Journal Pre-proof

A Personalised Clinical Dynamic Prediction Model to Characterise Prognosis for Patients with Localised Prostate Cancer: analysis of the CHHiP Phase III Trial

Harry Parr MSc , Nuria Porta PhD , Alison C Tree MD(Res) , David Dearnaley FRCR , Emma Hall PhD

PII: S0360-3016(23)00167-0
DOI: https://doi.org/10.1016/j.ijrobp.2023.02.022
Reference: ROB 28091

To appear in: *International Journal of Radiation Oncology, Biology, Physics*

Received date: 28 November 2022
Revised date: 24 January 2023
Accepted date: 7 February 2023

Please cite this article as: Harry Parr MSc , Nuria Porta PhD , Alison C Tree MD(Res) , David Dearnaley FRCR , Emma Hall PhD , A Personalised Clinical Dynamic Prediction Model to Characterise Prognosis for Patients with Localised Prostate Cancer: analysis of the CHHiP Phase III Trial, *International Journal of Radiation Oncology, Biology, Physics* (2023), doi: https://doi.org/10.1016/j.ijrobp.2023.02.022

Title

A Personalised Clinical Dynamic Prediction Model to Characterise Prognosis for Patients with Localised Prostate Cancer: analysis of the CHHiP Phase III Trial

Short running title

Dynamic predictive model for prostate cancer

Authors

Harry Parr[1] MSc,
Nuria Porta[1] PhD,
Alison C Tree[2,1] MD(Res),
David Dearnaley[1,2] FRCR,
Emma Hall[1] PhD

1) Clinical Trials and Statistics Unit, The Institute of Cancer Research, London, United Kingdom
2) The Royal Marsden NHS Foundation Trust, London, United Kingdom

Corresponding author

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

Statistical analysis author

Mr Harry Parr, MSc; email: harry.parr@icr.ac.uk

Disclosures: ACT declares grant funding from Accuray, Elekta and Varian, and honoraria/travel assistance from Elekta, Accuray and Janssen. DD declares personal fees through the Rewards to Discoverer's Scheme from The Institute of Cancer Research which receives royalty income from abiraterone, honoraria for advisory boards from Janssen and has a patent EP1933709B1 issued for a localisation and stabilization device. EH reports grants from Accuray Inc, Varian Medical Systems Inc, Merck Sharp & Dohm, Janssen-Cilag, Aventis Pharma Limited (Sanofi), and Roche Products Ltd, and grants and nonfinancial support from Astra Zeneca, Bayer, all outside the submitted work. HP & NP have nothing to disclose.

Trial Management Group: Isabel Syndikus (chair), Angela Baker, Margaret Bidmead, Ananya Choudhury, David Dearnaley, Vincent Khoo, Helen Mayles, Philip Mayles, Olivia Naismith, Paul Ridley, Christopher Scrase, Chris South, John Staffurth, Alison Tree, Jean Tremlett, Emma Hall, Clare Cruickshank, Clare Griffin, and Julia Pugh.

Data Sharing Statement: Research data are stored in an institutional repository and will be shared upon reasonable request to the corresponding author.

The Institute of Cancer Research (ICR) Clinical Trials and Statistics Unit (CTSU) supports wider dissemination of information from the research it conducts to increase cooperation between investigators. Trial data are obtained, managed, stored, shared, and archived according to ICR-CTSU standard operating procedures to ensure the enduring quality, integrity, and utility of the data. Formal requests for data sharing are considered in line with ICR-CTSU procedures, with due regard given to funder and sponsor guidelines. Requests are via a standard proforma describing the nature of the proposed research and extent of data requirements. Data recipients are required to enter a formal data sharing agreement, which describes the conditions for release and requirements for data transfer, storage, archiving, publication, and Intellectual Property. Requests are reviewed by the Trial Management Group in terms of scientific merit and ethical considerations, including patients' consent. Data sharing is undertaken if proposed projects have a sound scientific or patients' benefit rationale, as agreed by the Trial Management Group and approved by the Independent Data Monitoring and Steering Committee, as required. Restrictions relating to patients' confidentiality and consent will be limited by aggregating and anonymizing identifiable patients' data. Additionally, all indirect identifiers that could lead to deductive disclosures will be removed in line with Cancer Research UK data sharing guidelines.

Acknowledgments: We thank ICR-CTSU CRUK current core programme grant and ICR for the PhD award funding without whom the research and this paper would not have been possible. We would also like to thank the patients and all the staff at the participating centres and the Trial Steering Committee for their continued oversight. CHHiP is sponsored by The Institute of Cancer Research (ICR) and co-ordinated by the ICR Clinical Trials and Statistics Unit. We acknowledge support from Cancer Research UK (C46/A3976, C46/A10588, C33589/A19727, C33589/A28284, C7224/A28724) and the National Institute for Health Research (NIHR) Cancer Research Network. This project represents independent research supported by the National Institute for Health and Care Research (NIHR) Biomedical Research Centre at The Royal Marsden NHS Foundation Trust and the Institute of Cancer Research, London. The views expressed are those of the author(s) and not necessarily those of the NIHR or the Department of Health and Social Care.

Abstract

Background

The CHHiP trial assessed moderately hypofractionated radiotherapy in localised prostate cancer. We utilised longitudinal prostate-specific antigen (PSA) measurements collected over time to evaluate and characterise patient prognosis.

Methods

We developed a clinical dynamic prediction joint model to predict the risk of biochemical or clinical recurrence. Modelling included repeated PSA values and adjusted for baseline prognostic risk factors of age, tumour characteristics and treatment received. We included 3,071 trial participants for model development using a mixed-effect submodel for the longitudinal PSAs, and a time-to-event hazard submodel for predicting recurrence of prostate cancer. We evaluated how baseline prognostic factor subgroups impacted on the nonlinear PSA levels over time and quantify the association of PSA on time-to-recurrence. We assessed bootstrapped optimism-adjusted predictive performance on calibration and discrimination. Additionally, we performed comparative dynamic predictions on patients with contrasting prognostic factors and investigated PSA thresholds over landmark times to correlate with prognosis.

Results

Patients that developed recurrence had generally higher baseline and overall PSA values during follow-up and had an exponentially rising PSA in the two-years before recurrence. Additionally, most baseline risk factors were significant in the mixed-effect- and relative risk submodels. PSA value- and rate-of-change was predictive of recurrence. Predictive performance of the model was good across different prediction times over an 8-year period, with an overall mean AUC of 0.70, mean

Brier score of 0.10, and mean integrated calibration index of 0.048; these were further improved for predictions after 5 years of accrued longitudinal post-treatment PSA assessments. PSA thresholds less than 0.23ng/mL after 3 years were indicative of a minimal risk of recurrence by 8 years.

Conclusions

We successfully developed a joint statistical model to predict prostate cancer recurrence, evaluating prognostic factors and longitudinal PSA. We showed dynamically updated PSA information can improve prognostication, which can be used to guide follow-up and treatment management options.

