

Joint Models for Dynamic Prediction in Localised Prostate Cancer: A Literature Review

Harry Parr^{1*}, Emma Hall¹, Nuria Porta¹

¹The Institute of Cancer Research – Clinical Trials & Statistics Unit (ICR-CTSU), 15 Cotswold Rd, London SM2 5NG UK

* Corresponding author email: harry.parr@icr.ac.uk

Prof Emma Hall email: emma.hall@icr.ac.uk

Dr Nuria Porta email: nuria.porta@icr.ac.uk

Abstract

Background

Prostate cancer is a very prevalent disease in men. Patients are monitored regularly during and after treatment with repeated assessment of prostate-specific antigen (PSA) levels. Prognosis of localised prostate cancer is generally good after treatment, and the risk of having a recurrence is usually estimated based on factors measured at diagnosis. Incorporating PSA measurements over time in a dynamic prediction joint model enables updates of patients' risk as new information becomes available. We review joint model strategies that have been applied to model time-dependent PSA trajectories to predict time-to-event outcomes in localised prostate cancer.

Methods

We identify articles that developed joint models for prediction of localised prostate cancer recurrence over the last two decades. We report, compare, and summarise the methodological approaches and applications that use joint modelling accounting for two processes: the longitudinal model (PSA), and the time-to-event process (clinical failure). The methods explored differ in how they specify the association between these two processes.

Results

Twelve relevant articles were identified. A range of methodological frameworks were found, and we describe in detail shared-parameter joint models (9 of 12, 75%) and joint latent class models (3 of 12, 25%). Within each framework, these articles presented model development, estimation of dynamic predictions and model validations.

Conclusions

Each framework has its unique principles with corresponding advantages and differing interpretations. Regardless of the framework used, dynamic prediction models enable real-time prediction of individual patient prognosis. They utilise all available longitudinal information, in addition to baseline prognostic risk factors, and are superior to traditional baseline-only prediction models.

Keywords

Dynamic prediction models, prostate cancer, PSA, joint modelling, dynamic predictions, personalised medicine.

1 Background

Prostate cancer is highly prevalent, as the 2nd most diagnosed cancer in men worldwide (1.3 million cases in 2018) [1] and the most common cancer among men in the UK [2,3]. Within the UK, one in eight men on average are diagnosed with prostate cancer, increasing to 1-in-6 for men born after 1960. In the period 2014-16 there were almost 48,000 diagnosed per-annum and over 11,500 deaths in 2016 from prostate cancer [4–6]. Over half of patients (56.5—61.3%) present at diagnosis with localised prostate cancer, where disease is confined to the prostate and has not yet spread to the nodes or other organs of the body [7,8]. Many treatment options are available to patients with localised prostate cancer, including external-beam radiotherapy (EBRT), brachytherapy, radical prostatectomy; or conservative management strategies for favourable-risk prostate cancer to delay or avoid aggressive treatment and potential side effects [9]. Potential treatment-related toxicities and side effects often affect management treatment choices [10]. EBRT is deemed most appropriate for those with moderate- or high risk disease [11]. A combination with hormonal therapy (HT) can be given for its neoadjuvant efficacy [9]: for low-grade localised prostate cancer, HT can be given 3-6 months before EBRT treatment, or longer for higher-risk stages [12].

Patients are monitored regularly during, and after treatment. In particular, for their prostate-specific antigen (PSA), a serine protease protein biomarker secreted by the prostate [13]. Repeated PSA readings are taken during patient check-ups. Patients present at diagnosis with elevated PSA levels, which decrease upon starting of HT and RT. Thus, increased levels of PSA after treatment suggest a growth of prostate cancer cells, reflecting a higher risk of prostate cancer recurrence. PSA is used to determine biochemical failure (BcF), defined as a PSA concentration greater than the nadir (the lowest observed PSA value) plus 2 ng/mL [14]. HT can be used as a salvage therapy following BcF, lowering PSA levels and decreasing risk of clinical failure. Local or distant recurrence is confirmed by imaging.

There are known patient and tumour risk factors that affect prognosis of localised prostate cancer. These include PSA levels at diagnosis, tumour stage (as per the TNM scoring system) and Gleason score/grade grouping [15]. These risk factors are used to categorise patients into the National Comprehensive Cancer Network (NCCN) low, intermediate and high risk groups [16] (Table 1). Prognosis of localised prostate cancer (T1–T2N0M0) is generally good after treatment, with 5-year disease-free survival rates around 76% (95% CI: 75%—76%) [17].

Table 1 – Risk stratification by clinical risk factors: Clinical T-stage, Gleason score, and presenting PSA. Locally advance prostate cancer includes high-risk localised patients, as defined by NCCN [16].

Risk level	Clinical T-stage	Gleason Grading Group	Presenting PSA	Condition to be met
Low risk	T1-T2a	1	< 10ng/mL	All three
Intermediate risk	T2b-c	2 or 3	10-20ng/mL	Any
High risk	T3a	4 or 5	>20ng/mL	Any

Clinical prediction models (CPMs) are developed from patient and tumour features at diagnosis, as well as information on short-term treatments, to predict future prognosis. To date, a plethora of CPMs guide management decisions for localised prostate cancer, visualised in nomograms and online calculators [18–26]. These CPMs only consider information available at the time of diagnosis and/or at start/end of treatment, and PSA values collected after that timepoint are rarely considered, if only for the definition of BcF. However, it is of interest to both patient and clinician to examine the association of the biomarker of interest over time to prognosis. Knowing the patient is alive and recurrence-free at the new visit, with an updated PSA value, is informative. Including this new information into a prediction model can elicit dynamic predictions that enable updated prognosis of patients.

A naïve approach would be to consider PSA as a time-dependent variable in an extended Cox/relative risk model [27]. However, this is not appropriate due to the endogenous nature of the biomarker of interest [28,29], which contains biological variation and measurement error. A further extension is to use landmark modelling [28,30–34]: dynamic predictions are obtained by fitting time-dependent Cox models to the patient subsample still at risk at several prediction, or landmark times of interest, together with the value of the longitudinal biomarker at that time. Landmark models are straightforward to fit with standard software, but no measurement error for the time-varying biomarker is considered nor is the entire longitudinal history of the biomarker utilised (due to using the last observation carried forward) [34]. To improve predictions, a two-stage approach to landmarking (also known as mixed model landmarking [28,35]) can be considered to model measurement error and incorporate the full biomarker history. However, uncertainties in the mixed-effect model estimates are not carried through to the survival submodel, resulting in overexact estimates [36].

Joint models (JMs) permit dynamic prediction in localised prostate cancer by considering two time-dependent processes simultaneously: the repeated longitudinal PSA biomarker over time (modelled using a mixed-effects submodel), and the time to an event of interest (modelled using a relative-risk, or Cox submodel). The event of interest can be BcF, recurrence of disease (either locally in the prostate gland, in the regional lymph nodes or in a distant organ), clinical failure (need to re-commence HT), death, or a composite of all these events. The association between these two processes can be captured by shared random effects in both the longitudinal and time-to-event submodels (shared-parameter joint models, SPJMs) or by assuming a latent association structure between them (joint latent class models, JLCMs).

In this paper, we synthesize a review for published applications of joint models to localised prostate cancer over the last two decades, focusing on the modelling of the time-to-event process(es), the functional form of PSA, validation strategies and evaluation of dynamic predictions. We describe the search strategy to identify papers, and we briefly describe the joint modelling methodology, as well as how to compute dynamic predictions, measures of predictive performance from joint models. Given the rapid popularity and use of dynamic prediction models, this article serves as a reference to assess and reflect the applied and dynamic methods used in localised prostate cancer. The main review of the identified articles is given in the results section and summarised on Table 2. Finally, an appraisal and conclusion of these models are given.

2 Methods

2.1 Literature search strategy

Our search strategy included linear combinations of, {"joint model*" OR "individual* prediction"} AND {"prostate cancer" OR "prostate-specific antigen" OR "PSA"} in the title or abstract, using Web of Science and PubMed databases up to and including June 2020. A flowchart depicting the study identification strategy is given in Figure 1. A total of 751 articles were identified from the initial search parameters, 703 and 48 articles came from Web of Science and PubMed respectively. Duplicated articles were removed leaving 702 unique papers. Novel and seminal papers that involve the joint modelling of the longitudinal biomarker PSA and time-to-event of clinical recurrence in localised prostate cancer were selected by the lead author, and selection discussed with co-authors, as the focus was to understand the PSA dynamics for this disease, which can be quite different from PSA dynamics for advanced prostate cancer. Further exclusions were made on inspecting the abstract, these included: advance/metastatic disease; different disease; no joint modelling undertaken, or alternative machine learning/artificial intelligence methods used; simulated data used; predicting alternative endpoints such as time to diagnosis or death; no dynamic predictions derived; and whether focus was on methodology development.

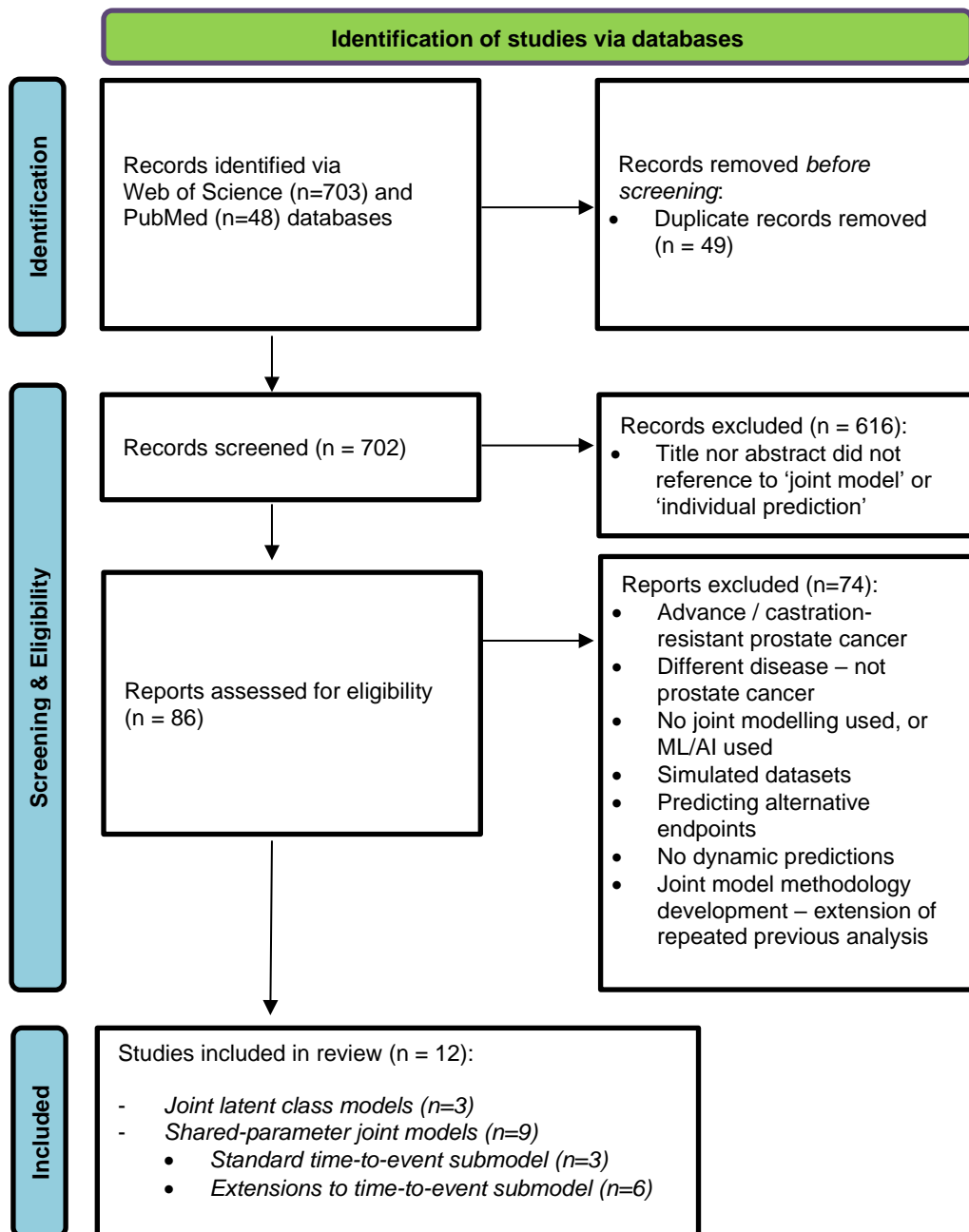


Figure 1 – A flowchart for identifying studies of the literature review.

2.2 Notation

In this section, we define the mathematical notation common to SPJMs & JLCMs frameworks. Let $\mathbf{y}_i = \{y_i(t_{ij}); i = 1, \dots, N; j = 1, \dots, n_i\}$ be a longitudinal response vector of the continuous biomarker measurements for the i^{th} patient and j^{th} biomarker reading taken at time t_{ij} . There are N patients with n_i longitudinal measurements per patient.

