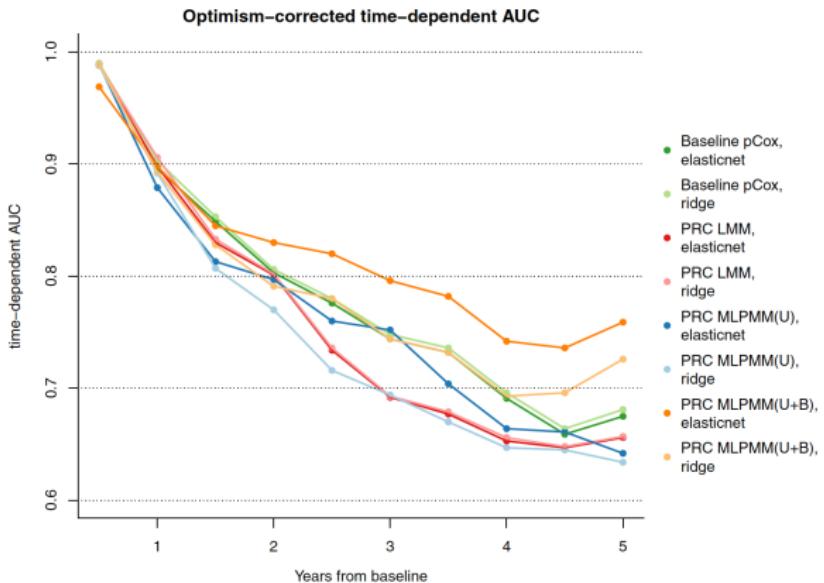


FIGURE 9 Optimism-corrected estimates of the tdAUC for the prediction of time to LoA in the MARK-MD dataset [Colour figure can be viewed at wileyonlinelibrary.com]

Parallelization and computing time

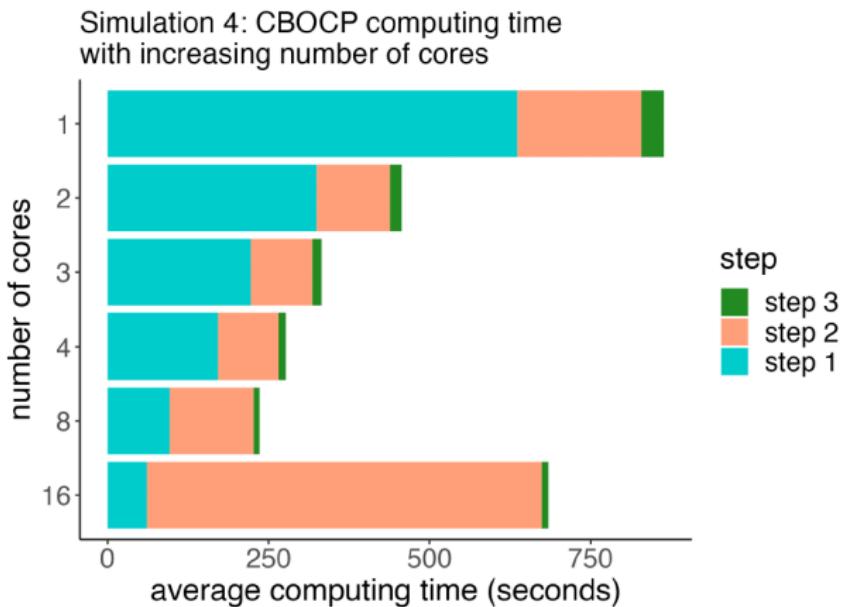
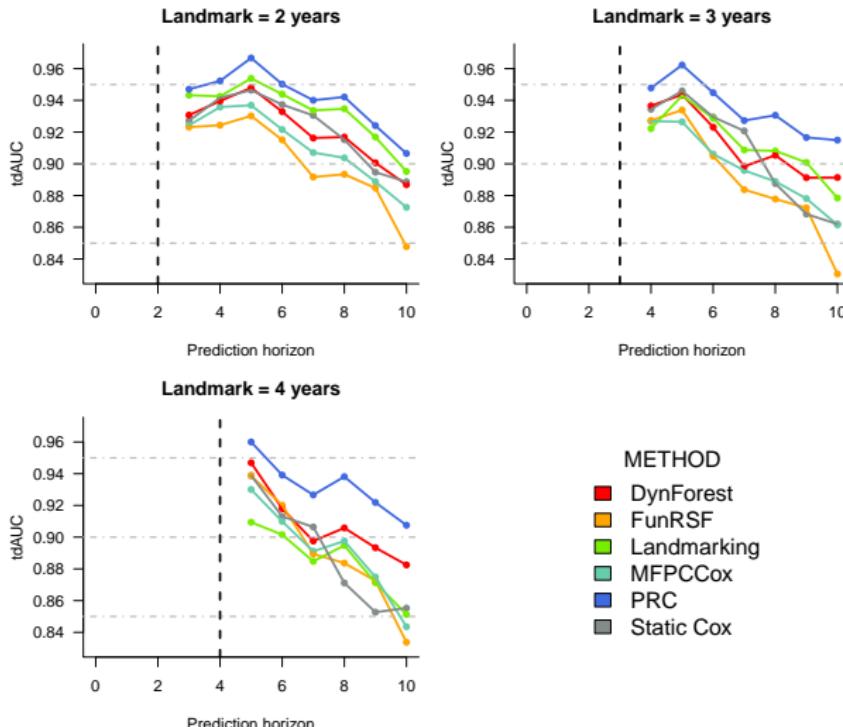
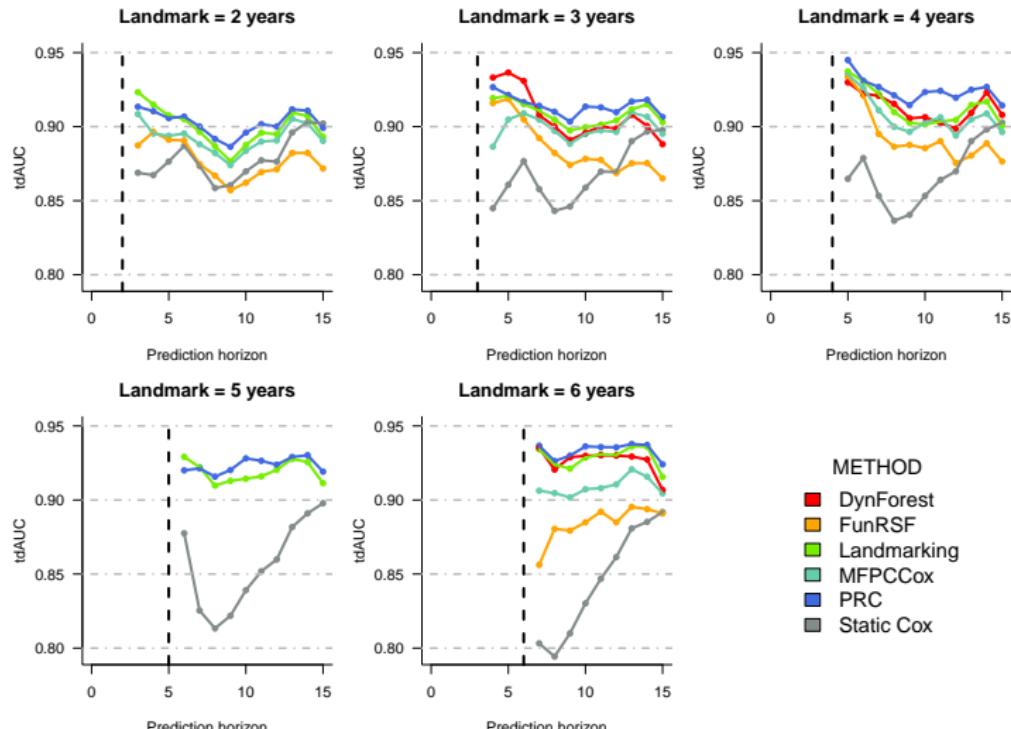


