Supplementary material of the article
"Penalized regression calibration: a method for the prediction of survival outcomes using complex longitudinal and high-dimensional data"

Mirko Signorelli[1], Pietro Spitali[2], Cristina Al-Khalili Szigyarto[3], The MARK-MD Consortium, and Roula Tsonaka[1]

[1]*Department of Biomedical Data Sciences, Leiden University Medical Center (NL)*
[2]*Department of Human Genetics, Leiden University Medical Center (NL)*
[3]*Department of Protein Science, KTH-Royal Institute of Technology (SW)*

1 Supplementary Tables

	Simulation 7		Simulation 10	
	$p = 50$	$p = 50$	$p = 150$	$p = 150$
Method	$n = 100$	$n = 300$	$n = 100$	$n = 300$
Baseline pCox	0.49	0.63	0.51	1.76
PRC LMM	2.46	4.92	11.14	23.14
PRC MLPMM(U)	20.51	47.30	60.19	163.57
PRC MLPMM(U+B)	20.73	47.58	60.34	166.26

Table 1: Model estimation: mean computing time (in seconds) in simulations 7 and 10. The table compares the mean computing time required to estimate each model for different number of longitudinal predictors (denoted by p) and different sample sizes (n).

	Simulation 7		Simulation 10	
	$p = 50$	$p = 50$	$p = 150$	$p = 150$
Method	$n = 100$	$n = 300$	$n = 100$	$n = 300$
Baseline pCox	20.18	29.34	24.14	56.27
PRC LMM	265.27	586.52	1631.55	3627.03
PRC MLPMM(U)	1303.48	3758.61	4923.15	13877.11
PRC MLPMM(U+B)	1303.02	3759.50	4932.86	13903.94

Table 2: Optimism correction: mean computing time (in seconds) in simulations 7 and 10. The table compares the mean computing time needed to compute the CBOCP (or, for the baseline pCox model, a simpler bootstrap optimism correction) for different number of longitudinal predictors (denoted by p) and different sample sizes (n).