Let the random variable T_i be the time-to-failure for the i^{th} patient, where $T_i = \min(T_i^*, C_i)$. The true event time is denoted T_i^* and C_i is the censoring time. An indicator variable $\delta_i = I(T_i^* \leq C_i)$ is unity if the event of interest is observed for that patient, or zero otherwise.

2.3 Shared-parameter joint model

Under the shared-parameter joint models framework, random effects are used to link the longitudinal and time-to-event components under study, whilst also accounting for the correlated repeated measurements within the longitudinal outcome.

The longitudinal process \mathbf{y}_i is assumed to follow a mixed-effects model, defined by a linear combination of possibly time-dependent main- and random effects $Y_i(t_{ij}) = m_i(t) + \epsilon_i(t_{ij}) = \boldsymbol{\beta}X_i(t_{ij}) + \mathbf{b}_iZ_i(t_{ij}) + \epsilon_i(t_{ij})$. The vector $\boldsymbol{\beta}$ are coefficients for the main- and time-effect covariates of the design matrix X_i , and the corresponding random effects \mathbf{b}_i for the Z_i design matrix. The measurement errors $\epsilon_i(t_{ij}) = \{\epsilon_i(t_{i1}), \dots, \epsilon_i(t_{in_i})\}^T$ are independent and identically distributed and assumed to follow $\epsilon_i(t_{ij}) \sim N(0, \sigma_\epsilon^2)$, or t-distribution with several degrees-of-freedom, with the fatter tails used to accommodate for possible outliers. The random effects, independent of $\epsilon_i(t_{ij})$, are usually assumed to follow a multivariate normal distribution, with an unknown square covariance matrix structure D , $\mathbf{b}_i \sim \text{MVN}(\mathbf{0}, D)$.

A relative risk, or proportional hazards model, is used for the parameterisation of the survival submodel:

$$\begin{aligned} h_i(t|\mathbf{M}_i(t), \mathbf{w}_i) &= \lim_{\Delta t \rightarrow 0} \frac{\Pr\{t \leq T_i^* < t + \Delta t | T_i^* \geq t, \mathbf{M}_i(t), \mathbf{w}_i\}}{\Delta t} \\ &= h_0(t) \exp\{\boldsymbol{\gamma}^T \mathbf{w}_i + f(\mathbf{M}_i(t), \mathbf{b}_i, \boldsymbol{\alpha})\}. \end{aligned}$$

Where $\mathbf{M}_i(t)$ denotes the true (unobserved) and entire longitudinal biomarker history up to time point t , with $m_i(t)$ indicating the true value at t (i.e. the mixed effect model not contaminated with measurement error). The baseline covariates in the hazard submodel are \mathbf{w}_i , with $\boldsymbol{\gamma}^T$ corresponding to the log-hazard ratio coefficients. An example of the parameterisation of the functional form $f(\mathbf{M}_i(t), \mathbf{b}_i, \boldsymbol{\alpha})$ can be a linear combination of value and gradient of the longitudinal biomarker, $f(\dots) = \alpha_1 m_i(t) + \alpha_2 \frac{dm_i(t)}{dt}$. The corresponding $\boldsymbol{\alpha}$ parameters quantify the intensity of association

between the two outcomes. Other functional forms of f exist, such as the (weighted) cumulative effect (1), or random effects association (2),

$$f = \alpha \int_0^t \omega(t-s) \times m_i(s) ds \quad (1)$$

$$f = \alpha^T \mathbf{b}_i \quad (2)$$

The former quantifies the risk of recurrence from the area under the biomarker trajectory and can allocate greater weights to more recent biomarker observations, e.g., using a standard normal density function for ω . The latter parameterisation uses only the random effects as a linear predictor, this requires no numerical integration which is computationally advantageous. Using a simple random intercept- and slopes structure is most interpretable, whereby patient deviations from the population average is expressed [37]. More elaborate structures are challenging to interpret [29,38,39]. A full parametric specification of the baseline hazard function, $h_0(t)$, is recommended (e.g. using constant-piecewise, or regression splines models), with an adequate number of knots for flexibly modelling the underlying baseline risk. Leaving $h_0(t)$ unspecified can lead to underestimating the precision of parameter estimates [40].

Given the random effects \mathbf{b}_i , Y_i and T_i become independent (conditional independence). Excellent overviews of shared-parameter joint modelling can be found in Rizopoulos [37] and Papageorgiou et al. [29].

2.4 Joint latent class model

The joint latent class models framework assumes the existence of latent classes that capture the association between the longitudinal biomarker trajectory and the relative risk of the endpoint of interest. Following the same notation as above, we can define the JLCM by the mixed-effect- and relative risk submodels for each latent class $\mathbf{c}_i \in \{1, \dots, G\}^T$:

$$(Y_i(t_{ij}) | \mathbf{c}_i = g) = \boldsymbol{\beta}_g X_i^T(t_{ij}) + \mathbf{b}_{ig} Z_i^T(t_{ij}) + \epsilon(t_{ij}); \epsilon(t_{ij}) \sim N(0, \sigma_e^2), \mathbf{b}_{ig} \sim \text{MVN}(\boldsymbol{\mu}_g, D)$$

$$h_i(t | \mathbf{c}_i = g) = h_{0g}(t) \exp(\boldsymbol{\gamma}_g^T \mathbf{w}_i),$$

where assignment to latent class g is given by a multinomial submodel,

$$\Pr(\mathbf{c}_i = g | X_i) = \frac{\exp(\boldsymbol{\lambda}_g^T X_i)}{\sum_{j=1}^G \exp(\boldsymbol{\lambda}_j^T X_i)}.$$

With X_i a fixed baseline design matrix associated with classification and corresponding coefficients $\boldsymbol{\lambda}_g^T = (\lambda_0^T = 0, \lambda_1^T, \dots, \lambda_G^T = 0)$. Given the latent class \mathbf{c}_i , conditional independence between the longitudinal and time-to-event outcomes is assumed.

The JLCM has some advantages compared to the SPJM: it does not need to specify a suitable functional form to link the two processes, and thus the conditional independence assumption in the JLCMs results in less onerous computations. However, as the number of latent classes are not known a priori, it is another component to be estimated, and as these are not observed, the conditional independence assumption is nontrivial to evaluate. Jacqmin-Gadda et al. [41], proposed a trivariate score test to evaluate this assumption, they showed that their score test was uniformly most powerful and simpler than all other considered tests.

2.5 Dynamic predictions

Given a sample from the population of interest, joint models permit to compute dynamic predictions of the event of interest at a future time u given the information available up to time $t > 0$. For a specific i^{th} individual, these are defined by $\pi_i(u|t) = \Pr(T^* \geq u | T^* > t, X_i, \mathbf{y}_i(t), T_i, \delta_i, \mathbf{w}_i, \boldsymbol{\theta})$. I.e., the conditional probability of being event-free at time $u > t$, given that the patient is still at risk of the event at time t ($T^* > t$), the baseline covariates / fixed effects design matrix X_i , the biomarker longitudinal values observed up to time t , $\mathbf{y}_i(t)$, and the parameters $\boldsymbol{\theta}$ estimated from the joint model. These predictions, which can be then *dynamically* updated when new biomarker information becomes available at $t' > t$ [28].

For the shared-parameter JMs, these are extracted by integrating the conditional event probability $\pi_i(u|t)$ over the random effects. Similarly for the JLCM, the predicted probabilities are given by summing over the latent classes. In both frameworks, this is difficult to compute analytically, therefore Markov chain Monte Carlo (MCMC) methods are implemented. MCMC extracts the predicted event posterior distribution of $\pi_i(u|t)$ and corresponding credible intervals from the Monte Carlo sample percentiles of interest [42,43].

2.6 Predictive performance

Measuring predictive ability is crucial to assess the proposed model(s) performance in producing accurate predictions, the end goal for any DPM (dynamic prediction model). Two aspects of modelling performance can be assessed: calibration (how well the model predicts the observed data) and discrimination (how well can the model distinguish between those patients that do and do not have an event).

Discrimination is typically assessed by considering the time-dependent AUROC (area under the receiver operating characteristic curve) [42,44–46]. Within a particular chosen prediction window, AUROC (or simply AUC) values of 0.5 indicate random chance assignment and values closer to unity indicate better model discrimination.

The prediction error (PE) focuses on assessing the calibration of the model, and it is defined as the expectation of the difference between the observed event status $N_i(u|t) = I(T_i^* > u|t)$ and the predicted event occurrence $\pi_i(u|t)$, at a specific time. A loss function can be incorporated within the expectation, e.g., the absolute- or mean squared-loss functions. The latter is also known as the Brier score (BS), which is an

overall measure of prognostic performance [47,48]. Under any loss function, as the difference between these two terms decrease and tends to zero, the closer the observed and predicted event align, resulting in better predictive performance of the model. In practice, one may want to consider predictions over a window of interest, rather than specified time points, by using weighted extensions of these estimators, e.g. weighted average absolute prediction error (WAPE) or integrated BS [49,50]. For any of these predictive measures to be valid, the censoring distributions need to be corrected for, e.g. using inverse probability weighted estimators [28,46,51,52].

Alternative measures of accuracy can be utilised, such as the expected prognostic observed cross-entropy (EPOCE) [53]. The EPOCE quantifies the prognostic information from the joint model at the landmark time of interest. When estimated internally, leave-one-out cross-validation of the prognostic observed log-likelihood (CVPOL) is used to correct for over-optimism [54]. For external validation, no cross-validation is required. Proust-Lima et al. [43] argue the advantages of EPOCE over the previously stated measures, including no censoring distribution nor a prediction window is assumed, direct comparison of two joint models can be made, and that it is more reasonable to evaluate directly on the likelihood density functions. Further formulation and discussion on this predictive accuracy metric can be found in [43,53].

3 Results

We identified 12 relevant full-text papers that best illustrated the joint modelling framework and summarised its applications in localised prostate cancer, these were selected to be included within this review. Table 2 summarises these twelve papers including details of the modelling framework used, sample sizes, parameterisations, the prediction windows of interest, whether validation was undertaken, and the code/software used.

Where available, the corresponding software and code with packages can also be found [55,56]. Nine papers (9 of 12, 75%) applied the shared-parameter joint modelling framework, with three of these presenting the standard joint model for a time-to-failure endpoint, while 6 of 9 papers presented extensions to the time-to-event submodel incorporating cure, competing risks, and multi-state models for localised prostate cancer (e.g., local- and distant recurrence, salvage therapy, and death). Three papers (3 of 12, 25%) described the joint latent class approach. In the following, we review and summarise these papers in detail around their model specification, estimation of dynamic predictions and model validations conducted.

Table 2 – Summary table of joint modelling articles applied to localised prostate cancer and clinical failure, in chronological order. Abbreviations include: joint latent class model (JLCM); shared-parameter joint model (SPJM); prostate-specific antigen (PSA); external-beam radiotherapy (EBRT); hormonal therapy (HT); expected prognostic observed cross-entropy (EPOCE); cross-validated prognostic observed log-likelihood (CVPOL); integrated Brier score (IBS).

Paper [ref]	Modelling Framework	Sample sizes (N) & Events (E) for model development	Joint model parametrisation	Dynamic prediction landmark and prediction window	Validation undertaken	Code & software used
1) Pauler & Finkelstein, 2002 [59]	Bayesian change-point SPJM.	N=676, E=176	PSA data during the first two years was dropped from analysis due to rapid drops of PSA post-EBRT & HT. The random effects include the intercept and the slopes (before & after the change-point). The change-point indicator predicts recurrence. Logged-PSA is modelled with covariates age, presenting PSA, T-stage, with change-point indicator.	Change-point occurring within 10 years. Relapse landmark by four years with a prediction horizon of 10 years.	None performed.	C routine <i>dfpmin</i> , and S-PLUS <i>surv.fit</i> function.