Figure 5: Average computing time (in seconds) for the estimation of the PRC LMM model and the computation of the CBOCP as a function of the number of cores.

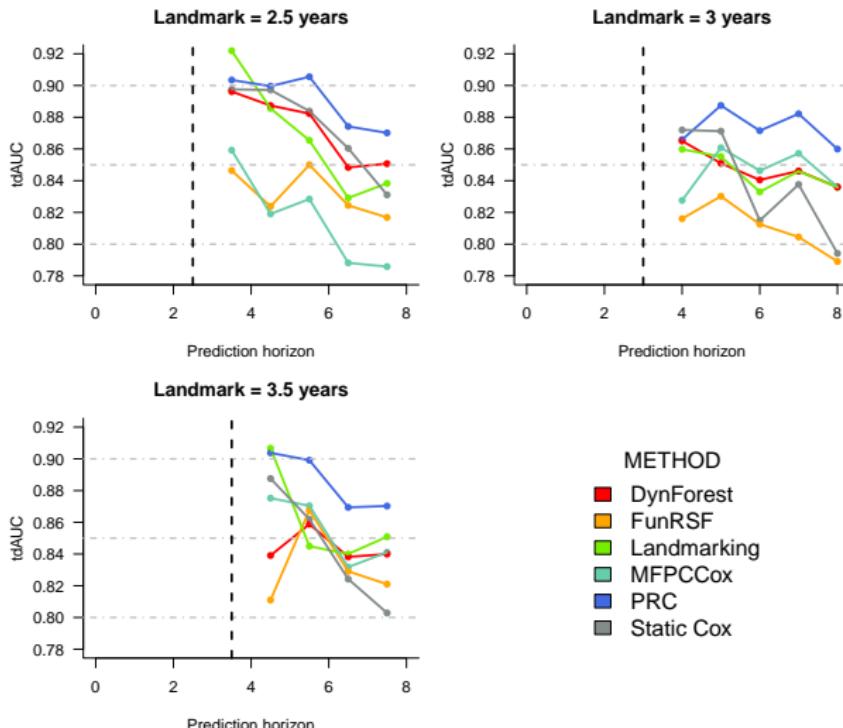
Benchmarking: ADNI



Benchmarking: ROSMAP



Benchmarking: PBC2



Benchmarking: computing time

Method	Landmark			
	2	3	4	Average
Static Cox	0.009	0.007	0.006	0.007
Landmarking	0.010	0.007	0.006	0.008
MFPCCox	0.080	0.046	0.046	0.057
PRC	0.776	0.482	0.453	0.571
FunRSF	0.240	0.122	0.125	0.163
DynForest	12.501	9.077	8.099	9.892

Table 4: Average computing time per CV fold (in **minutes**) for the ADNI dataset.

Method	Landmark					
	2	3	4	5	6	Average
Static Cox	0.022	0.019	0.016	0.014	0.011	0.017
Landmarking	0.023	0.020	0.017	0.014	0.011	0.017
MFPCCox	0.194	0.135	0.066	-	0.064	0.115
PRC	3.755	1.151	1.138	1.115	1.124	1.657
FunRSF	0.604	0.367	0.115	-	0.124	0.302
DynForest	-	11.504	17.657	-	34.147	21.102

Table 5: Average computing time per CV fold (in **minutes**) for the ROSMAP dataset.