2 Supplementary Figures

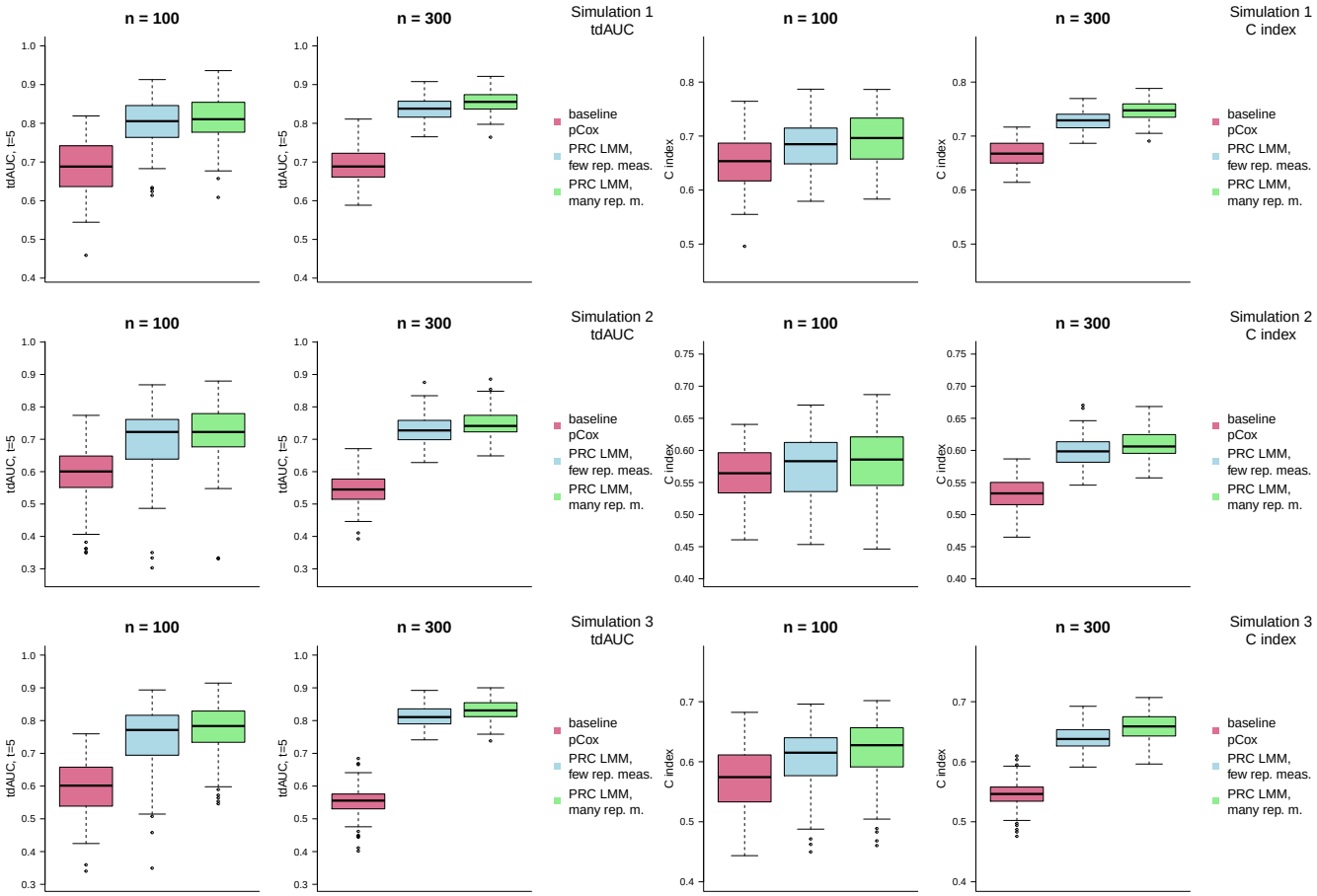

Figure 1: Results of simulations 1, 2 and 3 using the elasticnet penalty. The boxplots compare the distribution over 100 random replications of the optimism-corrected tdAUC (left) and C index (right) of the PRC LMM model when few (lightblue) or many (lightgreen) repeated measurements are available to that of a penalized Cox model where only baseline measurements are used (violet red).