Paper [ref]	Modelling Framework	Sample sizes (N) & Events (E) for model development	Joint model parametrisation	Dynamic prediction landmark and prediction window	Validation undertaken	Code & software used
2) Law et al., 2002 [63]	Frequentist cure SPJM.	N=458, E=92	Two models are fitted, joint-cure and logistic-Cox (no longitudinal PSA consideration). Nonlinear exponential- decay & growth modelled longitudinal logged-PSAs using presenting PSA, T-stage, and Gleason.	Not specified, estimated probabilities of recurrence are given for each patient at some time in the future.	Simulation study performed showing that joint-cure model has better sensitivity and discrimination compared to logistic-Cox model.	MATLAB
3) Yu et al., 2004 [64]	Cure SPJM (comparing Bayesian and Frequentist).	N=458, E=92	Modelled current PSA value and the PSA gradient trajectory. Random effects are modelled parametrically by exponential- decay & growth models adjusting for presenting PSA, T-stage, and Gleason.	Not specified, estimated probabilities of recurrence are given for each patient at some time in the future.	Not done – comparisons are made between the two estimation methods and are shown to be similar to one another.	MATLAB & C++

Paper [ref]	Modelling Framework	Sample sizes (N) & Events (E) for model development	Joint model parametrisation	Dynamic prediction landmark and prediction window	Validation undertaken	Code & software used
4) Taylor et al. 2005 [62]	Bayesian cure SPJM.	N=934, E=140	PSA value & slope and time-dependent hormone therapy commencement indicator is considered, adjusting for baseline covariates: presenting PSA, T-stage, Gleason, age, total dose (Gy), and treatment duration.	Landmarks from last contact, with a prediction window of four years.	Validation performed on data of the same patients used for development, but with further follow-up. The model is shown to be well calibrated and accurately predict new PSA values and recurrence risk.	C++

Paper [ref]	Modelling Framework	Sample sizes (N) & Events (E) for model development	Joint model parametrisation	Dynamic prediction landmark and prediction window	Validation undertaken	Code & software used
5) Yu et al., 2008 [65]	Bayesian cure SPJM.	N=928, E=146	PSA value & slope and time-dependent hormone therapy commencement indicator is considered, adjusting for baseline covariates: presenting PSA, T-stage, Gleason, age, total dose (Gy), and treatment duration.	Landmarks from last contact, with a prediction window of four years.	Validation performed on data of the same patients used for development, but with further follow-up. The model is shown to be well calibrated and accurately predict new PSA values and recurrence risk. Kaplan-Meier plot shows the higher predicted risks go on to have more recurrences indicating its validity.	C++

Paper [ref]	Modelling Framework	Sample sizes (N) & Events (E) for model development	Joint model parametrisation	Dynamic prediction landmark and prediction window	Validation undertaken	Code & software used
6) Proust-Lima & Taylor, 2009 [61]	Frequentist JLCM.	<i>Model development and validation:</i> N=2,386, E=317	Baseline covariates included: presenting PSA, T-stage, and Gleason. The main- and random effects are of the biphasic initial decline and long-term rise. Five latent classes were identified.	Landmarks taken at every six months from 1—3½ years, with a prediction window of three years.	External validation of prediction is performed on two external cohorts. A range of models are explored, the 5-JLCM shows consistently lower absolute- and weighted prediction errors in both cohorts, using prediction windows of 1 and three years.	Not stated but presumably R using the <i>lcmm</i> package.
7) Jacqmin-Gadda et al., 2010 [41]	Frequentist JLCM.	N=459, E=74	Similar to [61] with biphasic longitudinal components for the logged-PSA, considering presenting PSA, T-stage, and Gleason. Four latent classes were identified to be best fitting where the proposed score test did not reject the null of conditional independence.	Only mean evolutions for each of the four classes are given with predicted recurrence-free survival. No windows are specified.	Simulation study performed to appraise score test, where baseline hazard function was misspecified. This methodology was applied to prostate cancer cohort.	Not stated but presumably R using the <i>lcmm</i> package.

Paper [ref]	Modelling Framework	Sample sizes (N) & Events (E) for model development	Joint model parametrisation	Dynamic prediction landmark and prediction window	Validation undertaken	Code & software used
8) Taylor et al. 2013 [57]	Bayesian SPJM.	<i>Model development and validation:</i> N=3,232, E=458	Covariates include presenting PSA, T-stage, and Gleason grade. Longitudinal parameterisation includes PSA value & slope, and time-dependent HT.	Landmarks are given from most recent PSA values with a prediction window of three years.	External validation is performed on fourth dataset. Simpler visual approaches are undertaken, focusing on estimated risk of recurrence three years after treatment using a three year prediction window. Patients are assigned to four risk groups, comparing the training and testing Kaplan-Meier plots, treating commencing hormone therapy as either censored and as an event. The model is deemed adequately calibrated with similar patterns being exhibited between training & testing datasets.	C

Paper [ref]	Modelling Framework	Sample sizes (N) & Events (E) for model development	Joint model parametrisation	Dynamic prediction landmark and prediction window	Validation undertaken	Code & software used
9) Proust-Lima et al., 2014 [43]	Frequentist JLCM & SPJM.	<i>Model development and validation:</i> N=1,178, E=200	Biphasic mixed-effect parameterisation of longitudinal logged-PSA. Baseline covariates: presenting PSA, T-stage and Gleason Four latent classes identified for the JLCM, SPJM included PSA value and slope association structure. All other components had the same model structure for direct comparison.	Landmarks taken at every six months from 1—3½ years, with a prediction window of three years.	Internal and external validation is performed. The EPOCE is estimated internally using CVPOL from 1 – 6 years after EBRT. The difference in EPOCE for 4-JLCM and SPJM shows the 4-JLCM to be a better prognostic model in the first four years. External EPOCEs and integrated BS are shown over the follow-up period. The IBSs and EPOCEs show reduced errors for ≥3-JLCM and SPJM with little difference between the two approaches.	R: using the <i>lcmm</i> and <i>JM</i> packages – code is available on request from authors.

Paper [ref]	Modelling Framework	Sample sizes (N) & Events (E) for model development	Joint model parametrisation	Dynamic prediction landmark and prediction window	Validation undertaken	Code & software used
10) Sène et al., 2016 [58]	Frequentist SPJM.	N=2,386, E=312	Similar to [61] with biphasic longitudinal components for the logged-PSA, considering presenting PSA, T-stage, Gleason, and corrected total EBRT dose. Several specifications of the time-dependent initiation of salvage HT, and the association structures of the longitudinal value and slope of PSA and random effects.	Landmarks from 1.2, 1.6, 2 and 2.6 years are given with a prediction window of recurrence within the next three years. The predicted recurrence probabilities are given under four scenarios of initiating salvage HT immediately, in 1 or 2 years, or not at all.	Internal validation is performed using cross validation for a prediction window of three years. The CVPOL, CV-BS, and CV-IBS are shown for the six model structures. A simpler random effects joint model is best and chosen for the absence of salvage HT; for immediate HT, the JM that separated the PSA trajectory before and after HT is deemed best.	R: <i>JM</i> package with modifications to source code.

Paper [ref]	Modelling Framework	Sample sizes (N) & Events (E) for model development	Joint model parametrisation	Dynamic prediction landmark and prediction window	Validation undertaken	Code & software used
11) Ferrer et al., 2016 [67]	Frequentist multi-state SPJMs	N=1,474; E=941* <small>* sum of all events</small>	Similar to [61] with biphasic longitudinal components for the logged-PSA, considering presenting PSA, T-stage, and Gleason. The longitudinal PSA value and slope was modelled.	For the multi-state process, transition probabilities are given from each transition to any of the other four transitions from the end of treatment throughout follow-up.	A simulation study was undertaken to ensure the estimation process was correct. Diagnostic plots for the residuals and observed/predicted of the longitudinal model.	R: <i>nlme</i> , <i>survival</i> , <i>mstate</i> and <i>JM</i> packages, with adaptations. Code is readily available at the author's GitHub account.

Paper [ref]	Modelling Framework	Sample sizes (N) & Events (E) for model development	Joint model parametrisation	Dynamic prediction landmark and prediction window	Validation undertaken	Code & software used
12) Ferrer et al., 2018 [66]	Frequentist landmarking and cause-specific SPJMs.	Not explicitly stated but as above. N=1,474; E=unknown.	Longitudinal logged-PSA modelled similarly as to [67]. Adjusting for: dataset cohort, age, T-stage, Gleason and presenting-PSA.	Predicted recurrence and competing risk of death probabilities for two patients at their landmarks of 1.3 to 2.5 years using a prediction window of 1½ and three years, comparing JM and landmark modelling.	Simulation study using the prostate patient cohorts to generate similar data. Evaluating robustness of JMs and landmark models, under different assumptions. JMs generally more robust to deviations in assumptions than landmark models, other than a strong violation in the longitudinal PSA biomarker specification where the landmark model performs better.	R: <i>nlme</i> , <i>JM</i> , <i>survival</i> , <i>pseudo</i> , and <i>geepack</i> packages. Code is readily available at the author's GitHub account.

3.1 Shared-parameter joint models to predict recurrence in localised prostate cancer

In this section, we focus our review on three relevant papers that investigated PSA dynamics to predict recurrence in localised prostate cancer using the SPJM framework: Taylor et al. [57], Sene et al. [58], and Pauler & Finkelstein [59]. All three articles develop models in localised prostate cancer patients treated with EBRT in the absence of neoadjuvant HT. Taylor focused on developing a model to creating a clinical prediction tool online; Sene explored the effect on initiating salvage treatments at different time points and its effect on the predicted dynamic probabilities of recurrence. Pauler & Finkelstein use a change-point model to capture any jump in PSA.

3.1.1 Model specification

In Taylor et al. [57], the functional form over time of the longitudinal PSA mixed model assumes three phases: baseline/presenting PSA (B_0), and the short-term (decrease, B_1), and long-term (increase, B_2) evolutions of PSA, $Y_i(t) = \log[\text{PSA}_i(t) + 0.1] = B_0 + B_1 f_1 + B_2 f_2$, with $f_1 = \{(1 + \text{time})^{-\frac{3}{2}} - 1\}$ and $f_2 = \text{time}$. For each of the three phases, $B_{k=\{0,1,2\}}$, are matrices containing linear combinations of the fixed baseline covariates T-stage, Gleason grade and presenting pre-treatment PSA, along with subject-specific random effects parameters. A t-distribution with five degrees-of-freedom for the error term is assumed. Time to prostate cancer clinical recurrence is measured from the end of RT. In the survival submodel, the functional form $f(\mathbf{M}_i(t), \mathbf{b}_i, \boldsymbol{\alpha})$ is a linear combination of the value of PSA concentration and its slope at time t, $f(\mathbf{M}_i(t), \mathbf{b}_i, \boldsymbol{\alpha}) = \alpha_1 \text{PSA}(t) + \alpha_2 \frac{d \text{PSA}(t)}{dt}$. Additionally, the survival submodel included a time-dependent indicator variable for when salvage hormonal treatment (ST) is initiated to account for the subsequent drop in hazard of clinical failure. PSA values after ST were excluded due to the sudden decrease in PSA trajectory and did not feature in the mixed-effect model; however, clinical recurrences after ST were considered. A piecewise constant function is assumed for the baseline hazard.

Sene et al. [58] made similar modelling assumptions as Taylor et al. [57] for the functional forms in the mixed and survival submodels. The model adjusted for presenting PSA, Gleason score, T-stage, and total corrected dose of EBRT (using the linear-quadratic model given in [60]). Initiation of ST was included as a time-dependent indicator variable to reflect the potential decrease in risk of progression; five functional forms of ST were considered. Three different association structures of f were fitted: a linear combination of PSA value and gradient (with and without a logistic transformation for PSA), and the random effect structure, which evaluated the individual deviations from the overall population's PSA trajectories. A combination of those different parametrisations yielded 12 models with varied complexity.

Pauler & Finkelstein [59] proposed a change-point parameterisations in the longitudinal model for PSA, by incorporating an unknown change-point indicator

variable for whether change in PSA has occurred. If a shift is indicated, a likely change-point time-range is estimated with a uniform distribution for PSA. A narrower posterior change-point range with larger differences in the slopes before- and after the change-point indicate prostate cancer recurrence is likely (before the formal clinical failure endpoint). Trivariate normal and uniform priors are used for four random effects, they included: intercept; change-point time (uniform); the slope before and after the change-point. For the survival submodel, a piecewise exponential hazard function was used. Baseline covariates included age, presenting PSA, and disease stage. For the joint model, non-informative priors were chosen.

3.1.2 Estimation, prediction and validation

In Taylor et al [57], estimation was undertaken under a Bayesian framework using C software. The joint model was developed on three pooled cohorts (totalling $N = 2,386$ patients) and externally tested using a separate fourth dataset ($N = 846$ patients). Dynamic predictions for an individual patient's PSA trajectory and risk of recurrence for the next three years were shown: no formal validation measures were presented. The authors opted for simpler graphical inspections to study the model, owing to the complicated nature of the time-dependent ST events within the validation cohort. An online prognostic calculator was developed, enabling individual dynamic predictions of disease recurrence given PSA trajectories for future patients (<http://psacalc.sph.umich.edu>¹).

In Sene et al [58], estimation was undertaken under a frequentist framework, and R software used for model development, again using the same three cohorts as in Taylor et al [57]. Internal approximated leave-one-out cross-validation was used to assess six of the 12 models' predictability, using BS and EPOCE accuracy measures [54]. The two best fitting models were the logistic-transformed PSA value and slope that separated the effect of PSA before- and after ST, whilst the model with the random effect association structure performed best when assumed that the patient would not start ST within three years. Exemplar individualised dynamic predictions used a prediction window of three years on an intermediate risk patient. Different scenarios when ST would be initiated were used to illustrate the impact of delays in ST initiation on risk of recurrence. External validation was not performed.