Keywords: joint model, prostate cancer prognosis, biochemical and clinical failure, prostate-specific antigen (PSA), intensity modulated radiotherapy (IMRT), hypofractionation.

1 Introduction

Prostate cancer is the second most common cancer in men globally [1]. In the UK, at diagnosis over half of men (56%—61%) present with localised prostate cancer (TNM staging: T1b–T3aN0M0) [2]. Radical treatment with neoadjuvant hormone therapy and intensity-modulated radiotherapy (IMRT) is less invasive and generally better tolerated for long-term quality-of-life than radical prostatectomy [3–5]. Following the publication of three randomised controlled trials, including CHHiP, showed that moderately hypofractionated radiotherapy was non-inferior to conventional (2 Gray/fraction) radiotherapy; hypofractionation is now used as a standard of care in Europe and North America [6–10].

At diagnosis, patients are routinely stratified into the National Comprehensive Cancer Network (NCCN) risk groups, which guide initial treatment management decisions [11]. NCCN risk groups are defined by presenting tumour features at diagnosis, including TNM staging, Gleason score (GS), and prostate-specific antigen (PSA), that are used to predict prognosis [12]. Although these prognostic factors may stratify patient groups, they do not always accurately predict the risk of biochemical or clinical failure of individual patients [13].

PSA is used in routine follow-up to monitor for cancer recurrence and define biochemical failure. After radiotherapy, typically nadir PSA values (the lowest concentration) are observed; the nadir directly defines each patient's biochemical failure threshold [14]. Data collected during routine follow-up, including repeat measures of PSA, may be of additional prognostic value to update the predicted risk of recurrence. Studies have investigated pre-treatment prognostic factors of the risk of biochemical failure in CHHiP [7]. In this work, in addition to (baseline) pre-treatment information, we propose the use of joint modelling methodology [15–18]. We incorporate

PSA values collected over time (the longitudinal process) to obtain updated predictions of the risk of biochemical- or clinical failure (the time-to-event process), as new information becomes available. This could lead to a more personalised approach to follow-up care and management. For instance, if a patient remains recurrence-free for a prolonged period, and the PSA trajectories would classify the patient as having a low recurrence risk, then a possible recommendation could be to reduce the patient's follow-up schedule, resulting in less burden for both patients and clinics. Conversely, if the patient's risk increases, it may enable the clinician to initiate more intensive follow-up or direct alternative therapies as appropriate [19].

The objective of our research is to develop a clinical dynamic prediction joint model utilising longitudinally collected PSAs to then predict the risk of future recurrence in the CHHiP trial. We present dynamic predictions of the developed model on prognosis for two patients contrasting in their baseline prognostic factors, then evaluate the predictions and performance with internal bootstrapped validation. We then deduce PSA thresholds that are indicative of good prognosis, with minimal $< 5\%$ risk of recurrence.

2 Methods & Materials

2.1 Study design & procedure

CHHiP is an international, multicentre, randomised, phase III, non-inferiority trial. Men with localised prostate cancer (T1b-T3aN0M0) were randomised (1:1:1) to receive conventional radiotherapy 74Gy in 37 fractions (f) over 7.4 weeks, or one of two hypofractionated radiotherapy schedules: 60Gy/20f in 4 weeks or 57Gy/19f over 3.8 weeks. The protocol mandated hormone therapy in men with NCCN intermediate and high-risk disease, for at least 3 months (maximum 6 months) before start of

radiotherapy and continued until the end of radiotherapy; this was optional for low-risk patients. Bicalutamide monotherapy or luteinising-hormone-releasing-hormone analogue plus possible short-term anti-androgen (LHRHa) were permitted according to patient and physician's choice. PSA values were recorded pre-hormone therapy and pre-radiotherapy; then at weeks 10, 18, and 26 after start of radiation therapy; and then at intervals of 6 months after end of radiotherapy for 5 years; then annually thereafter. The trial was registered (ISRCTN97182923), approved by the London Multicentre Research Ethics Committee (04/MRE02/10) and by the institutional research board of each participating international site. This study was conducted in accordance with principles of good clinical practice; full details of the trial design have been described previously [7].

2.2 Outcomes

Prostate cancer recurrence was defined as the composite of biochemical or clinical failure or death due to prostate cancer. Biochemical failure was defined using the Phoenix definition of a $\mathrm{PSA} > \mathrm{the\,nadir} + 2\mathrm{ng/mL}$ [14]. Clinical failure included: recommencement of hormone therapy, local recurrence, lymph node or pelvic recurrence, and distant metastases. Time-to-recurrence was calculated as the time between the patient's closest pre-treatment PSA before hormone therapy (time origin $t = 0$), and the first primary endpoint event. The median time between the closest pre-treatment PSA and randomisation was 15 weeks. Patients who were alive and recurrence-free or died due to causes unrelated to prostate cancer were censored at their last known follow-up date, with administrative censoring of longitudinal follow-up taking place at 10 years after time origin.

For this study, only patients who received hormone therapy and had at least one post-treatment PSA were included for model development. Complete-case analysis

was undertaken for the baseline prognostic factors. We based analyses on a data snapshot taken on October 9, 2019.

2.3 Joint modelling statistical methodology

A Bayesian shared-parameter joint modelling framework was used to develop the clinical dynamic prediction joint model (CDPJM). We specify a mixed-effects submodel to model PSA trajectories over time, and a Cox hazard submodel to model the time-to-recurrence endpoint. The shared parameters link the two submodels together, allowing us to quantify how a specific PSA trajectory is associated with risk of prostate cancer recurrence.

In the mixed effect submodel, baseline prognostic factors are included as fixed effects, while random effects are included to capture the individual variability of each patient's presenting PSA (random intercept) and deviation over time (random slopes). Deviations of PSA from the predicted trajectory are assumed to follow normally distributed measurement errors; PSA is log-transformed to conform to the distributional assumptions. Natural (restricted) cubic splines are employed to capture the nonlinear PSA over time. Implementing these splines is advantageous as they allow nonlinear PSAs to be flexibly modelled, without specific parametric assumptions needed, such as exponential-decay-growth, or biphasic parameterisations [20,21]. The splines split the range of the continuous PSA values into sub-intervals.

In the Cox hazard submodel it is assumed that the risk of recurrence depends on the trajectory of the longitudinal PSA biomarker. Thus, the trajectory parametrised via the mixed-effect submodel, considering the entire longitudinal history up to a time point t for each patient, is imputed into the Cox parameterisation as a linear

predictor, with a specified association structure where the features of the longitudinal PSA biomarker outcome are included. Typical associations structures are the value or the rate-of-change of PSA, or a linear combination of the two. I.e., this combinatory association structure of the PSA trajectory (at time t) could be associated with the hazard of recurrence at that same point in time.