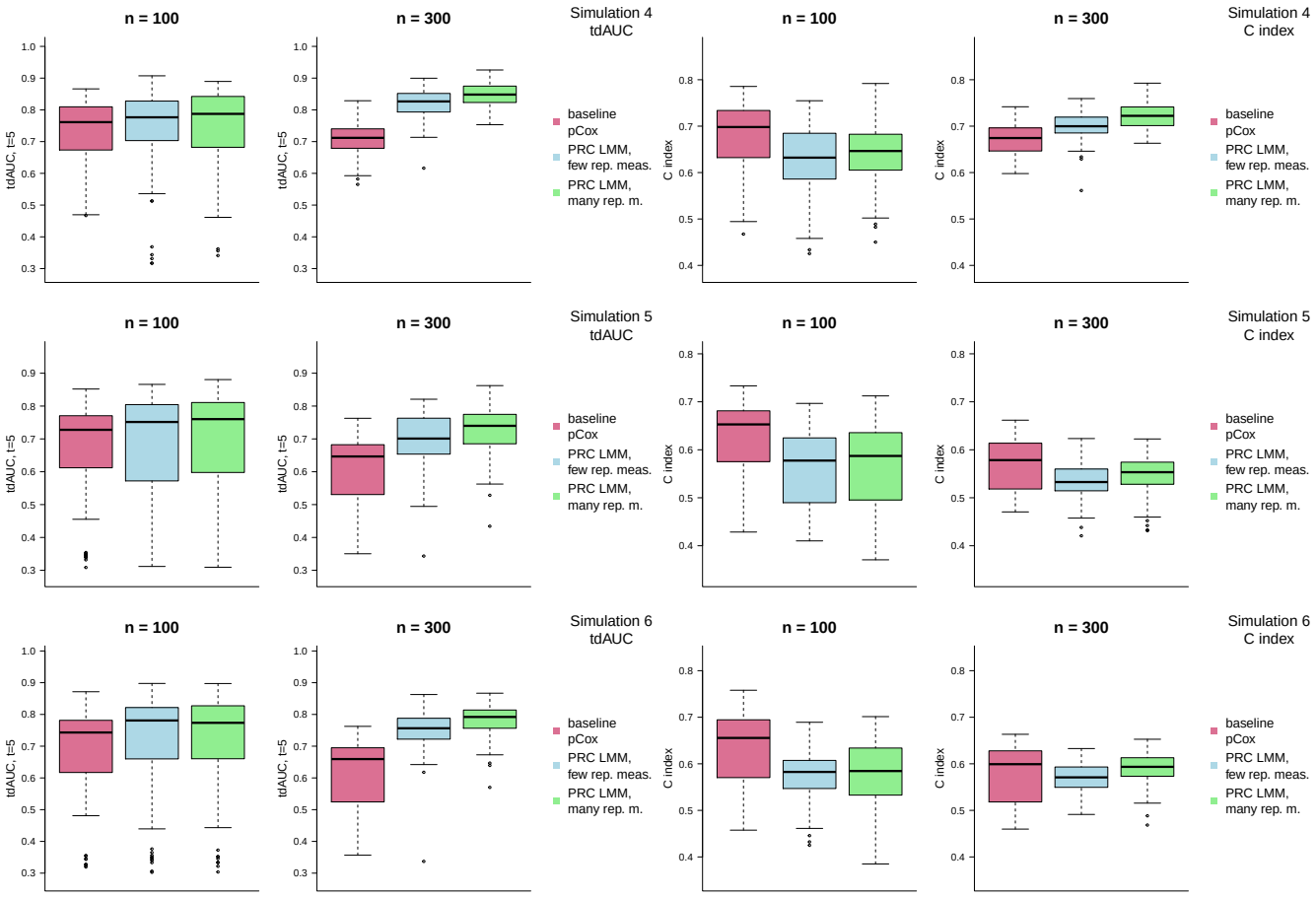

Figure 2: Results of simulations 4, 5 and 6 using the elasticnet penalty. The boxplots compare the distribution over 100 random replications of the optimism-corrected tdAUC (left) and C index (right) of the PRC LMM model when few (lightblue) or many (lightgreen) repeated measurements are available to that of a penalized Cox model where only baseline measurements are used (violet red).