We cannot make direct comparisons between the predictive performances of the two papers as they used different assessment methods (graphical approaches in Taylor, EPOCE & BS presented in Sene). In Sene et al., patients who did not receive HT nor ST within three years were mainly used in order to assess predictive performance. Sene noted that this may not be a representative situation for all patients, so they performed a sensitivity analysis using Taylor's approach to widen the sample on HT-free patients at the landmark prediction time only, then with subsequent ST initiation within the three-year prediction window, as either a recurrence event or dependent

¹ Last accessed in January 2022.

censoring. The relative predictive performance was largely unchanged in both papers under this approach and therefore can be considered robust.

In Pauler & Finkelstein [59], estimation was done in a Bayesian framework, using C and S-plus software. The joint model was developed on a cohort of $N = 676$ patients. As the majority of patients do not exhibit clinical failure, the slope after the change-point was non-significantly negative, indicating PSA trajectories generally remain constant over the follow-up period. The regression coefficients from the relative risk component are not straightforward to interpret due to the number of pairwise and three-way interactions, the authors' noted that coefficients are in the expected directions. Sensitivity analysis was done on three differing definitions of recurrence based on PSA rises. They showed that regardless of rule followed, there was little difference to their optimal joint model. The AIC rose when considering only a relative risk model using indicator covariates for each rule, this provided justification on using the joint change-point model, as the longitudinal PSAs substantively improve the goodness-of-fit. The posterior distributions of four individual patient change-points were shown. For two patients who do relapse, sharp change-points are given between 2-4 years, who then go on to recur at six and four years of follow-up. For stable PSA patients, the change-point is imprecise with very wide uniform posteriors. Individualised predictions are performed on two hypothetical patients showing each's posterior predictive distributions of time to relapse. Although discussed, the model was not validated.

3.2 Latent class joint models to predict recurrence in localised prostate cancer

In this section, we focus our review on relevant papers that investigated PSA dynamics using the JLCM framework. There are three papers of interest reviewed in this section, by Proust-Lima & Taylor [61], Jacqmin-Gadda et al. [41], and a third paper by Proust-Lima et al. [43], which is appraised separately in 3.2.1 as compares SPJM and JLCM.

Proust-Lima & Taylor [61] modelled the functional longitudinal PSA similarly to Taylor et al. [57] (described in 3.1). Baseline covariates T-stage, Gleason score, and pre-treatment PSA were included into both submodels. The survival submodel also includes an exogenous time-dependent indicator variable for initiation of ST, and a class-specific Weibull baseline hazard function.

Model development was performed on a single cohort of patients ($N = 1,268$), and external validation was performed on two additional smaller cohorts (with $N = 503$ and $N = 615$ patients respectively). Several JLCMs were fitted with ranging classes (2—6), with the five-class model (5-JLCM) producing the lowest Bayesian Information Criterion (BIC); the optimal model included estimation of 75 parameters. Predicted PSA evolutions and survival curves for each of the five classes illustrate how PSA trajectories with long-term rise of PSA correspond to greater risk of failure. Dynamic predictions were made within a prediction window of three years for two patients with contrasting baseline risk factors: a lower-risk patient who recurs and a higher-risk patient with no observed recurrence.

Within each external validation cohort, measures of predictive accuracy (absolute prediction errors EP and WAEP) for the five-class JLCM were computed, and compared to a relative risk model with baseline information only, and a two-stage landmark model. The JLCM was shown to be the best fitting at various landmark times, and accounting for the longitudinal biomarker reduced both the EP and WAEP, particularly at earlier landmarks.

For Jacqmin-Gadda et al. [41], the score test methodology introduced in 2.4 is applied to develop a prognostic joint model for prostate cancer recurrence (with the same dataset used as in Taylor et al., [62]). They develop the JLCM similarly to Proust-Lima et al., [35,61]. They show that the more flexible 4-class JLCM did not reject conditional independence, whereas the less powerful alternative Wald test for dependence failed to reject the null for a 3-class JLCM.

3.2.1 Comparison between latent-class and shared-parameter joint models

A direct comparison is made between the two types of joint models applied to prostate cancer by Proust-Lima and colleagues [43]. Three prognostic baseline factors were adjusted for, logged initial-PSA, T-stage, and Gleason score using the same Michigan hospital cohort dataset. The three-component parameterisation of PSA in the mixed-effect model was done in the same manner to Proust-Lima & Taylor, and Taylor et al. [57,61] for both joint models for direct comparison. The developed 4-JLCM adjusting for PSA value and slope was chosen from information criteria and conditional independence being met. The BIC favoured the 4-JLCM compared to the shared-parameter JM.

For direct comparisons between the JLCM and SPJM, evaluation of dynamic predictions (for the entire follow-up) are made using the cross-validated EPOCE framework in the first six years. The 4-JLCM is superior to the SPJM in the first four years on internal validation, and also slightly better in the first three years on external validation.

3.3 Extensions to the shared-parameter joint model

We present some further extensions to the joint model in the following subsections. In particular we comment and review four papers with a cured fraction [62–65]; a competing risk joint model [66], where clinical recurrence is competing with a non-related cancer death; and a multi-state joint model [67], whereby patients can go through a pathway of disease states throughout follow-up.

3.3.1 Joint-Cure models

A natural extension to the SPJM is to incorporate a cure component to the time-to-event submodel, whereby patients are considered to be susceptible to experience the event under study (e.g. recurrence), or, on the contrary, to be cured after initial treatment, and thus never susceptible of recurrence. Allocation into the two groups is typically modelled using a logistic classifier submodel:

$$\Pr(D = 1 | X_i) = \frac{\exp(\boldsymbol{\beta}^T X_i)}{1 + \exp(\boldsymbol{\beta}^T X_i)},$$

where $D = 1$ refers to the susceptible group (observed only when the event of interest occurs), X_i is the fixed baseline design matrix with their corresponding vector of coefficients, $\boldsymbol{\beta}$. Patients that have been allocated to the ‘cured’ group are coded $D = 0$.

There can be a high proportion of patients that are recurrence-free after long follow-up, resulting in heavy censoring. This may compromise the predictions of a joint model given the lack of events observed. It therefore may be appropriate to model these patients that appear to have prolonged event-free survival as ‘cured’, using a cure joint model.

There are four articles that consider a joint cure model for the risk of clinical recurrence [62–65]. The four papers have a similar model specification: a nonlinear parametric exponential decay-growth (U-shaped) model is used to capture the log PSA trajectory $m_i(t, \mathbf{r}_i) = r_{i1} \exp(-tr_{i2}) + r_{i3} \exp(tr_{i4})$, where $r_{i1, \dots, 4}$ are the random effects to be estimated. Those that have been allocated to the cure group (from the logistic incidence submodel) have $r_{i4} | (D = 0) = 0$, as this assumption reduces the PSA trajectory, $m_i(t, \mathbf{r}_i)$, to an exponential decay cure SPJM. The conditional failure time model is given by $h(t | D_i, X_i, \mathbf{r}_i, \boldsymbol{\beta}, \boldsymbol{\alpha}, g(m_i)) = h_0(t) \exp(\boldsymbol{\beta}^T X_i + \boldsymbol{\alpha} g(m_{i|D}, t))$, where $g(m_i)$ can be given by including the trajectory function and its slope given in [62,65].

Baseline covariates included pre-treatment PSA, T-stage, and Gleason score. Additionally Taylor et al. [62] considered PSA value & slope as time-dependent covariates, age, EBRT total delivered dose (in Gy) and treatment duration as baseline covariates. Yu and colleagues [65] included an exogenous time-dependent variable to indicate start of salvage HT, similarly to [57], and used a generalised Weibull model for the baseline hazard function. Both frequentist and Bayesian approaches are directly compared by Yu et al. [64].

In Law et al. [63], the joint cure model is compared to the standard cure model without longitudinal time-dependent information, and to the shared parameter joint model without the cure component. They showed better predictions and discrimination, together with reducing biases from informative censoring. Taylor et al. and Yu et al. [62,65] compared the predictions of the model with updated information on the same patients who were initially used to develop the model, that is whereby more longitudinal PSAs and events on the same patients are gathered.

The extended shared-parameter joint-cure model offers additional flexibility to model the inherent heterogeneity of patients that go on to have extended event-free survival. Yu et al. directly compared joint models with and without a cure component. They showed a standard JM tends to overestimate the number of clinical events. They compared the two models using the conditional predictive ordinate and BIC, both

favouring the additional cure submodel component, despite an extra 30 parameters needed to be estimated [65]. This however may over-parameterise the model without adequate event sizes [68]. Also as the prostate cancer disease pathway is complicated, clinical input is recommended with regards to plausibility of the cure component and its definition.

3.3.2 Competing risks joint models

The event of interest may be precluded by the occurrence of a competing event, for instance, non-cancer related deaths in the case of recurrence. It is well known that biases are elicited by censoring these competing event deaths [69,70]; joint models can be extended considering the presence of a competing event.

Ferrer et al. [66] perform individual dynamic predictions and validate the robustness of the estimators in the presence of competing risk of death (from a non-related cancer cause), within a frequentist framework. A cause-specific proportional hazards submodel is proposed for each competing event, and thus the relationship of the longitudinal biomarker with each competing event can be assessed. Individual dynamic predictions were estimated and compared to landmarking estimators. Two simulation studies were performed using simulated data that was alike to the applied prostate cancer dataset. Each approach validated the estimators, then compared and assess their robustness to misspecification of the joint model. Both the AUC and mean-squared prediction error were employed to characterise the predictive accuracy. An extension of the AUC was adapted to the competing risk setting, proposed by Blanche et al. [51]. It was shown that in almost all cases, the joint models were superior to the landmark models. The landmark models were only superior to the joint models when the longitudinal biomarker was heavily misspecified. Ferrer's competing risk paper is the only study to present validation metrics, using simulated studies. Code is available at <https://github.com/LoicFerrer/Individual-dynamic-predictions>².

3.3.3 Multi-state joint models

The evolution of localised prostate cancer over time can be characterised by the occurrence of different events of interest, such as biochemical failure, local recurrence, distant recurrence and death. One way to jointly model all these events is via multi-state models, in which the event progressions of interest define the transition between different disease states [71]. As longitudinal process such as PSA trajectories can have an impact on several of these event transitions, multi-state models can be generalised to the joint modelling framework.

Ferrer et al. [67] proposed modelling the longitudinal PSA process using a mixed-effect submodel, similarly to Proust-Lima and Taylor [61], Sene et al. [58], and Ferrer et al. [66]. They used a non-homogeneous Markov multi-state model for the intensity of the transitions between five states: 0) end of EBRT treatment, 1) local recurrence, 2) salvage HT, 3) distant recurrence, and 4) death (the absorbing state). Intermediate

² Last accessed in January 2022.

states could be skipped (e.g. ending EBRT₀ → death₄), and backward transitions were not allowed. Two properties were considered: 1) the Markov property whereby the future process is only dependent on the present state and not the preceding transitions / states; 2) the non-homogeneous property ensures the time since entering the study influences the evolution of the process.

Each transition intensity is modelled assuming proportional hazards and incorporated the biomarker trajectory. For each transition from state i to j , only patients visiting the state i are included in the analysis. The baseline intensity function was modelled parametrically. The maximum likelihood framework was used to estimate the corresponding parameters.

The multi-state joint model was fitted with the same two study datasets as in Ferrer et al. [66]. Four covariates (presenting PSA, Gleason score, T-stage, and study cohort) were adjusted for in the models. Worse baseline risk factors were associated with higher values in their long-term PSA trajectories, and reaching clinical failure states earlier. Higher presenting PSA was associated with a higher instantaneous risk to salvage hormone therapy or death. A linear combination of increases in PSA value & slope were associated with significant deterioration for all clinical progression transitions (local, salvage HT, or distant failures) from the initial state.

After adjusting for all other covariates and PSA slope, a unit increase in the log PSA gave rise to a 43% increase in risk to local recurrence. Patients with continually high PSAs or increasingly steep PSA gradients after EBRT treatment, led to earlier and higher hazards to clinical failure states. Conversely, higher PSA levels had a protective effect on the transition to direct death after EBRT but were more likely to progress through the prostate cancer progression states.

Predictions were compared with the observed data. The observed values were averaged at each decile with corresponding predicted values computed, they show the observed values lay within the 95% CIs, with very similar predicted values. The predicted transition probabilities over time, in a given state to another other feasible state are presented, comparing similar parametric estimated probabilities to the observed. The only exception was between transitions 1→2 (from local recurrence to receiving HT) where the spike after EBRT was not adequately captured with the splines, it shows there is a very near-immediate initiation of HT after localised recurrence to control the disease. It is worth noting that PSA dynamics were only collected until the patient's first clinical event and not thereafter and were extrapolated according to their posterior trajectories.