Models were developed using R software (v4.1.0). The individual submodels were fitted by maximum likelihood estimation using the *survival* (v3.2-11) and *nlme* (v3.1-152) R packages; the fully specified joint model was estimated with Bayesian Markov chain Monte Carlo sample algorithm, using *JMbayes2* (v0.1-64–0.2-3) [22–24]. Computation was performed on a CentOS 8 Linux high-performance computer and Windows 10 Intel Core i9-8950HK CPU. Further details can be found in the *supplementary materials*.

2.4 Dynamic predictions

For each individual patient, we considered their longitudinal PSA biomarker values up to the landmark time point t, where we assumed they are recurrence-free. We wish to make predictions about their prognosis, within some clinically relevant prediction window in the future $[t, u]$; $u > t$, say two-, five-, or ten years from present landmark time t. The joint model estimates the probability of recurrence within time u, given the information available up to time t. These predictions are dynamic, as they can be updated as new follow-up and PSA information becomes available for that patient (that is, by increasing the landmark time t) [17].

2.5 Assessing predictive performance and risk thresholds

Predictive performance of the joint model was evaluated at varying landmark times, by assessing its discrimination via time-dependent area under the curve (AUC) and

its calibration via the integrated calibration index (ICI) metrics [25,26]. The ICI is the absolute mean difference between the predicted- and observed event probabilities. Overall prognostic performance was measured by estimating the Brier score, which is the expectation of the squared difference between the predicted and observed event probabilities, comprised of both calibration and discrimination [16,27]. Higher AUC metrics indicate superior discrimination, for both ICI and Brier, smaller measures indicate closer predicted and observed agreement and better model calibration. Internal validation of the proposed CDPJM was pursued by internal bootstrapping (50 repetitions) to account for any over-optimism, and to correct biases accordingly. We then compared the CDPJM predictions at future landmark times, with the predictions obtained when no longitudinal PSA biomarker information is available (i.e., at $t = 0$), to assess the improvement that longitudinal PSAs make.

As well as personalised predictions, it is often useful for clinicians to have threshold values of PSA which give acceptable risk profiles following radiotherapy and short-course hormone therapy. We used linear regression to quantify the association of PSA values from zero (baseline) to five years to correlate to the predicted risk of recurrence by eight years.

3 Results

3.1 Dataset for model building

The CHHiP trial randomised 3216 participants, of which data from 3071 (95%) were used to develop the statistical CDPJM. We excluded 104 participants who did not receive hormone therapy ($n = 90$) or had missing hormone therapy allocation ($n = 14$); 5 who received maximal androgen blockade; 3 with at least 1 baseline prognostic factor missing; 9 with no baseline pre-treatment PSA available, and 24

with missing PSA values beyond baseline over time (non-mutually exclusive). Table 1 presents the baseline characteristics of the included patients. Median follow-up of this subset was 8.6 years (IQR=[6.3–10.1]), and the median number of PSA values per patient was 16 (IQR=[13—18]).

Of the 3071 patients, 607 (20%) had recurrence, a composite endpoint of biochemical (n=541, 18%), clinical failure (n=65, 2%), or prostate cancer death (n=1). There were an additional 148 patients that exhibited PSA values that met the biochemical failure threshold but were not confirmed by a subsequent PSA observation. A further 355 (12%) patients died due to causes unrelated to prostate cancer. These patients were censored at the time of last follow-up for the primary analysis. This outcome was not considered a competing risk as the estimated cumulative incidence function accounting for competing risks and without (1 minus Kaplan-Meier estimate) yielded almost overlapping curves (e.g. the maximum difference was found at 10 years, between 0.216 and 0.23, respectively). [28].

3.2 Modelling of PSA trajectories

In Figure 1(a), PSA levels and boxplot distributions are presented, aggregated by years since starting treatment and outcome, with patients still at risk in the table below. There is much more variability and an increase in PSA values for those patients that recur at any time, compared to those that are alive and free from recurrence at their last follow-up. Presenting PSA values ($t = 0$) are higher for patients who recur. Apparent separation between the distributions of PSA is evident from year four onwards. Figure 1(b) shows the smoothed reverse-year PSA trajectories (i.e., the PSA course in the years before a recurrence or end of follow-up) of those that are recurrence-free and those patients that develop recurrence. Patients who develop a recurrence have higher presenting PSA levels, and do not

achieve the same PSA reduction after treatment as patients who do not recur. In the final two-years before patients who developed recurrence, PSA increases at an exponential rate, compared with recurrence-free / censored patients whose PSA remains at a very low plateau.

Figure 2 shows the mixed-effect joint model predictions and how each baseline factor impacts on the PSA trajectory, by outcome. We see initial high levels of PSA at diagnosis which drop for both groups during treatment. When treatment stops, PSA recovery/bounce is seen at 1-2 years after treatment, the slight bump around 2 years is likely due to the effects of testosterone recovery. For those that go on to remain event-free, a slight decrease is seen and then a stable plateau.

For fractionation schedule, there is generally little difference between the PSA trajectories for each schedule in the first year, and then PSA slightly deviates post-two years with systematically lower predicted PSA values in the 60Gy/20f arm for those with no recurrence, but highest predicted PSA values for those that do recur. Visually there does not appear to be much predicted difference in the GS and T-stage for the lower risk factors, but GS≥8 and T3 subgroups appear to exhibit lower PSA trajectories. Patients who received LHRHa appear to have lower predicted PSA values than those receiving bicalutamide; noting that allocation to hormone therapy was not randomised, with most patients (87%) receiving LHRHa. The biggest effect on PSA trajectories is age at diagnosis, with younger patients (ages 40-49, n=6) exhibiting higher post-treatment PSAs; for those who do not relapse, a stable PSA after 4 years is seen across all age groups. Parameter estimates for the final mixed-effects submodel component can be found in the *supplementary materials tables S1 & S2*.

3.3 Joint modelling time-to-recurrence

Conditioning on the PSA trajectory, fractionation schedule did not show a statistically significant effect (ref 74Gy/37f; 57Gy/19f: HR=0.99 95% credible interval (CI)=[0.65, 1.50]; 60Gy/20f: HR=1.01 95%CI=[0.71, 1.44]). GS (ref ≤ 6; 3+4: HR=1.81 95%CI=[1.33, 2.49]; 4+3: HR=2.76 95%CI=[1.95, 3.95]; ≥ 8: HR=2.49 95%CI=[1.27, 4.96]), T-stage (ref T1; T2: HR=1.47 95%CI=[1.10, 1.97]; T3: HR=2.41 95%CI=[1.52, 3.76]) and age (HR=1.05 95%CI=[1.03, 1.08]) were associated with the risk of recurrence. Patients who received bicalutamide appeared to have lower risk of recurrence (HR=0.70 95%CI=[0.48, 1.00]), although this was not statistically significant (p=0.053), in line with previous results [29].