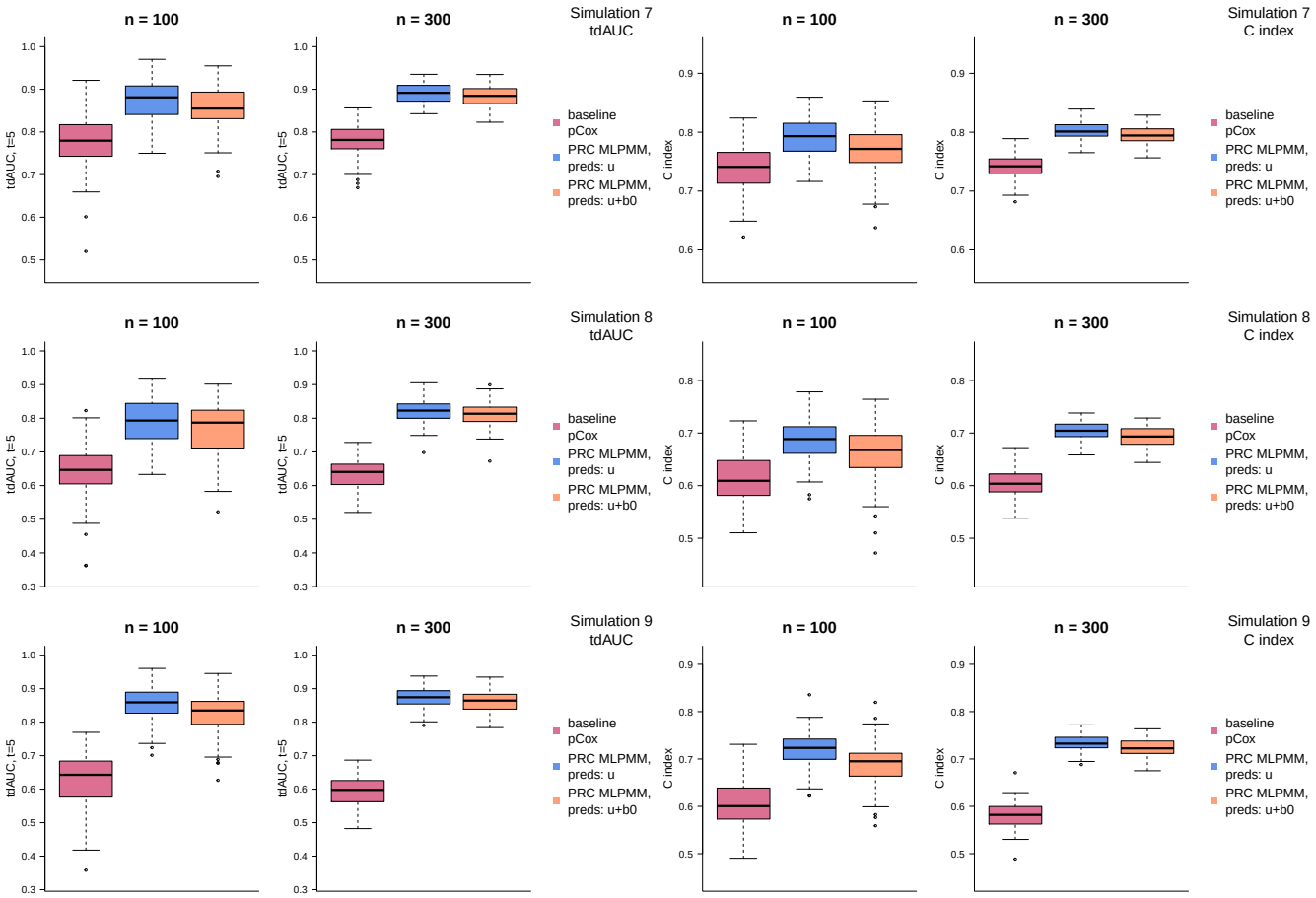

Figure 3: Results of simulations 7, 8 and 9 using the elasticnet penalty. The boxplots compare the distribution over 100 random replications of the optimism-corrected tdAUC (left) and C index (right) of the baseline pCox (violet red), PRC-MLPMM(U) (blue) and PRC-MLPMM(U+B) (orange) models.