Diagnostics of the joint multi-state model were evaluated visually. Residuals vs fitted values, observed and predicted PSA trajectories, and predicted vs non-parametric transition probabilities between states were presented. In general, they showed the model fits particularly well to the longitudinal, and multi-state submodels. The models themselves were not externally validated nor stated any predictive performance

measures, only the estimation process via simulation studies. Although equations for obtaining individual dynamic predictions for patients were presented in the paper, these were not demonstrated with specific examples.

The code to apply these multistate models to a simulated dataset and adapt for use is freely available at <https://github.com/LoicFerrer/JMstateModel>³ and could be used to derive patient predictions and be adapted for the reader's need.

4 Discussion

Over the last two decades there has been a plethora of research on PSA protein concentration and its association to recurrence, or prolonged event-free survival (effectively cure). We have reviewed and assessed 12 papers that report joint models of longitudinal PSA trajectories and time-to-event endpoints that aim to describe how these trajectories impact and predict prostate cancer recurrence. We found two broad frameworks (SPJMs & JLCMs) that were utilised and assessed different methodologies. We synthesize these different approaches applied to similar dataset cohorts of prostate cancer patients receiving EBRT without HT, which allow the methodology to be compared. Due to the long-term nature of prostate cancer recurrence and progression, the datasets to develop the DPMs comprise patients treated in the 1980s. As long-term follow-up is necessary the historical nature of the datasets is unavoidable but the impact of changes in clinical practice should be considered when utilising DPMs for contemporary patients.

There are limitations to our work, as this report was not initially intended to be a systematic review on all the available literature, but a synthesised summary of what we considered relevant articles of modelling both PSA longitudinally, and time-to-recurrence in localised prostate cancer; in preparation for an application for these methods in our own dataset (in a publication to follow). For instance, we focused on specific key words within the title and abstract only, so we may have missed reports if the use of these terms was not explicit in these fields. Further joint modelling papers not included here were due to, for instance, no dynamic predictions presented [72], a mix of non-radiotherapy treatments (e.g, radical prostatectomy); methodology development focused but repeated analysis referred to [53,73,74]; or exclusive use of simulated datasets [75]. It was noted that not all papers were expectedly populated by the search strategy [76]. In the localised prostate cancer setting, where PSA is used to monitor recurrence after radical treatment of disease, joint models have also been used in the context of prostate cancer screening [77–82] or advanced (metastatic) disease [33,83–87]. We did not consider these scenarios as the PSA dynamics differ greatly. These models could also be extended to accommodate more than one longitudinal biomarker, such as PSA and testosterone, or the sequential findings on digital rectal exams, in a joint multivariate model. Regardless of disease stage, these papers highlight and emphasise the use of longitudinal information, such as the PSA biomarker. This increases the prognostic power of the prediction model to help inform

³ Last accessed in January 2022.

and update predictions of the event of interest, compared to solely using baseline risk factors that are imprecise [88].

Modern typical first-line treatment of localised prostate cancer include HT before (neoadjuvant) and concurrently with external-beam radiotherapy [12,89], and PSA trajectories are known to be more homogeneous with combined treatment [57]. Furthermore, given recent advances in radiotherapy techniques and the use of moderate- and ultra-hypofractionation (fewer but larger doses of radiotherapy) [90,91], treatment exposures of RT are less than the average treatment durations presented in these papers. The tool in Taylor et al. [57], reviewed in 3.1 was developed in the absence of neoadjuvant HT, therefore predictions from these models have limited applicability within current treatment pathways. Further development of these models for patients receiving HT are needed.

The papers reviewed provide a very good exposition and rationale to their model development and clinical usage. Regardless of the functional form used in the joint modelling framework, a fully parametric form was fitted for the mixed-effects model. There are possibly more appropriate and flexible forms that may exist, compared to the biphasic form for PSA trajectories they postulate throughout [35,41,43,57,58,61,66]. Many of the reviewed articles present an appraisal of their models, either by validation or contain a simulation study. External validation is seen as the gold-standard, to ensure model suitability and generalisability in other patient populations and to assess overfitting [92]. However, when rigorous measures of predictive performance have not been reported in these papers, these would not be considered validated by today's standards [93].

As with any specification of modelling, there are pros and cons to the joint modelling approach taken and several differences exist. For JLCMs, the maximum likelihood approach contains closed-form solutions and are computationally feasible to compute. They are advantageous for the use of developing a predictive joint model for dynamic predictions, whilst not having to impose specific parametric assumptions for the biomarker's functional form (e.g. current value, slope, area), unlike SPJMs [43]. Robustness to deviations of the imposed functional form have been rigorously assessed in Ferrer et al. [66]. In this paper, they demonstrated that no method (joint modelling nor landmarking) was particularly robust to misspecification in the longitudinal biomarker. However, when there was heavy misspecification, landmarking methods did perform better than joint modelling.

The SPJMs assume a homogenous population with a singular average PSA biomarker trajectory, whereas JLCMs account for further population heterogeneity through the latent classes. Both JLCMs and SPJMs account for the variability of the PSA biomarker through the random effects in the longitudinal model. The purpose of the random effects in the SPJM is two-fold, accounting for the correlation of the repeated measures in the mixed-effect model, and the association between the PSA biomarker

and time-to-recurrence, whilst in the JLCM only the latent classes account of the association between the biomarker and event.

Disadvantages of the JLCMs approach include the possibility of having multiple local maxima for the maximum likelihood estimates, and several models are needed to be fitted in order to find the optimal number of latent classes (by comparing multiple information criteria) [37]. Some of these issues can be circumnavigated via parallelisation of the computation for more optimal resourcing, e.g., make use of parallel computing by using search grid methods for JLCMs as computations are independent (see the *mpjlcm* function from R package *lcmm*); or implementing multiple MCMC chains performed in parallel using Bayesian SPJMs (*jm* function from R package *JMbayes2*) [55,94].

Both Frequentist and Bayesian paradigms were used for the SPJMs, whereas we only reviewed frequentist methods for the JLCMs. In their direct comparison of JLCMs and SPJMs [43], the authors showed that the JLCMs had less assumptions and performed better. However when adjusting for the same patient cohort dataset, baseline covariates, prediction times, and biphasic components for the longitudinal PSA component: the prognostic accuracy measures for EPOCE in Sene et al. [58] using SPJMs appear superior than those obtained with the JLCM in Proust-Lima et al. [43].

All models reviewed in this paper can produce dynamic predictions for prostate cancer prognosis. The JLCMs do not assume a specific association structure nor quantify those associations, (like the SPJMs do), they describe the trajectories in a heterogeneous population. If the main goal is to quantify the associations assuming a homogenous population, then SPJMs are recommended. There is not one overarching or standout model to always use by default. The choice of model may be primarily driven by the research question and personal choice. If the purpose is solely for prediction, then combining several frameworks for dynamic predictions using some weighted model averaging methodology could be applied [95]. Indeed one type of framework may outperform another at certain time intervals and then vice-versa at different time windows. Each model has its own advantages, depending on the end goal of the reader. It is hard to compare each model's framework with another in terms of superior predictive performance as not all these papers present these metrics.

This review focused on radiotherapy, however there are other treatments for prostate cancer including hormone therapy, prostatectomy and combinations therein, though optimal timing of these combinational therapies appears unclear [96,97]. There have been recent advances in using sophisticated machine learning/artificial intelligence (ML/AI) techniques on imaging data to predict whether patients require biopsies, or to predict clinical failure or death under these alternative treatment pathways. Some recent articles include development of artificial neural networks, support vector machines, and random forests for predicting diagnoses [98,99], optimal timing of biopsies [100], and clinical failure [101] or death [102]. However, it is not apparent that the longitudinal nature of time-varying markers like PSA have been considered,

nor produce dynamic predictions. A review of these AI and ML methods is given in Tătaru et al [103]. Some authors refer to joint modelling itself is an AI approach [104]. Other studies have suggested combining the boosting approaches of machine learning to joint models, to create a unified framework using mechanistic data-driven approaches [105]. ML/AI techniques are not a panacea and need to be correctly developed and incorporate all available information, be rigorously validated, and to have clinical utility [106–109]. Reporting guidance, based on TRIPOD & PROBAST statements, have been developed for AI & ML (TRIPOD-AI/ML & PROBAST-AI/ML) [110–113].

5 Conclusions

To conclude, we reviewed, summarised, and synthesised principal methodologies on twelve seminal papers over the last two decades on dynamic prediction joint models applied to the prognosis of prostate cancer patients, using PSA to dynamically update prognosis. This article supports the use of utilising longitudinally collected PSA, in addition to baseline prognostic factors to improve predictions in a joint modelling framework. There have been many advancements in computational processing, methodologies, with improvements in clinical practice and treatments. Combining all these developments together with utilising all available information, the future of dynamic prediction models is encouraging.

List of abbreviations

- AI – artificial intelligence
- AIC – Akaike information criterion
- AUROC (or AUC) – area under the receiver operating characteristic curve
- BcF – biochemical failure
- BIC – Bayesian information criterion
- (I)BS – (Integrated) Brier score
- CI – confidence interval (frequentist) *or* credible interval (Bayesian)
- (CV)POL – (cross-validated) prognostic observed log-likelihood
- CPM – clinical prediction model
- DPM – dynamic prediction model
- EBRT – external beam radiotherapy
- EP – prediction error (or error of prediction)
- EPOCE – expected prognostic observed cross-entropy
- HT – hormonal therapy
- JLCM – joint latent class model
- JM – joint model
- ML – Machine Learning
- NCCN – National Comprehensive Cancer Network
- PSA – prostate-specific antigen
- SPJM – shared-parameter joint model
- ST – salvage therapy
- WAEP – weighted average absolute error of prediction

Declarations

Ethics approval and consent to participate

Not applicable

Consent for publication

Not applicable

Availability of data and materials

Data sharing is not applicable to this article as no datasets were generated or analysed during the current study.

Competing interests

The authors declare that they have no competing interests.

Funding

Not applicable

Authors' contributions

HP carried out the review and wrote the manuscript. EH & NP reviewed previous drafts of the manuscript, offering critical appraisal and comment. The final manuscript was approved by all authors.

Acknowledgements

Not applicable

References

- [1] P. Rawla, Epidemiology of Prostate Cancer, *World J Oncol.* 10 (2019) 63–89. <https://doi.org/10.14740/wjon1191>.
- [2] Cancer registration statistics, England: first release, 2018, GOV.UK. (n.d.). <https://www.gov.uk/government/publications/cancer-registration-statistics-england-2018/cancer-registration-statistics-england-first-release-2018> (accessed February 24, 2020).
- [3] C.E. Lovegrove, O. Musbahi, N. Ranasinha, A. Omer, F. Lopez, A. Campbell, R.J. Bryant, T. Leslie, R. Bell, S. Brewster, F.C. Hamdy, B. Wright, A.D. Lamb, Implications of celebrity endorsement of prostate cancer awareness in a tertiary referral unit - the 'Fry-Turnbull' effect, *BJU Int.* 125 (2020) 484–486. <https://doi.org/10.1111/bju.14992>.
- [4] About prostate cancer, Prostate Cancer UK. (n.d.). <https://prostatecanceruk.org/prostate-information/about-prostate-cancer> (accessed August 6, 2019).
- [5] Are you at risk?, Prostate Cancer UK. (n.d.). <https://prostatecanceruk.org/prostate-information/are-you-at-risk> (accessed August 6, 2019).
- [6] Prostate cancer risk, Cancer Research UK. (2015). <https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/prostate-cancer/risk-factors> (accessed August 6, 2019).
- [7] S. McPhail, S. Johnson, D. Greenberg, M. Peake, B. Rous, Stage at diagnosis and early mortality from cancer in England, *British Journal of Cancer.* 112 (2015) S108–S115. <https://doi.org/10.1038/bjc.2015.49>.
- [8] Survival by stage, (n.d.). http://www.ncin.org.uk/publications/survival_by_stage (accessed November 23, 2020).
- [9] Recommendations | Prostate cancer: diagnosis and management | Guidance | NICE, (n.d.). <https://www.nice.org.uk/guidance/ng131/chapter/Recommendations> (accessed August 6, 2019).
- [10] F.C. Hamdy, J.L. Donovan, J.A. Lane, M. Mason, C. Metcalfe, P. Holding, M. Davis, T.J. Peters, E.L. Turner, R.M. Martin, J. Oxley, M. Robinson, J. Staffurth, E. Walsh, P. Bollina, J. Catto, A. Doble, A. Doherty, D. Gillatt, R. Kockelbergh, H. Kynaston, A. Paul, P. Powell, S. Prescott, D.J. Rosario, E. Rowe, D.E. Neal, 10-Year Outcomes after Monitoring, Surgery, or Radiotherapy for Localized Prostate Cancer, <https://doi.org/10.1056/NEJMoa1606220>. (2016). <https://doi.org/10.1056/NEJMoa1606220>.
- [11] S.C. Morgan, K. Hoffman, D.A. Loblaw, M.K. Buyyounouski, C. Patton, D. Barocas, S. Bentzen, M. Chang, J. Efstathiou, P. Greany, P. Halvorsen, B.F. Koontz, C. Lawton, C.M. Leyrer, D. Lin, M. Ray, H. Sandler, Hypofractionated Radiation Therapy for Localized Prostate Cancer: An ASTRO, ASCO, and AUA Evidence-Based Guideline, *JCO.* 36 (2018) 3411–3430. <https://doi.org/10.1200/JCO.18.01097>.
- [12] J.W. Denham, D. Joseph, D.S. Lamb, N.A. Spry, G. Duchesne, J. Matthews, C. Atkinson, K.-H. Tai, D. Christie, L. Kenny, S. Turner, N.K. Gogna, T. Diamond, B. Delahunt, C. Oldmeadow, J. Attia, A. Steigler, Short-term androgen suppression and radiotherapy versus intermediate-term androgen suppression and radiotherapy, with or without zoledronic acid, in men with locally advanced prostate cancer (TROG 03.04 RADAR): 10-year results from a randomised,