For the association with the mixed-effect model, the log-hazard ratio parameter estimates for both the PSA value- and PSA gradient are 4.52 (95%CI=[4.07, 4.99]), and 2.08 (95%CI=[1.74, 2.43]), respectively, indicating that both absolute PSA value and its gradient at a given time as parameterised in the mixed-effects model are highly predictive of recurrence. Parameter estimates for the joint model can be found in the *supplementary materials tables S3 & S4,* including comparison with a Cox model with baseline-only covariates.

3.4 Dynamic predictions

We demonstrate how the model updates prognosis over time on two selected patients who received the same treatment (57Gy/19f radiotherapy schedule, LHRHa hormone therapy) and PSA follow-up schedule, similar age at diagnosis and contrasting NCCN risk groups at presentation (patient *A*: GS=8, T3, presenting PSA=5.3ng/mL, vs patient *B*: GS=6, T1, presenting PSA=9ng/mL), and outcome. Dynamic predictions for these two patients are presented in Figure 3 over five panels (**V—Z**) for different prediction landmark times ($t = 0, 1, 3.5, 4.5, 5$ years), to predict

risk ten years after initiating treatment. On each panel, the left-side of each figure depicts PSA (in blue, observed PSA values in dots, while line depicts estimated predicted PSA) and the right-side shows the point estimate of the cumulative risk of recurrence up to ten years from the landmark time (in green the curve for *A* who does not experience recurrence, the red for *B* who does). The shaded areas show the 95% credible intervals of the estimated predictions for each outcome.

At baseline (at $t = 0$ years, Figure **3V**), patient *A* has poorer baseline prognostic factors and worse ten-year prognosis (~45% recurrence risk) than patient *B* (~30% recurrence risk), despite having a lower presenting PSA. For both patients, using only presenting PSA gives very wide credible risk intervals for the predictions beyond two years. A year since starting treatment (Figure **3W**), both patients exhibit a similar drop in PSA with patient prognosis slightly improving for *B*. In Figure **3X** (landmark $t = 3.5$), *A*'s PSA remains low whilst *B*'s PSA level starts to increase beyond the plateau. In Figure **3Y** (landmark $t = 4.5$), A's PSA continues to remain low and stable, with their risk substantially dropping, whilst B's PSA continues to increase thereby further increasing his risk of recurrence. In Figure **3Z** after 5 years follow-up, *A*'s PSA is very stable around 0.1 ng/mL, thus his updated prognosis is very good, with reduced credible intervals for his predictions, compared to *B*'s, whose post-treatment PSA presents more variability and increases over time. The risk of recurring by 10 years for *A* is very small (~5%) compared to *B*'s risk of recurrence (>60%), with jumps in estimated risk at each previous landmark after a year. This is driven by the accrued PSA levels before 5 years, approaching biochemical failure. Patient *A* was recurrence-free by 9 years follow-up, whilst *B* had a recurrence by 5½ years.

3.5 Assessing predictive performance

We assessed the CDPJM's calibration and discrimination for predictions of risk to recurrence by 8 years. Presented in Table 2, the 50-times-repeated bootstrapped optimism-corrected metrics (mean for each time point) for the CDPJM to predict biochemical / clinical failure for landmark times at years zero ($t = 0$, baseline) to seven ($t = 7$), with a fixed horizon prediction of eight years. Discrimination improves as more longitudinal PSA information becomes available after 3 years' worth and AUC was maximised after 5 years of follow-up 0.84, 95% bootstrapped CI (bCI)=[0.81—0.87]. Similarly, calibration and Brier improves considerably after four years. The overall corrected AUC is 0.70, 95% bCI=[0.51—0.86]; ICI=0.05, 95% bCI=[0.014—0.089]; Brier=0.10, 95% bCI =[0.025—0.164].

3.6 PSA risk thresholds

In Figure 4, we performed linear regression analysis between the predicted risk of recurrence by eight years from the joint model, given the accrued longitudinal biomarker information up to landmark time t ($t = 0, 1, \ldots, 5$ years), and the latest PSA value available prior to the landmark time. At all landmarks there is a strong positive correlation between latest PSA value and predicted risk of recurrence by eight years. As the latest PSA value nearest to landmark time increases, predicted prognosis worsens with an increased recurrence probability. The recurrence risk threshold is minimised at the origin for each landmark time t. This gives an approximate level of an 'acceptable' average PSA threshold on a continuous scale.

For instance, we see that at the start of treatment (landmark $t = 0$ years) in Figure 4 (top-left), for a minimal PSA, the lowest predicted risk is 13% (y-intercept) and a relatively small R^2 value as there is some heterogeneity here at baseline time origin with wider 95% prediction interval (PI) bands (8-year recurrence risk PI 0% to 29% at

the intercept). As follow-up continues ($1 \leq t \leq 2$ years), PSA drops to the nadir (the lowest recorded PSA) which is near-zero. The intercept implies a minimal PSA predicts a recurrence risk of 11% and 7% at landmark time 1 and 2 years respectively. At landmark times 3, 4, and 5 years, the regression intercepts are negative (a nil PSA implying an infeasible negative risk), though their magnitudes are very small; PSA levels less than 0.23, 0.34, and 0.41ng/mL respectively predict a small ($< 5\%$) risk of recurrence by 8 years.

4 Discussion & conclusions

In this study, we have developed a dynamically updated clinical prediction joint model for the risk of prostate cancer relapse in patients treated with both hormone therapy and IMRT in the CHHiP trial. We showed that incorporating longitudinal PSA values collected over time into the model, in addition to baseline prognostic factors and treatment schedules, aids and improves prediction of individual patient prognosis. We explored and quantified the effect of hypofractionation (3Gy/f) compared to conventional fractions (2Gy/f) on patients' longitudinal PSA trajectories and on recurrence. There was no statistical evidence of a difference between either of the hypofractionation schedules, compared to the conventional fractionation arm, in terms of the PSA trajectories or reducing recurrence risk as expected due to the non-inferiority hypothesis of the study design.

PSA levels typically started to rise exponentially approximately 1½-2 years before formal biochemical failure. We quantified the association of PSA values and its rate-of-change, both being highly significant and predictive of recurrence. The rationale to include PSA gradient is that there may be non-recurring patients who have a higher post-radiotherapy PSA value but continues to be stable (non-increasing over time), compared with a patient who may have a lower PSA post-radiotherapy that

continues to increase post-treatment. With the entire PSA trajectory captured and supplied to the CDPJM, PSA to nadir is directly modelled and has previously been shown to be an important predictor of event-free survival [30]. Similarly, inference in changes of the minimum (nadir) PSA between patients can be made. The nadir often occurred by two years since treatment commenced, with PSA value and gradient at the nadir both being close to zero. E.g., take two similar patients where their only clinical difference is a nadir of nil and 0.1 after a year of starting treatment. The predicted recurrence risk by 8 years of the patient with a higher nadir is 4.85%, over doubled from the risk of the patient with lower nadir, 2.06%. However, in absolute values, this is still a small increase, and would still be considered to have a good prognosis.