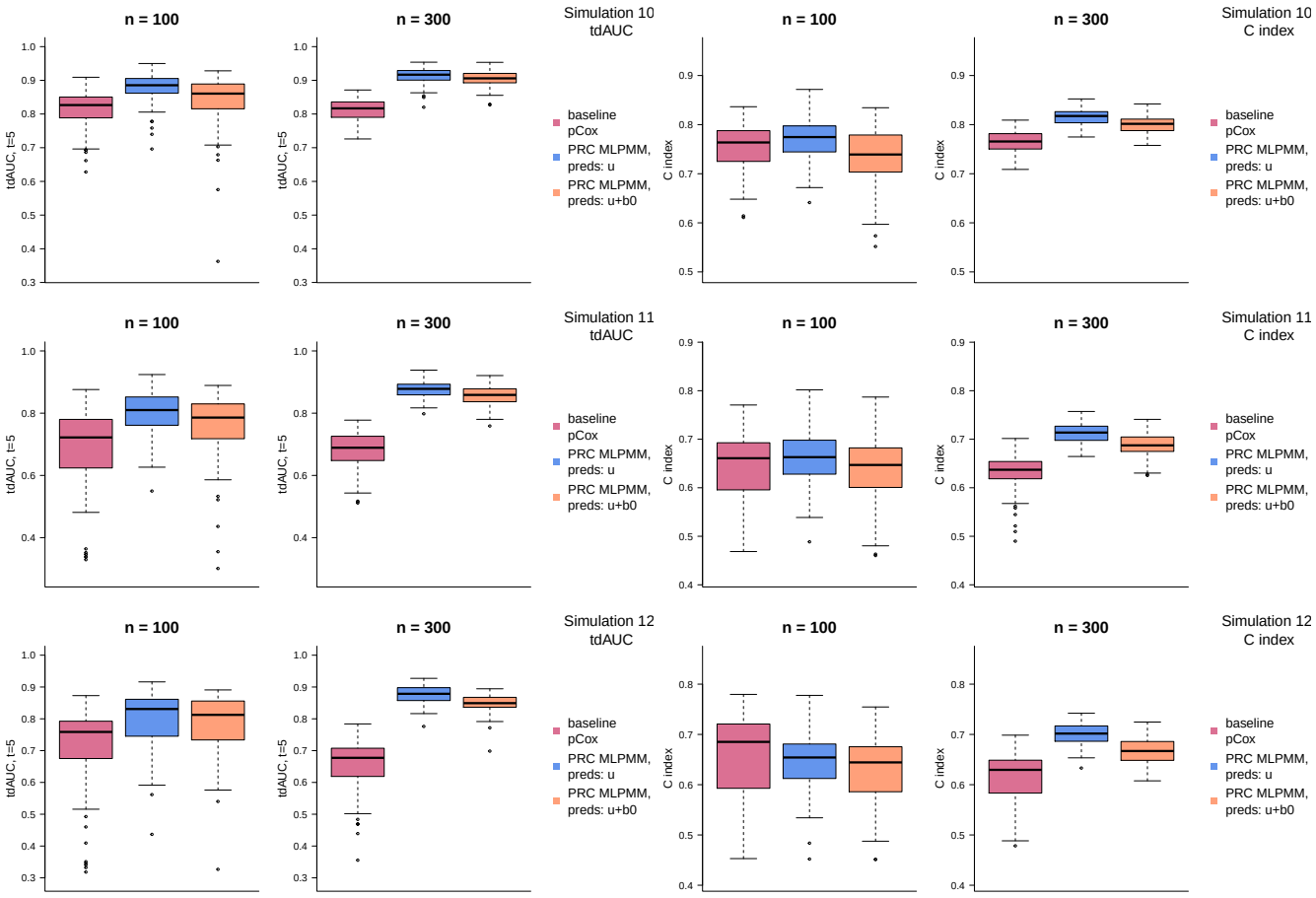

Figure 4: Results of simulations 10, 11 and 12 using the elasticnet penalty. The boxplots compare the distribution over 100 random replications of the optimism-corrected tdAUC (left) and C index (right) of the baseline pCox (violet red), PRC-MLPMM(U) (blue) and PRC-MLPMM(U+B) (orange) models.

3 Comparison to the predictive performance of joint modelling in a low-dimensional settings

The primary motivation for the development of PRC is the fact that the estimation of joint models is computationally intensive, and it is practically unfeasible with more than a handful of longitudinal covariates. In situations where a larger number of longitudinal covariates is available, PRC represents a computationally feasible alternative to joint models for the prediction of survival probabilities. Nevertheless it may be still be interesting to compare the predictive performance of PRC to that of joint models *in low-dimensional settings*, where both approaches can be pursued.