- phase 3, factorial trial, *The Lancet Oncology*. 20 (2019) 267–281.
[https://doi.org/10.1016/S1470-2045\(18\)30757-5](https://doi.org/10.1016/S1470-2045(18)30757-5).
- [13] A.R. Rao, H.G. Motiwala, O.M.A. Karim, The discovery of prostate-specific antigen, *BJU International*. 101 (2008) 5–10. <https://doi.org/10.1111/j.1464-410X.2007.07138.x>.
- [14] M. Roach, G. Hanks, H. Thames, P. Schellhammer, W.U. Shipley, G.H. Sokol, H. Sandler, Defining biochemical failure following radiotherapy with or without hormonal therapy in men with clinically localized prostate cancer: Recommendations of the RTOG-ASTRO Phoenix Consensus Conference, *International Journal of Radiation Oncology*Biophysics*. 65 (2006) 965–974. <https://doi.org/10.1016/j.ijrobp.2006.04.029>.
- [15] M.B. Amin, S. Edge, F. Greene, D.R. Byrd, R.K. Brookland, M.K. Washington, J.E. Gershenwald, C.C. Compton, K.R. Hess, D.C. Sullivan, J.M. Jessup, J.D. Brierley, L.E. Gaspar, R.L. Schilsky, C.M. Balch, D.P. Winchester, E.A. Asare, M. Madera, D.M. Gress, L.R. Meyer, eds., *AJCC Cancer Staging Manual*, 8th ed., Springer International Publishing, 2017.
<https://www.springer.com/gp/book/9783319406176> (accessed November 23, 2020).
- [16] J.L. Mohler, E.S. Antonarakis, A.J. Armstrong, A.V. D’Amico, B.J. Davis, T. Dorff, J.A. Eastham, C.A. Enke, T.A. Farrington, C.S. Higano, E.M. Horwitz, M. Hurwitz, J.E. Ippolito, C.J. Kane, M.R. Kuettel, J.M. Lang, J. McKenney, G. Netto, D.F. Penson, E.R. Plimack, J.M. Pow-Sang, T.J. Pugh, S. Richey, M. Roach, S. Rosenfeld, E. Schaeffer, A. Shabsigh, E.J. Small, D.E. Spratt, S. Srinivas, J. Tward, D.A. Shead, D.A. Freedman-Cass, Prostate Cancer, Version 2.2019, *NCCN Clinical Practice Guidelines in Oncology*, *Journal of the National Comprehensive Cancer Network*. 17 (2019) 479–505.
<https://doi.org/10.6004/jnccn.2019.0023>.
- [17] W. Xie, M.M. Regan, M. Buyse, S. Halabi, P.W. Kantoff, O. Sartor, H. Soule, N.W. Clarke, L. Collette, J.J. Dignam, K. Fizazi, W.R. Paruleker, H.M. Sandler, M.R. Sydes, B. Tombal, S.G. Williams, C.J. Sweeney, Metastasis-Free Survival Is a Strong Surrogate of Overall Survival in Localized Prostate Cancer, *JCO*. 35 (2017) 3097–3104. <https://doi.org/10.1200/JCO.2017.73.9987>.
- [18] S.F. Shariat, P.I. Karakiewicz, C.G. Roehrborn, M.W. Kattan, An updated catalog of prostate cancer predictive tools, *Cancer*. 113 (2008) 3075–3099.
<https://doi.org/10.1002/cncr.23908>.
- [19] S.F. Shariat, K. Mw, V. Aj, K. Pi, S. Pt, Critical review of prostate cancer predictive tools., *Future Oncol*. 5 (2009) 1555–1584.
<https://doi.org/10.2217/fon.09.121>.
- [20] S.F. Shariat, P.I. Karakiewicz, V. Margulis, M.W. Kattan, Inventory of prostate cancer predictive tools, *Current Opinion in Urology*. 18 (2008) 279–296.
<https://doi.org/10.1097/MOU.0b013e3282f9b3e5>.
- [21] M. Aladwani, A. Lophatananon, W. Ollier, K. Muir, Prediction models for prostate cancer to be used in the primary care setting: a systematic review, *BMJ Open*. 10 (2020) e034661. <https://doi.org/10.1136/bmjopen-2019-034661>.
- [22] D.R. Thurtle, D.C. Greenberg, L.S. Lee, H.H. Huang, P.D. Pharoah, V.J. Gnanapragasam, Individual prognosis at diagnosis in nonmetastatic prostate cancer: Development and external validation of the PREDICT Prostate multivariable model, *PLOS Medicine*. 16 (2019) e1002758.
<https://doi.org/10.1371/journal.pmed.1002758>.

- [23] J.A. Brockman, S. Alanee, A.J. Vickers, P.T. Scardino, D.P. Wood, A.S. Kibel, D.W. Lin, F.J. Bianco, D.M. Rabah, E.A. Klein, J.P. Ciezki, T. Gao, M.W. Kattan, A.J. Stephenson, Nomogram Predicting Prostate Cancer-specific Mortality for Men with Biochemical Recurrence After Radical Prostatectomy, *Eur. Urol.* 67 (2015) 1160–1167. <https://doi.org/10.1016/j.eururo.2014.09.019>.
- [24] D.P. Ankerst, J. Straubinger, K. Selig, L. Guerrios, A. De Hoedt, J. Hernandez, M.A. Liss, R.J. Leach, S.J. Freedland, M.W. Kattan, R. Nam, A. Haese, F. Montorsi, S.A. Boorjian, M.R. Cooperberg, C. Poyet, E. Vertosick, A.J. Vickers, A Contemporary Prostate Biopsy Risk Calculator Based on Multiple Heterogeneous Cohorts, *Eur. Urol.* 74 (2018) 197–203. <https://doi.org/10.1016/j.eururo.2018.05.003>.
- [25] C.R. Pound, A.W. Partin, M.A. Eisenberger, D.W. Chan, J.D. Pearson, P.C. Walsh, Natural history of progression after PSA elevation following radical prostatectomy, *JAMA.* 281 (1999) 1591–1597. <https://doi.org/10.1001/jama.281.17.1591>.
- [26] Prostate Cancer Nomograms, Memorial Sloan Kettering Cancer Center. (n.d.). <https://www.mskcc.org/nomograms/prostate> (accessed September 16, 2019).
- [27] T.M. Therneau, P.M. Grambsch, Modeling Survival Data: Extending the Cox Model, Springer Science & Business Media, 2000.
- [28] D. Rizopoulos, G. Molenberghs, E.M.E.H. Lesaffre, Dynamic predictions with time-dependent covariates in survival analysis using joint modeling and landmarking, *Biometrical Journal.* 59 (2017) 1261–1276. <https://doi.org/10.1002/bimj.201600238>.
- [29] G. Papageorgiou, K. Mauff, A. Tomer, D. Rizopoulos, An Overview of Joint Modeling of Time-to-Event and Longitudinal Outcomes, *Annual Review of Statistics and Its Application.* 6 (2019). <https://doi.org/10.1146/annurev-statistics-030718-105048>.
- [30] J.R. Anderson, K.C. Cain, R.D. Gelber, Analysis of survival by tumor response., *JCO.* 1 (1983) 710–719. <https://doi.org/10.1200/JCO.1983.1.11.710>.
- [31] H.C.V. Houwelingen, Dynamic Prediction by Landmarking in Event History Analysis, *Scandinavian Journal of Statistics.* 34 (2007) 70–85. <https://doi.org/10.1111/j.1467-9469.2006.00529.x>.
- [32] H. van Houwelingen, H. Putter, Dynamic Prediction in Clinical Survival Analysis, CRC Press, 2011.
- [33] K. Suresh, J.M.G. Taylor, D.E. Spratt, S. Daignault, A. Tsodikov, Comparison of joint modeling and landmarking for dynamic prediction under an illness-death model, *Biom J.* 59 (2017) 1277–1300. <https://doi.org/10.1002/bimj.201600235>.
- [34] H. Putter, H.C. van Houwelingen, Landmarking 2.0: Bridging the gap between joint models and landmarking, *Statistics in Medicine.* n/a (2022). <https://doi.org/10.1002/sim.9336>.
- [35] C. Proust-Lima, J.M.G. Taylor, S.G. Williams, D.P. Ankerst, N. Liu, L.L. Kestin, K. Bae, H.M. Sandler, Determinants of Change in Prostate-Specific Antigen Over Time and Its Association With Recurrence After External Beam Radiation Therapy for Prostate Cancer in Five Large Cohorts, *International Journal of Radiation Oncology*Biophysics*Physics.* 72 (2008) 782–791. <https://doi.org/10.1016/j.ijrobp.2008.01.056>.
- [36] M.J. Sweeting, S.G. Thompson, Joint modelling of longitudinal and time-to-event data with application to predicting abdominal aortic aneurysm growth and rupture, *Biom J.* 53 (2011) 750–763. <https://doi.org/10.1002/bimj.201100052>.

- [37] D. Rizopoulos, *Joint Models for Longitudinal and Time-to-Event Data: With Applications in R*, Chapman and Hall/CRC, Boca Raton, UNITED KINGDOM, 2012. <http://ebookcentral.proquest.com/lib/icruk/detail.action?docID=952042> (accessed November 7, 2018).
- [38] A.L. Gould, M.E. Boye, M.J. Crowther, J.G. Ibrahim, G. Quartey, S. Micallef, F.Y. Bois, Joint modeling of survival and longitudinal non-survival data: current methods and issues. Report of the DIA Bayesian joint modeling working group, *Stat Med.* 34 (2015) 2181–2195. <https://doi.org/10.1002/sim.6141>.
- [39] E.-R. Andrinopoulou, D. Rizopoulos, J.J.M. Takkenberg, E. Lesaffre, Joint modeling of two longitudinal outcomes and competing risk data, *Statistics in Medicine.* 33 (2014) 3167–3178. <https://doi.org/10.1002/sim.6158>.
- [40] F. Hsieh, Y.-K. Tseng, J.-L. Wang, Joint modeling of survival and longitudinal data: likelihood approach revisited, *Biometrics.* 62 (2006) 1037–1043. <https://doi.org/10.1111/j.1541-0420.2006.00570.x>.
- [41] H. Jacqmin-Gadda, C. Proust-Lima, J.M.G. Taylor, D. Commenges, Score test for conditional independence between longitudinal outcome and time to event given the classes in the joint latent class model, *Biometrics.* 66 (2010) 11–19. <https://doi.org/10.1111/j.1541-0420.2009.01234.x>.
- [42] D. Rizopoulos, Dynamic Predictions and Prospective Accuracy in Joint Models for Longitudinal and Time-to-Event Data, *Biometrics.* 67 (2011) 819–829. <https://doi.org/10.1111/j.1541-0420.2010.01546.x>.
- [43] C. Proust-Lima, M. Séne, J.M. Taylor, H. Jacqmin-Gadda, Joint latent class models for longitudinal and time-to-event data: A review, *Statistical Methods in Medical Research.* 23 (2014) 74. <https://doi.org/10.1177/0962280212445839>.
- [44] Y. Zheng, P.J. Heagerty, Prospective Accuracy for Longitudinal Markers, *Biometrics.* 63 (2007) 332–341. <https://doi.org/10.1111/j.1541-0420.2006.00726.x>.
- [45] P.J. Heagerty, Y. Zheng, Survival Model Predictive Accuracy and ROC Curves, *Biometrics.* 61 (2005) 92–105. <https://doi.org/10.1111/j.0006-341X.2005.030814.x>.
- [46] P. Blanche, J.-F. Dartigues, H. Jacqmin-Gadda, Estimating and comparing time-dependent areas under receiver operating characteristic curves for censored event times with competing risks, *Statistics in Medicine.* 32 (2013) 5381–5397. <https://doi.org/10.1002/sim.5958>.
- [47] T.A. Gerds, M. Schumacher, Consistent Estimation of the Expected Brier Score in General Survival Models with Right-Censored Event Times, *Biom. J.* 48 (2006) 1029–1040. <https://doi.org/10.1002/bimj.200610301>.
- [48] M.-C. Fournier, E. Dantan, P. Blanche, An R2-curve for evaluating the accuracy of dynamic predictions, *Statistics in Medicine.* 37 (2018) 1125–1133. <https://doi.org/10.1002/sim.7571>.
- [49] M. Schemper, R. Henderson, Predictive Accuracy and Explained Variation in Cox Regression, *Biometrics.* 56 (2000) 249–255. <https://doi.org/10.1111/j.0006-341X.2000.00249.x>.
- [50] R. Henderson, P. Diggle, A. Dobson, Identification and efficacy of longitudinal markers for survival, *Biostatistics.* 3 (2002) 33–50. <https://doi.org/10.1093/biostatistics/3.1.33>.
- [51] P. Blanche, C. Proust-Lima, L. Loubère, C. Berr, J.-F. Dartigues, H. Jacqmin-Gadda, Quantifying and comparing dynamic predictive accuracy of joint models for longitudinal marker and time-to-event in presence of censoring and