We also attempted to quantify the relationship between PSA values at particular landmarks from starting treatment and subsequent recurrence. This is not straightforward as it is difficult to define precise or best cut-offs for PSA which need clinical (and patient) value judgements. For example, we saw some implausible predicted risk values from the regression parameter estimates (the intercept) between landmark times 3—5 years after treatment. Perhaps more careful consideration and sophisticated methods could be applied here or forcing the intercept to be zero. However, this was to give an indication of upper PSA bounds to predict a recurrence risk of $< 5\%$ by 8 years at various landmarks in a simple 'rule-of-thumb' without overcomplicating the interpretation. It is worth noting that the extracted risk predictions are obtained from the joint model. It is not just the PSA value by the landmark time which is considered, it is also its rate-of-change and its history modelled by the mixed-effects submodel. Using only the most recent PSA value predictor at the landmark time is a simplified approach, as the raw

concentration is a proxy to each patient's PSA trajectory by the landmark time point. Additionally for personalised predictions it is difficult to give one-size-fits-all cut-offs, and a balance must be made to the weighting and importance of false-positives and false-negative predictions. Our data suggest that PSA levels ≤0.23, 0.34, and 0.41ng/mL at 3, 4, and 5 years respectively give a reasonable indication of having a $< 5\%$ risk of recurrence by 8 years. For those same landmark times and risk thresholds, 27%, 40%, and 51% proportion of patients have this PSA threshold (or less). It is encouraging that these thresholds are consistent with previous findings in the context of prognosis after brachytherapy [31] and mono-external-beam radiotherapy [32,33]. Yock et al similarly state that a 5-year PSA ≤0.5ng/mL has very good prognosis (97% progression-free rates by 8 years) [33]. These studies [32,33] have some differences with ours, being over 20 years old, lower radiation doses were delivered, with no hormone therapy, and they used PSA categories at a fixed time at 5 years using a simplified Kaplan-Meier landmarking approach. This may explain the slightly lower threshold we found; our continuous method for ascertaining these thresholds is more flexible, without arbitrarily categorised PSAs.

Conversely a PSA of 1ng/mL at 5 years gives a predicted risk of recurrence of 20% (95% PI=[6%–33%]). The reason we report prediction intervals rather than the smaller confidence interval is to have a prediction range for new patients entering this treatment pathway where most patients are within the PIs. There is reasonable heterogeneity in the Figure 4 scatter plots at the earlier landmark times (indicated by the lower R^2 values). It is worth noting the individualised predictions directly from the joint model will give bespoke credible intervals too. This study supports the importance of presenting, nadir, and post-treatment recovery levels of PSA. Findings

from this CDPJM suggest that patients with a PSA ≲ 0.23ng/mL and stable (low or nil gradient) PSA from 3 years onwards have good prognosis.

We chose to validate up to a fixed horizon time of 8 years, given that the median time at the date of data snapshot was 8.6 years, despite being able to extend this as seen in the dynamic predictions. We chose the fixed horizon approach to exemplify how predictions improve as more PSA information is collected. The model also allows predictions at fixed prediction windows, such as two- and five years from fixed landmark times (e.g., given data up to three years, what is the predicted risk of recurrence in the next two years). These (non-corrected) metrics are presented in the *supplementary materials table S5.* There is relatively little difference in the validation between the two methods (apparent vs bootstrapped-corrected) AUC and Brier scores – with some notable differences in the earlier landmarks. There are slightly bigger differences in the first two years for the ICI metrics.

Diagnostics of the joint model were performed (not shown). The longitudinal component conforms to the assumptions, though there were some departures observed in the tails of the quantile-quantile plot, suggesting t-distributed residuals could be appropriate. The random effects themselves conform to normally distributed residuals. The Cox submodel proportional hazards assumption for baseline covariates were checked and found not to be violated; a joint model including time-varying $\alpha_1(t)$ showed some departures of PH for time-varying PSA process, which become reasonably constant after the nadir. Previous work has shown the joint model is highly robust to departures in the proportional hazards assumption [34].

A limitation of the study is the inherent association of the longitudinal process and the outcome, as we included biochemical recurrence in the definition of the outcome.

This is because the primary endpoint in CHHiP captured failure-free survival, which is time free of any event that would trigger further treatment for the patient (or prostate cancer death). For this reason, for patients whose biochemical failure triggered treatment, it was not always possible to confirm clinical radiological progression.

Furthermore, we acknowledge the relative complexity of the joint model, namely in the mixed-effect submodel. The parameterisation of the longitudinal mixed-effects model is complicated with four internal cubic spline knots over time to capture the exhibited nonlinear PSA, with 15 main-effect parameters and a total of 71 parameters to estimate. We do feel that the complexity of the model is warranted. We further investigated (not presented) pairwise interactions of the baseline variables with time; an additional 45 parameters to be estimated and as most curves were reasonably parallel, this was considered adding unnecessary complexity and did not improve the deviance information criterion.

When conditioning on PSA trajectory, it appeared that receiving Bicalutamide magnified a reduced risk of recurrence, compared to LHRHa (not randomised) when using the regular Cox model (see *supplementary materials table S3*, and Tree et al [29]). Conversely, PSAs remained higher than LHRHa patients (Figure 2). Surprisingly, we saw patients with worse prognostic factors (Gleason ≥ 8 & T3) have the lowest PSA trajectory, however there were relatively few patients in these subgroups (n=97 and 270 respectively).

Dynamic prediction models that incorporated longitudinal PSA levels to predict risk of recurrence in prostate cancer have been previously explored. A full review of these relevant studies can be found in Parr et al with similar dynamic prediction windows

and expected predictions to our study [18]. However, these previous studies use PSA dynamics after standard external-beam radiotherapy has ceased and without neoadjuvant or concurrent hormone therapy; therefore, in their setting the PSA dynamics are different, with lower PSA values at $t = 0$, slower PSA decrease to nadir, and elongation of its trajectory, compared to the PSA dynamics we observed from pre-hormone treatment PSA values. As recruitment of these previous studies occurred in the 1980s, there have been significant advances in treatment, with 5- and 10-year survival rates doubling (in the UK) since then [35]. Additionally, the majority of CHHiP patients received neoadjuvant and concurrent hormone therapy, with hypofractionated radiotherapy regime, therefore our model and analysis is applicable to the current standard-of-care.