In this Section we compare the predictive performance of PRC to that of joineRML (Hickey et al., 2018), a multivariate joint modelling approach, in a scenario with 3 longitudinal covariates. PRC and joineRML model the relationship between the longitudinal outcomes and the survival time differently, but they are equivalent when the random effects structure comprises only random intercepts. Additionally, joineRML does not offer a built-in strategy to compute optimism-corrected estimates of predictive performance. Therefore, to ensure that the two models are comparable we simulate the longitudinal predictors from LMM models where $y_{sij} = \beta_{s0} + b_{s0i} + (\beta_{s1})a_{ij} + \varepsilon_{sij}$, letting the survival time depend on the random intercepts through a Weibull model. Moreover, we use a split-sample validation approach to estimate predictive performance, with a training set comprising $n = 300$ subjects and a validation set with 200 subjects. joineRML computes conditional survival probabilities from the last available longitudinal measurement; therefore, for simplicity we assume a balanced longitudinal design with 5 longitudinal measurements taken at $t_{ij} = 0, 0.1, 0.3, 0.5, 1$, and that events and censoring occur starting from $t > 1$. We estimate both models on the training set, and we compute the conditional survival probabilities $\hat{S}(t|1)$ for subjects in the validation set.

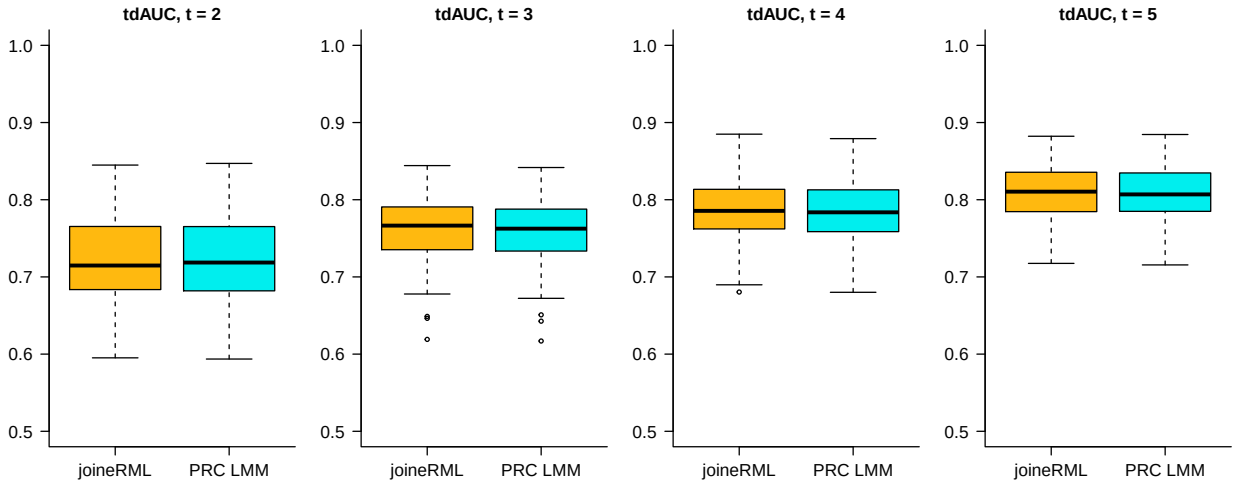

Figure 5: Optimism-corrected time-dependent AUC of joineRML and PRC LMM for predictions at $t = 2, 3, 4, 5$. Results based on 100 replications.

Figure 5 compares the distribution of the optimism-corrected tdAUC for predictions at $t = 2, 3, 4, 5$ for the two models. In short, we can observe that the predictive performance of PRC and joineRML is very similar at each of the considered prediction horizons. It should be noted, however, that estimation of joineRML is considerably slower than that of the PRC LMM: on average, the esti-

mation of PRC LMM took 0.47 seconds, whereas that of joineRML 453.84 seconds (computations were run using a single core on an Intel E7-4890 processor with 2.2 GhZ CPU).

References

Hickey, G. L., Philipson, P., Jorgensen, A., and Kolamunnage-Dona, R. (2018). joinerml: a joint model and software package for time-to-event and multivariate longitudinal outcomes. *BMC medical research methodology*, 18(1):1–14.