- competing risks, *Biometrics*. 71 (2015) 102–113.
<https://doi.org/10.1111/biom.12232>.
- [52] R. Schoop, E. Graf, M. Schumacher, Quantifying the Predictive Performance of Prognostic Models for Censored Survival Data with Time-Dependent Covariates, *Biometrics*. 64 (2008) 603–610. <https://doi.org/10.1111/j.1541-0420.2007.00889.x>.
- [53] D. Commenges, B. Lique, C. Proust-Lima, Choice of Prognostic Estimators in Joint Models by Estimating Differences of Expected Conditional Kullback–Leibler Risks, *Biometrics*. 68 (2012) 380–387. <https://doi.org/10.1111/j.1541-0420.2012.01753.x>.
- [54] D. Commenges, C. Proust-Lima, C. Samieri, B. Lique, A Universal Approximate Cross-Validation Criterion for Regular Risk Functions, *The International Journal of Biostatistics*. 11 (2015) 51–67. <https://doi.org/10.1515/ijb-2015-0004>.
- [55] C. Proust-Lima, V. Philipps, B. Lique, Estimation of Extended Mixed Models Using Latent Classes and Latent Processes: The R Package lcomm, *Journal of Statistical Software*. 78 (2017) 1–56. <https://doi.org/10.18637/jss.v078.i02>.
- [56] D. Rizopoulos, JM: An R Package for the Joint Modelling of Longitudinal and Time-to-Event Data, *Journal of Statistical Software*. 35 (2010) 1–33. <https://doi.org/10.18637/jss.v035.i09>.
- [57] J.M.G. Taylor, Y. Park, D.P. Ankerst, C. Proust-Lima, S. Williams, L. Kestin, K. Bae, T. Pickles, H. Sandler, Real-Time Individual Predictions of Prostate Cancer Recurrence Using Joint Models, *Biometrics*. 69 (2013) 206–213. <https://doi.org/10.1111/j.1541-0420.2012.01823.x>.
- [58] M. Sene, J.M.G. Taylor, J.J. Dignam, H. Jacqmin-Gadda, C. Proust-Lima, Individualized dynamic prediction of prostate cancer recurrence with and without the initiation of a second treatment: development and validation, *Stat Methods Med Res*. 25 (2016) 2972–2991. <https://doi.org/10.1177/0962280214535763>.
- [59] D.K. Pauler, D.M. Finkelstein, Predicting time to prostate cancer recurrence based on joint models for non-linear longitudinal biomarkers and event time outcomes, *Statistics in Medicine*. 21 (2002) 3897–3911. <https://doi.org/10.1002/sim.1392>.
- [60] C. Proust-Lima, J.M.G. Taylor, S. Sécher, H. Sandler, L. Kestin, T. Pickles, K. Bae, R. Allison, S. Williams, Confirmation of a Low α/β Ratio for Prostate Cancer Treated by External Beam Radiation Therapy Alone Using a Post-Treatment Repeated-Measures Model for PSA Dynamics, *International Journal of Radiation Oncology*Biology*Physics*. 79 (2011) 195–201. <https://doi.org/10.1016/j.ijrobp.2009.10.008>.
- [61] C. Proust-Lima, J.M.G. Taylor, Development and validation of a dynamic prognostic tool for prostate cancer recurrence using repeated measures of posttreatment PSA: a joint modeling approach, *Biostatistics*. 10 (2009) 535–549. <https://doi.org/10.1093/biostatistics/kxp009>.
- [62] J.M.G. Taylor, M. Yu, H.M. Sandler, Individualized Predictions of Disease Progression Following Radiation Therapy for Prostate Cancer, *JCO*. 23 (2005) 816–825. <https://doi.org/10.1200/JCO.2005.12.156>.
- [63] N.J. Law, J.M.G. Taylor, H. Sandler, The joint modeling of a longitudinal disease progression marker and the failure time process in the presence of cure, *Biostatistics*. 3 (2002) 547–563. <https://doi.org/10.1093/biostatistics/3.4.547>.
- [64] M. Yu, N.J. Law, J.M.G. Taylor, H.M. Sandler, Joint Longitudinal-survival-cure Models and Their Application to Prostate Cancer, *Statistica Sinica*. (2004) 28.

- [65] M. Yu, J.M.G. Taylor, H.M. Sandler, Individual Prediction in Prostate Cancer Studies Using a Joint Longitudinal Survival–Cure Model, *Journal of the American Statistical Association*. 103 (2008) 178–187. <https://doi.org/10.1198/016214507000000400>.
- [66] L. Ferrer, H. Putter, C. Proust-Lima, Individual dynamic predictions using landmarking and joint modelling: Validation of estimators and robustness assessment, *Stat Methods Med Res.* (2018) 0962280218811837. <https://doi.org/10.1177/0962280218811837>.
- [67] L. Ferrer, V. Rondeau, J.J. Dignam, T. Pickles, H. Jacqmin-Gadda, C. Proust-Lima, Joint modelling of longitudinal and multi-state processes: application to clinical progressions in prostate cancer, *Stat Med.* 35 (2016) 3933–3948. <https://doi.org/10.1002/sim.6972>.
- [68] S.D. Collins, N. Peek, R.D. Riley, G.P. Martin, Sample sizes of prediction model studies in prostate cancer were rarely justified and often insufficient, *Journal of Clinical Epidemiology*. 133 (2021) 53–60. <https://doi.org/10.1016/j.jclinepi.2020.12.011>.
- [69] P.K. Andersen, R.B. Geskus, T. de Witte, H. Putter, Competing risks in epidemiology: possibilities and pitfalls, *Int J Epidemiol.* 41 (2012) 861–870. <https://doi.org/10.1093/ije/dyr213>.
- [70] C. van Walraven, F.A. McAlister, Competing risk bias was common in Kaplan–Meier risk estimates published in prominent medical journals, *Journal of Clinical Epidemiology*. 69 (2016) 170-173.e8. <https://doi.org/10.1016/j.jclinepi.2015.07.006>.
- [71] L.C. de Wreede, M. Fiocco, H. Putter, The mstate package for estimation and prediction in non- and semi-parametric multi-state and competing risks models, *Computer Methods and Programs in Biomedicine*. 99 (2010) 261–274. <https://doi.org/10.1016/j.cmpb.2010.01.001>.
- [72] R.A. Mohammadpour, A. Alizadeh, M. Barzegartahamtan, A. Akbarzadeh Pasha, Association between prostate specific antigen change over time and prostate cancer recurrence risk: a joint model, *Caspian Journal of Internal Medicine*. 11 (2020) 324–328. <https://doi.org/10.22088/cjim.11.3.324>.
- [73] X. Lin, J.M.G. Taylor, W. Ye, A penalized likelihood approach to joint modeling of longitudinal measurements and time-to-event data, *Stat. Interface.* 1 (2008) 33–45. <https://doi.org/10.4310/SII.2008.v1.n1.a4>.
- [74] S. Kim, D. Zeng, J.M.G. Taylor, Joint partially linear model for longitudinal data with informative drop-outs, *Biometrics*. 73 (2017) 72–82. <https://doi.org/10.1111/biom.12566>.
- [75] H. Li, C. Gatsonis, Dynamic optimal strategy for monitoring disease recurrence, *Sci. China Math.* 55 (2012) 1565–1582. <https://doi.org/10.1007/s11425-012-4475-y>.
- [76] M. Sène, C.A. Bellera, C. Proust-Lima, Shared random-effect models for the joint analysis of longitudinal and time-to-event data: application to the prediction of prostate cancer recurrence, *Journal de la société française de statistique*. 155 (2014) 134–155.
- [77] A. Tomer, D. Nieboer, M.J. Roobol, A. Bjartell, E.W. Steyerberg, D. Rizopoulos, Movember Foundation’s Global Action Plan Prostate Cancer Active Surveillance (GAP3) consortium, Personalized Biopsy Schedules Based on Risk of Gleason Upgrading for Low-Risk Prostate Cancer Active Surveillance Patients, *BJU Int.* (2020). <https://doi.org/10.1111/bju.15136>.

- [78] A. Tomer, D. Nieboer, M.J. Roobol, E.W. Steyerberg, D. Rizopoulos, Personalized schedules for surveillance of low-risk prostate cancer patients, *Biometrics*. 75 (2019) 153–162. <https://doi.org/10.1111/biom.12940>.
- [79] A. Tomer, D. Rizopoulos, D. Nieboer, F.-J. Drost, M.J. Roobol, E.W. Steyerberg, Personalized Decision Making for Biopsies in Prostate Cancer Active Surveillance Programs, *Med Decis Making*. (2019) 0272989X19861963. <https://doi.org/10.1177/0272989X19861963>.
- [80] C. Serrat, M. Rué, C. Armero, X. Piulachs, H. Perpiñán, A. Forte, Á. Páez, G. Gómez, Frequentist and Bayesian approaches for a joint model for prostate cancer risk and longitudinal prostate-specific antigen data, *Journal of Applied Statistics*. 42 (2015) 1223–1239. <https://doi.org/10.1080/02664763.2014.999032>.
- [81] H. Lin, B.W. Turnbull, C.E. McCulloch, E.H. Slate, Latent class models for joint analysis of longitudinal biomarker and event process data: Application to longitudinal prostate-specific antigen readings and prostate cancer, *Journal of the American Statistical Association*; Alexandria. 97 (2002) 53–65.
- [82] R.Y. Coley, A.J. Fisher, M. Mamawala, H.B. Carter, K.J. Pienta, S.L. Zeger, A Bayesian hierarchical model for prediction of latent health states from multiple data sources with application to active surveillance of prostate cancer, *Biometrics*. 73 (2017) 625–634. <https://doi.org/10.1111/biom.12577>.
- [83] K. Suresh, J.M.G. Taylor, A. Tsodikov, A copula-based approach for dynamic prediction of survival with a binary time-dependent covariate, *Stat Med*. (2021). <https://doi.org/10.1002/sim.9102>.
- [84] A. Finelli, T.M. Beer, S. Chowdhury, C.P. Evans, K. Fizazi, C.S. Higano, J. Kim, L. Martin, F. Saad, O. Saarela, 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 Network Open*. 4 (2021) e2112426–e2112426. <https://doi.org/10.1001/jamanetworkopen.2021.12426>.
- [85] S. Desmée, F. Mentré, C. Veyrat-Follet, B. Sébastien, J. Guedj, Using the SAEM algorithm for mechanistic joint models characterizing the relationship between nonlinear PSA kinetics and survival in prostate cancer patients, *Biometrics*. 73 (2016) 305–312. <https://doi.org/10.1111/biom.12537>.
- [86] S. Desmée, F. Mentré, C. Veyrat-Follet, B. Sébastien, J. Guedj, Nonlinear joint models for individual dynamic prediction of risk of death using Hamiltonian Monte Carlo: application to metastatic prostate cancer, *BMC Med Res Methodol*. 17 (2017). <https://doi.org/10.1186/s12874-017-0382-9>.
- [87] S. Desmée, F. Mentré, C. Veyrat-Follet, J. Guedj, Nonlinear Mixed-Effect Models for Prostate-Specific Antigen Kinetics and Link with Survival in the Context of Metastatic Prostate Cancer: a Comparison by Simulation of Two-Stage and Joint Approaches, *AAPS J*. 17 (2015) 691–699. <https://doi.org/10.1208/s12248-015-9745-5>.
- [88] E. Lalonde, A.S. Ishkanian, J. Sykes, M. Fraser, H. Ross-Adams, N. Erho, M.J. Dunning, S. Halim, A.D. Lamb, N.C. Moon, G. Zafarana, A.Y. Warren, X. Meng, J. Thoms, M.R. Grzadkowski, A. Berlin, C.L. Have, V.R. Ramnarine, C.Q. Yao, C.A. Malloff, L.L. Lam, H. Xie, N.J. Harding, D.Y.F. Mak, K.C. Chu, L.C. Chong, D.H. Sendorek, C. P'ng, C.C. Collins, J.A. Squire, I. Jurisica, C. Cooper, R. Eeles, M. Pintilie, A.D. Pra, E. Davicioni, W.L. Lam, M. Milosevic, D.E. Neal, T. van der Kwast, P.C. Boutros, R.G. Bristow, Tumour genomic and microenvironmental heterogeneity for integrated prediction of 5-year