We compared the prognostic performance of the CDPJM to other published articles. Arguably most similar to our work is Taylor et al who propose a joint model using real-time evaluation of predicting recurrence of prostate cancer [36]. The longitudinal PSA biomarker was modelled using a biphasic exponentially decreasing-increasing parametric function. Some of their parameter estimates were remarkably similar to ours, namely in the log-hazard ratios to the PSA level, T-stage and Gleason. Their prediction time focuses on a window of no more than three years ahead, whereas we present a fixed horizon prediction time of eight years (see *supplementary materials table S5*). Their exhibited PSA trajectory is an elongated tick-shape, typical of radiotherapy-only treatment. Direct model comparison cannot be made due to their differing validation appraisal methods, and lack of androgen deprivation therapy. There is likely not much difference in predictability between mono-radiotherapy and dual-therapy at earlier landmarks. However, the nadir may occur later for

monotherapy patients, which could slightly decrease the predictability compared to dual-therapy at the nadir.

As follow-up continues, and there are an ageing population of patients in CHHiP, we assessed whether deaths from non-prostate cancer related causes may represent a competing risk for our outcome of interest, but finding it was not an issue, these deaths were treated as censored in our model. Extensions exist for joint models accounting for competing risks, but extracting dynamic predictions and assessing their predictive performance is not trivial [37,38]. Other extensions to our model could include an additional multivariate longitudinal process (e.g. with both PSA and testosterone), which is known to be prognostic in later disease stages [39], or novel biomarkers of early detection of recurrence, such as circulating-tumour DNA fraction [40]; or additional histopathological prognostic factors, such as Ki67 [41]. These however were not routinely collected in this trial.

In this work, we proposed joint modelling to characterise PSA trajectories and how these impact on predicting risk of recurrence in the CHHiP trial. Further work includes developing a gold-standard dynamic predictive tool to be used within the clinic. Other sophisticated predictive algorithms exist, however typically use only baseline factors under varying treatment modalities and outcomes [42,43]. For example, we saw that patient *B* in Figure 3 had poor prognosis evident from their increasing PSA from 3½ years, despite their relatively good baseline prognostic factors. Although worse baseline prognostic factors, PSA trajectory indicated patient *A's* prognosis was good, and continued to be so after 4 years of follow-up; the model could be further extended to recommend and reduce follow-up frequency and burden. For instance, amongst the patients recurrence-free and alive at five years, the median time to failure for patients who recurred after 5 years was 7 years. In this

cohort, the median predicted cumulative incidence of recurrence is 4% by year 6, 12% by year 7, 20% by year 8, 27% by year 9 and 34% by year 10. Amongst patients who do fail by 7 years $[t = 5, u = 7]$ (n=136), their median cumulative risk of failure is 30%; compared to a median of 2% risk of failure for equivalent patients who are censored by 7 years (n=423). This demonstrates the predictive difference of the two outcomes and the two-year lead-time capabilities of the model, suggested by the reverse-time plot (Figure 1(b)).

Our clinical calculator would allow the clinician to visualise each patient's personalised risk of recurrence over time; if the risk surpasses an unacceptable threshold, further investigation could be considered, and personalised follow-up schedules could be designed [19,44]. To achieve this, we plan for the CDPJM to undergo robust external validation so we can assess its clinical utility in differing patient populations and alternative treatment modalities where similar PSA dynamics are expected. For instance, we hope to explore the generalisability of our proposed model with stereotactic radiotherapy or using longer hormone therapy schedules [45,46]. It may be that with differing treatments and disease stage, alternative model development and/or recalibration is required. Additionally decision analysis could quantify net benefit at various thresholds, versus a 'do-all-or-nothing' approach [47].

To conclude, we quantified the impact of an increase in PSA value- and rate-of-change on prostate cancer recurrence, adjusting for baseline prognostic factors and treatments in the CHHiP trial. Our model will be applicable to future patients who undergo hormone therapy with either conventional or hypofractionated IMRT. As expected, PSA is predictive of recurrence, as previous studies have shown. We also assessed the performance of the prediction model, which showed good calibration and discrimination, optimised after 4-5 years' of accrued longitudinal PSA biomarker

information to predict recurrence by 8 years. We demonstrated the practical aspect of these models in performing dynamic predictions from the relevant patient population that can help to guide patient care and allocate limited resource more effectively. We have also proposed clinical thresholds at various landmarks, with simple continuous calculations to determine alternative PSA thresholds given the recurrence risk clinicians might be willing to accept, which is easily applicable in clinical practice.

References

1. Rawla P. Epidemiology of Prostate Cancer. *World J Oncol*. 2019;10(2):63-89. doi:10.14740/wjon1191

2. McPhail S, Johnson S, Greenberg D, Peake M, Rous B. Stage at diagnosis and early mortality from cancer in England. *British Journal of Cancer*. 2015;112(1):S108-S115. doi:10.1038/bjc.2015.49

3. Wang Z, Ni Y, Chen J, et al. The efficacy and safety of radical prostatectomy and radiotherapy in high-risk prostate cancer: a systematic review and meta-analysis. *World Journal of Surgical Oncology*. 2020;18(1):42. doi:10.1186/s12957-020-01824-9

4. Mazariego CG, Egger S, King MT, et al. Fifteen year quality of life outcomes in men with localised prostate cancer: population based Australian prospective study. *BMJ*. 2020;371:m3503. doi:10.1136/bmj.m3503

5. Donovan JL, Hamdy FC, Lane JA, et al. Patient-Reported Outcomes after Monitoring, Surgery, or Radiotherapy for Prostate Cancer. *N Engl J Med*. 2016;375(15):1425-1437. doi:10.1056/NEJMoa1606221

6. Recommendations | Prostate cancer: diagnosis and management | Guidance | NICE. Accessed August 6, 2019. https://www.nice.org.uk/guidance/ng131/chapter/Recommendations

7. Dearnaley D, Syndikus I, Mossop H, et al. 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*. 2016;17(8):1047-1060. doi:10.1016/S1470-2045(16)30102-4

8. Lee WR, Dignam JJ, Amin M, et al. NRG Oncology RTOG 0415: A randomized phase III non-inferiority study comparing two fractionation schedules in patients with low-risk prostate cancer. *JCO*. 2016;34(2_suppl):1-1. doi:10.1200/jco.2016.34.2_suppl.1

9. Catton CN, Lukka H, Gu CS, et al. Randomized Trial of a Hypofractionated Radiation Regimen for the Treatment of Localized Prostate Cancer. *JCO*. 2017;35(17):1884-1890. doi:10.1200/JCO.2016.71.7397

10. Rodin D, Tawk B, Mohamad O, et al. Hypofractionated radiotherapy in the real-world setting: An international ESTRO-GIRO survey. *Radiotherapy and Oncology*. 2021;157:32-39. doi:10.1016/j.radonc.2021.01.003

11. Mohler JL, Antonarakis ES, Armstrong AJ, et al. Prostate Cancer, Version 2.2019, NCCN Clinical Practice Guidelines in Oncology. *Journal of the National Comprehensive Cancer Network*. 2019;17(5):479-505. doi:10.6004/jnccn.2019.0023