- biochemical recurrence of prostate cancer: a retrospective cohort study, *The Lancet Oncology*. 15 (2014) 1521–1532. [https://doi.org/10.1016/S1470-2045\(14\)71021-6](https://doi.org/10.1016/S1470-2045(14)71021-6).
- [89] D. Dearnaley, I. Syndikus, H. Mossop, V. Khoo, A. Birtle, D. Bloomfield, J. Graham, P. Kirkbride, J. Logue, Z. Malik, J. Money-Kyrle, J.M. O'Sullivan, M. Panades, C. Parker, H. Patterson, C. Scrase, J. Staffurth, A. Stockdale, J. Tremlett, M. Bidmead, H. Mayles, O. Naismith, C. South, A. Gao, C. Cruickshank, S. Hassan, J. Pugh, C. Griffin, E. Hall, Conventional versus hypofractionated high-dose intensity-modulated radiotherapy for prostate cancer: 5-year outcomes of the randomised, non-inferiority, phase 3 CHHiP trial, *The Lancet Oncology*. 17 (2016) 1047–1060. [https://doi.org/10.1016/S1470-2045\(16\)30102-4](https://doi.org/10.1016/S1470-2045(16)30102-4).
- [90] D.H. Brand, A.C. Tree, P. Ostler, H. van der Voet, A. Loblaw, W. Chu, D. Ford, S. Tolan, S. Jain, A. Martin, J. Staffurth, P. Camilleri, K. Kancherla, J. Frew, A. Chan, I.S. Dayes, D. Henderson, S. Brown, C. Cruickshank, S. Burnett, A. Duffton, C. Griffin, V. Hinder, K. Morrison, O. Naismith, E. Hall, N. van As, D. Dodds, E. Lartigau, S. Patton, A. Thompson, M. Winkler, P. Wells, T. Lymberiou, D. Saunders, M. Vilarino-Varela, P. Vavassis, T. Tsakiridis, R. Carlson, G. Rodrigues, J. Tanguay, S. Iqbal, M. Winkler, S. Morgan, A. Mihai, A. Li, O. Din, M. Panades, R. Wade, Y. Rimmer, J. Armstrong, M. Panades, N. Oommen, Intensity-modulated fractionated radiotherapy versus stereotactic body radiotherapy for prostate cancer (PACE-B): acute toxicity findings from an international, randomised, open-label, phase 3, non-inferiority trial, *The Lancet Oncology*. 20 (2019) 1531–1543. [https://doi.org/10.1016/S1470-2045\(19\)30569-8](https://doi.org/10.1016/S1470-2045(19)30569-8).
- [91] A. Widmark, A. Gunnlaugsson, L. Beckman, C. Thellenberg-Karlsson, M. Hoyer, M. Lagerlund, J. Kindblom, C. Ginman, B. Johansson, K. Björnlinger, M. Seke, M. Agrup, P. Fransson, B. Tavelin, D. Norman, B. Zackrisson, H. Anderson, E. Kjellén, L. Franzén, P. Nilsson, Ultra-hypofractionated versus conventionally fractionated radiotherapy for prostate cancer: 5-year outcomes of the HYPO-RT-PC randomised, non-inferiority, phase 3 trial, *The Lancet*. 394 (2019) 385–395. [https://doi.org/10.1016/S0140-6736\(19\)31131-6](https://doi.org/10.1016/S0140-6736(19)31131-6).
- [92] S. Halabi, C. Li, S. Luo, Developing and Validating Risk Assessment Models of Clinical Outcomes in Modern Oncology, *JCO Precision Oncology*. (2019) 1–12. <https://doi.org/10.1200/PO.19.00068>.
- [93] G.S. Collins, J.B. Reitsma, D.G. Altman, K.G. Moons, Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD Statement, *BMC Medicine*. 13 (2015) 1. <https://doi.org/10.1186/s12916-014-0241-z>.
- [94] D. Rizopoulos, G. Papageorgiou, P.M. Afonso, *JMbayes2: Extended Joint Models for Longitudinal and Time-to-Event Data*, 2022.
- [95] D. Rizopoulos, L.A. Hatfield, B.P. Carlin, J.J.M. Takkenberg, Combining Dynamic Predictions From Joint Models for Longitudinal and Time-to-Event Data Using Bayesian Model Averaging, *Journal of the American Statistical Association*. 109 (2014) 1385–1397. <https://doi.org/10.1080/01621459.2014.931236>.
- [96] D. Tilki, A.V. D'Amico, Timing of radiotherapy after radical prostatectomy, *The Lancet*. 396 (2020) 1374–1375. [https://doi.org/10.1016/S0140-6736\(20\)31957-7](https://doi.org/10.1016/S0140-6736(20)31957-7).

- [97] F. Zattoni, I. Heidegger, V. Kasivisvanathan, A. Kretschmer, G. Marra, A. Magli, F. Preisser, D. Tilki, I. Tsaur, M. Valerio, R. van den Bergh, C. Kesch, F. Ceci, C. Fankhauser, G. Gandaglia, Radiation Therapy After Radical Prostatectomy: What Has Changed Over Time?, *Frontiers in Surgery*. 8 (2021) 245. <https://doi.org/10.3389/fsurg.2021.691473>.
- [98] H.A. Elmarakeby, J. Hwang, R. Arafeh, J. Crowdis, S. Gang, D. Liu, S.H. AlDubayan, K. Salari, S. Kregel, C. Richter, T.E. Arnoff, J. Park, W.C. Hahn, E.M. Van Allen, Biologically informed deep neural network for prostate cancer discovery, *Nature*. 598 (2021) 348–352. <https://doi.org/10.1038/s41586-021-03922-4>.
- [99] S. Bernatz, J. Ackermann, P. Mandel, B. Kaltenbach, Y. Zhdanovich, P.N. Harter, C. Döring, R. Hammerstingl, B. Bodelle, K. Smith, A. Bucher, M. Albrecht, N. Rosbach, L. Basten, I. Yel, M. Wenzel, K. Bankov, I. Koch, F.K.-H. Chun, J. Köllermann, P.J. Wild, T.J. Vogl, Comparison of machine learning algorithms to predict clinically significant prostate cancer of the peripheral zone with multiparametric MRI using clinical assessment categories and radiomic features, *Eur Radiol*. 30 (2020) 6757–6769. <https://doi.org/10.1007/s00330-020-07064-5>.
- [100] S. Yu, J. Tao, B. Dong, Y. Fan, H. Du, H. Deng, J. Cui, G. Hong, X. Zhang, Development and head-to-head comparison of machine-learning models to identify patients requiring prostate biopsy, *BMC Urol*. 21 (2021) 80. <https://doi.org/10.1186/s12894-021-00849-w>.
- [101] S.B. Ginsburg, M. Rusu, J. Kurhanewicz, A. Madabhushi, Computer extracted texture features on T2w MRI to predict biochemical recurrence following radiation therapy for prostate cancer, in: *Medical Imaging 2014: Computer-Aided Diagnosis*, SPIE, 2014: pp. 69–81. <https://doi.org/10.1117/12.2043937>.
- [102] N. Momenzadeh, H. Hafezalseheh, M.R. Nayeypour, M. Fathian, R. Noorossana, A hybrid machine learning approach for predicting survival of patients with prostate cancer: A SEER-based population study, *Informatics in Medicine Unlocked*. 27 (2021) 100763. <https://doi.org/10.1016/j.imu.2021.100763>.
- [103] O.S. Tătaru, M.D. Vartolomei, J.J. Rassweiler, O. Virgil, G. Lucarelli, F. Porpiglia, D. Amparore, M. Manfredi, G. Carrieri, U. Falagario, D. Terracciano, O. de Cobelli, G.M. Busetto, F.D. Giudice, M. Ferro, Artificial Intelligence and Machine Learning in Prostate Cancer Patient Management—Current Trends and Future Perspectives, *Diagnostics (Basel)*. 11 (2021) 354. <https://doi.org/10.3390/diagnostics11020354>.
- [104] M. Raynaud, O. Aubert, G. Divard, P.P. Reese, N. Kamar, D. Yoo, C.-S. Chin, É. Bailly, M. Buchler, M. Ladrrière, M. Le Quintrec, M. Delahousse, I. Juric, N. Basic-Jukic, M. Crespo, H.T. Silva, K. Linhares, M.C. Ribeiro de Castro, G. Soler Pujol, J.-P. Empana, C. Ulloa, E. Akalin, G. Böhmig, E. Huang, M.D. Stegall, A.J. Bentall, R.A. Montgomery, S.C. Jordan, R. Oberbauer, D.L. Segev, J.J. Friedewald, X. Jouven, C. Legendre, C. Lefaucheur, A. Loupy, Dynamic prediction of renal survival among deeply phenotyped kidney transplant recipients using artificial intelligence: an observational, international, multicohort study, *The Lancet Digital Health*. 3 (2021) e795–e805. [https://doi.org/10.1016/S2589-7500\(21\)00209-0](https://doi.org/10.1016/S2589-7500(21)00209-0).
- [105] E. Waldmann, D. Taylor-Robinson, N. Klein, T. Kneib, T. Pressler, M. Schmid, A. Mayr, Boosting joint models for longitudinal and time-to-event data,

- Biometrical Journal. 59 (2017) 1104–1121.
<https://doi.org/10.1002/bimj.201600158>.
- [106] P. Dhiman, J. Ma, C.A. Navarro, B. Speich, G. Bullock, J.A. Damen, S. Kirtley, L. Hooft, R.D. Riley, B. Van Calster, K.G.M. Moons, G.S. Collins, Reporting of prognostic clinical prediction models based on machine learning methods in oncology needs to be improved, *Journal of Clinical Epidemiology*. 138 (2021) 60–72. <https://doi.org/10.1016/j.jclinepi.2021.06.024>.
- [107] C.L. Andaur Navarro, J.A.A. Damen, T. Takada, S.W.J. Nijman, P. Dhiman, J. Ma, G.S. Collins, R. Bajpai, R.D. Riley, K.G.M. Moons, L. Hooft, Completeness of reporting of clinical prediction models developed using supervised machine learning: a systematic review, *BMC Medical Research Methodology*. 22 (2022) 12. <https://doi.org/10.1186/s12874-021-01469-6>.
- [108] A.J. Vickers, E.B. Elkin, Decision Curve Analysis: A Novel Method for Evaluating Prediction Models, *Med Decis Making*. 26 (2006) 565–574. <https://doi.org/10.1177/0272989X06295361>.
- [109] E.W. Steyerberg, A.J. Vickers, N.R. Cook, T. Gerds, M. Gonen, N. Obuchowski, M.J. Pencina, M.W. Kattan, Assessing the Performance of Prediction Models: A Framework for Traditional and Novel Measures, *Epidemiology*. 21 (2010) 128–138. <https://doi.org/10.1097/EDE.0b013e3181c30fb2>.
- [110] X. Liu, L. Faes, M.J. Calvert, A.K. Denniston, Extension of the CONSORT and SPIRIT statements, *The Lancet*. 394 (2019) 1225. [https://doi.org/10.1016/S0140-6736\(19\)31819-7](https://doi.org/10.1016/S0140-6736(19)31819-7).
- [111] The Lancet Digital Health, Walking the tightrope of artificial intelligence guidelines in clinical practice, *The Lancet Digital Health*. 1 (2019) e100. [https://doi.org/10.1016/S2589-7500\(19\)30063-9](https://doi.org/10.1016/S2589-7500(19)30063-9).
- [112] G.S. Collins, P. Dhiman, C.L.A. Navarro, J. Ma, L. Hooft, J.B. Reitsma, P. Logullo, A.L. Beam, L. Peng, B.V. Calster, M. van Smeden, R.D. Riley, K.G. Moons, Protocol for development of a reporting guideline (TRIPOD-AI) and risk of bias tool (PROBAST-AI) for diagnostic and prognostic prediction model studies based on artificial intelligence, *BMJ Open*. 11 (2021) e048008. <https://doi.org/10.1136/bmjopen-2020-048008>.
- [113] G.S. Collins, K.G.M. Moons, Reporting of artificial intelligence prediction models, *The Lancet*. 393 (2019) 1577–1579. [https://doi.org/10.1016/S0140-6736\(19\)30037-6](https://doi.org/10.1016/S0140-6736(19)30037-6).