12. Dess RT, Suresh K, Zelefsky MJ, et al. Development and Validation of a Clinical Prognostic Stage Group System for Nonmetastatic Prostate Cancer Using Disease-Specific Mortality Results From the International Staging Collaboration for Cancer of the Prostate. *JAMA Oncology*. 2020;6(12):1912-1920. doi:10.1001/jamaoncol.2020.4922

13. Lalonde E, Ishkanian AS, Sykes J, et al. Tumour genomic and microenvironmental heterogeneity for integrated prediction of 5-year biochemical recurrence of prostate cancer: a retrospective cohort study. *The Lancet Oncology*. 2014;15(13):1521-1532. doi:10.1016/S1470-2045(14)71021-6

14. Roach M, Hanks G, Thames H, et al. 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*Biology*Physics*. 2006;65(4):965-974. doi:10.1016/j.ijrobp.2006.04.029

15. Tsiatis AA, Davidian M. Joint Modeling of Longitudinal and Time-to-Event Data: An Overview. *Statistica Sinica*. 2004;14(3):809-834. Accessed December 3, 2018. https://www.jstor.org/stable/24307417

16. Proust-Lima C, Taylor JMG. Development and validation of a dynamic prognostic tool for prostate cancer recurrence using repeated measures of posttreatment PSA: a joint modeling approach. *Biostatistics*. 2009;10(3):535-549. doi:10.1093/biostatistics/kxp009

17. Rizopoulos D, Molenberghs G, Lesaffre EMEH. Dynamic predictions with time-dependent covariates in survival analysis using joint modeling and landmarking. *Biometrical Journal*. 2017;59(6):1261-1276. doi:10.1002/bimj.201600238

18. Parr H, Hall E, Porta N. Joint models for dynamic prediction in localised prostate cancer: a literature review. *BMC Medical Research Methodology*. 2022;22(1):245. doi:10.1186/s12874-022-01709-3

19. Rizopoulos D, Taylor JMG, Van Rosmalen J, Steyerberg EW, Takkenberg JJM. Personalized screening intervals for biomarkers using joint models for longitudinal and survival data. *Biostatistics*. 2016;17(1):149-164. doi:10.1093/biostatistics/kxv031

20. Taylor JMG, Yu M, Sandler HM. Individualized Predictions of Disease Progression Following Radiation Therapy for Prostate Cancer. *JCO*. 2005;23(4):816-825. doi:10.1200/JCO.2005.12.156

21. Proust-Lima C, Taylor JMG, Williams SG, et al. 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*Biology*Physics*. 2008;72(3):782-791. doi:10.1016/j.ijrobp.2008.01.056

22. Pinheiro J, Bates D, DebRoy S, Sarkar D, R Core Team. *Nlme: Linear and Nonlinear Mixed Effects Models*.; 2021. https://CRAN.R-project.org/package=nlme

23. Therneau TM. *A Package for Survival Analysis in R*.; 2021. https://CRAN.R-project.org/package=survival

24. Rizopoulos D, Papageorgiou G, Afonso PM. *JMbayes2: Extended Joint Models for Longitudinal and Time-to-Event Data*.; 2022.

25. Rizopoulos D. Dynamic Predictions and Prospective Accuracy in Joint Models for Longitudinal and Time-to-Event Data. *Biometrics*. 2011;67(3):819-829. doi:10.1111/j.1541-0420.2010.01546.x

26. Austin PC, Harrell FE, Klaveren D van. Graphical calibration curves and the integrated calibration index (ICI) for survival models. *Statistics in Medicine*. 2020;39(21):2714-2742. doi:https://doi.org/10.1002/sim.8570

27. Fournier MC, Dantan E, Blanche P. An R2-curve for evaluating the accuracy of dynamic predictions. *Statistics in Medicine*. 2018;37(7):1125-1133. doi:10.1002/sim.7571

28. Putter H, Fiocco M, Geskus RB. Tutorial in biostatistics: competing risks and multi-state models. *Statistics in Medicine*. 2007;26(11):2389-2430. doi:10.1002/sim.2712

29. Tree A, Griffin C, Syndikus I, et al. Nonrandomized Comparison of Efficacy and Side Effects of Bicalutamide Compared With Luteinizing Hormone-Releasing Hormone (LHRH) Analogs in Combination With Radiation Therapy in the CHHiP Trial. *International Journal of Radiation Oncology*Biology*Physics*. Published online January 10, 2022. doi:10.1016/j.ijrobp.2021.12.160

30. Geara FB, Bulbul M, Khauli RB, et al. Nadir PSA is a strong predictor of treatment outcome in intermediate and high risk localized prostate cancer patients treated by definitive external beam radiotherapy and androgen deprivation. *Radiat Oncol*. 2017;12. doi:10.1186/s13014-017-0884-y

31. Crook JM, Tang C, Thames H, et al. A biochemical definition of cure after brachytherapy for prostate cancer. *Radiotherapy and Oncology*. 2020;149:64-69. doi:10.1016/j.radonc.2020.04.038

32. Davis JW, Kolm P, Wright GL, Kuban D, El -Mahdi Anas, Schellhammer PF. The durability of external beam radiation therapy for prostate cancer: can it be identified? *Journal of Urology*. 1999;162(3 Part 1):758-761. doi:10.1097/00005392-199909010-00036

33. Yock TI, Zietman AL, Shipley WU, Thakral HK, Coen JJ. Long-term durability of PSA failure-free survival after radiotherapy for localized prostate cancer. *International Journal of Radiation Oncology, Biology, Physics*. 2002;54(2):420-426. doi:10.1016/S0360-3016(02)02957-7

34. Ferrer L, Putter H, Proust-Lima C. Individual dynamic predictions using landmarking and joint modelling: Validation of estimators and robustness assessment: *Statistical Methods in Medical Research*. Published online November 22, 2018. doi:10.1177/0962280218811837

35. Prostate cancer survival statistics. Cancer Research UK. Published May 15, 2015. Accessed September 5, 2022. https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/prostate-cancer/survival

36. Taylor JMG, Park Y, Ankerst DP, et al. Real-Time Individual Predictions of Prostate Cancer Recurrence Using Joint Models. *Biometrics*. 2013;69(1):206-213. doi:10.1111/j.1541-0420.2012.01823.x

37. Blanche P, Proust-Lima C, Loubère L, Berr C, Dartigues JF, Jacqmin-Gadda H. Quantifying and comparing dynamic predictive accuracy of joint models for longitudinal marker and time-to-event in presence of censoring and competing risks. *Biometrics*. 2015;71(1):102-113. doi:10.1111/biom.12232

38. Andrinopoulou ER, Rizopoulos D, Takkenberg JJ, Lesaffre E. Combined dynamic predictions using joint models of two longitudinal outcomes and competing risk data. *Stat Methods Med Res*. 2017;26(4):1787-1801. doi:10.1177/0962280215588340

39. Klotz L, O'Callaghan C, Ding K, et al. Nadir Testosterone Within First Year of Androgen-Deprivation Therapy (ADT) Predicts for Time to Castration-Resistant Progression: A Secondary Analysis of the PR-7 Trial of Intermittent Versus Continuous ADT. *JCO*. 2015;33(10):1151-1156. doi:10.1200/JCO.2014.58.2973

40. Sumanasuriya S, Seed G, Parr H, et al. Elucidating Prostate Cancer Behaviour During Treatment via Low-pass Whole-genome Sequencing of Circulating Tumour DNA. *European Urology*. Published online June 5, 2021. doi:10.1016/j.eururo.2021.05.030

41. Wilkins AC, Gusterson B, Szijgyarto Z, et al. Ki67 Is an Independent Predictor of Recurrence in the Largest Randomized Trial of 3 Radiation Fractionation Schedules in Localized Prostate Cancer. *International Journal of Radiation Oncology • Biology • Physics*. 2018;101(2):309-315. doi:10.1016/j.ijrobp.2018.01.072

42. Thurtle DR, Greenberg DC, Lee LS, Huang HH, Pharoah PD, Gnanapragasam VJ. Individual prognosis at diagnosis in nonmetastatic prostate cancer: Development and external validation of the PREDICT Prostate multivariable model. *PLOS Medicine*. 2019;16(3):e1002758. doi:10.1371/journal.pmed.1002758

43. Shariat SF, Mw K, Aj V, Pi K, Pt S. Critical review of prostate cancer predictive tools. *Future Oncol*. 2009;5(10):1555-1584. doi:10.2217/fon.09.121

44. Tomer A, Nieboer D, Roobol MJ, Steyerberg EW, Rizopoulos D. Shared decision making of burdensome surveillance tests using personalized schedules and their burden and benefit. *Statistics in Medicine*. 2022;41(12):2115-2131. doi:10.1002/sim.9347

45. Brand DH, Tree AC, Ostler P, et al. 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*. 2019;20(11):1531-1543. doi:10.1016/S1470-2045(19)30569-8

46. Denham JW, Joseph D, Lamb DS, et al. 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*. 2019;20(2):267-281. doi:10.1016/S1470-2045(18)30757-5

47. Vickers AJ, van Calster B, Steyerberg EW. A simple, step-by-step guide to interpreting decision curve analysis. *Diagnostic and Prognostic Research*. 2019;3(1):18. doi:10.1186/s41512-019-0064-7

Figure captions

PSA distribution and reverse-year plot

Figure 1 – top (a): aggregated PSAs and boxplots by year and outcome since starting treatment. Patient numbers still at risk are presented below the plot. Bottom (b): Smoothed reverse-years PSA trajectory plot, stratified by outcome, natural cubic spline smoothers shown. In the non-recurring patients, a few PSAs >5ng/mL are recorded; these PSAs were considered bounces/flares and therefore did not achieve the protocol's definition of biochemical failure.

Effect plots

Figure 2 –The predicted effect plots of PSA, stratified by outcome (solid – recurrence, dashed – censored) and each baseline subgroup over time. The top-left panel are the overall PSA trajectories by outcome. The natural cubic spline smoother is depicted.

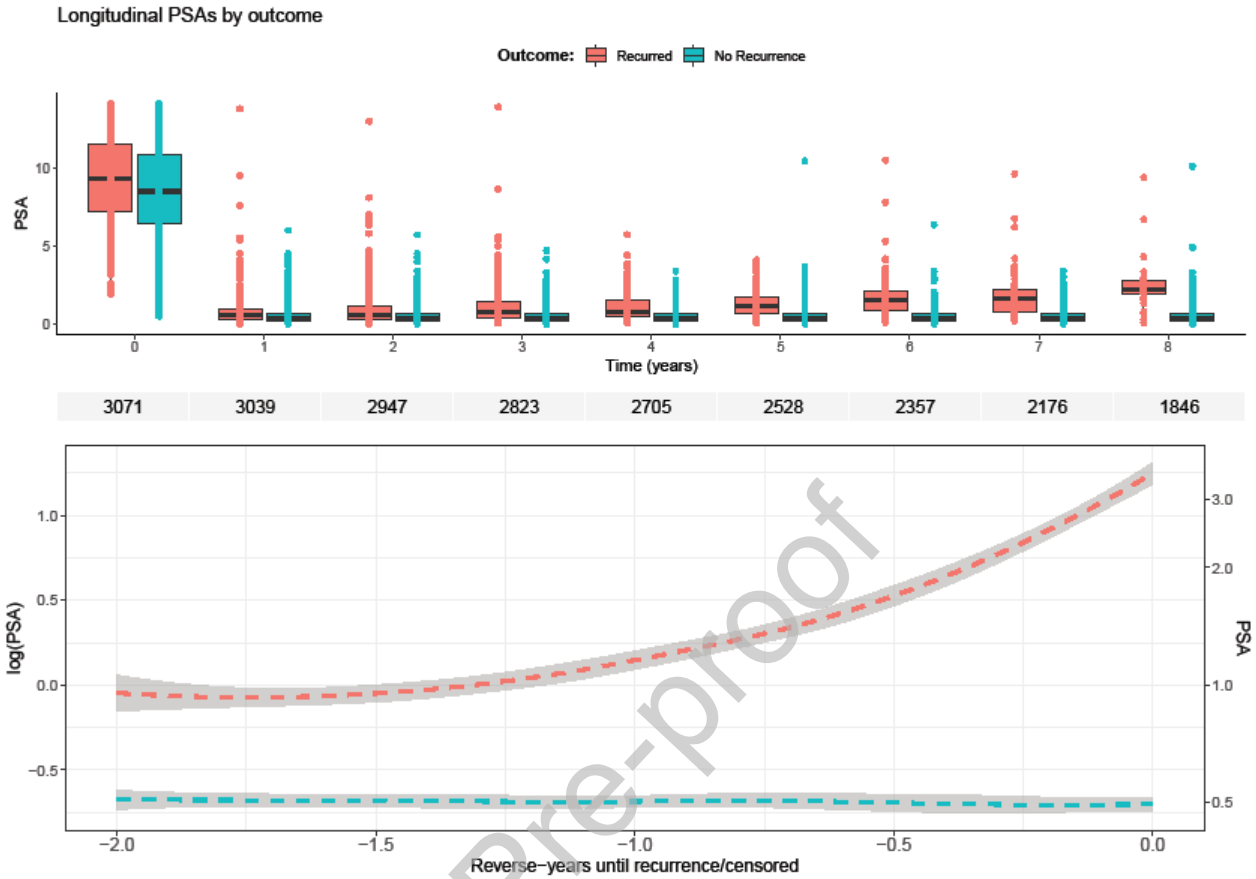
Dynamic predictions

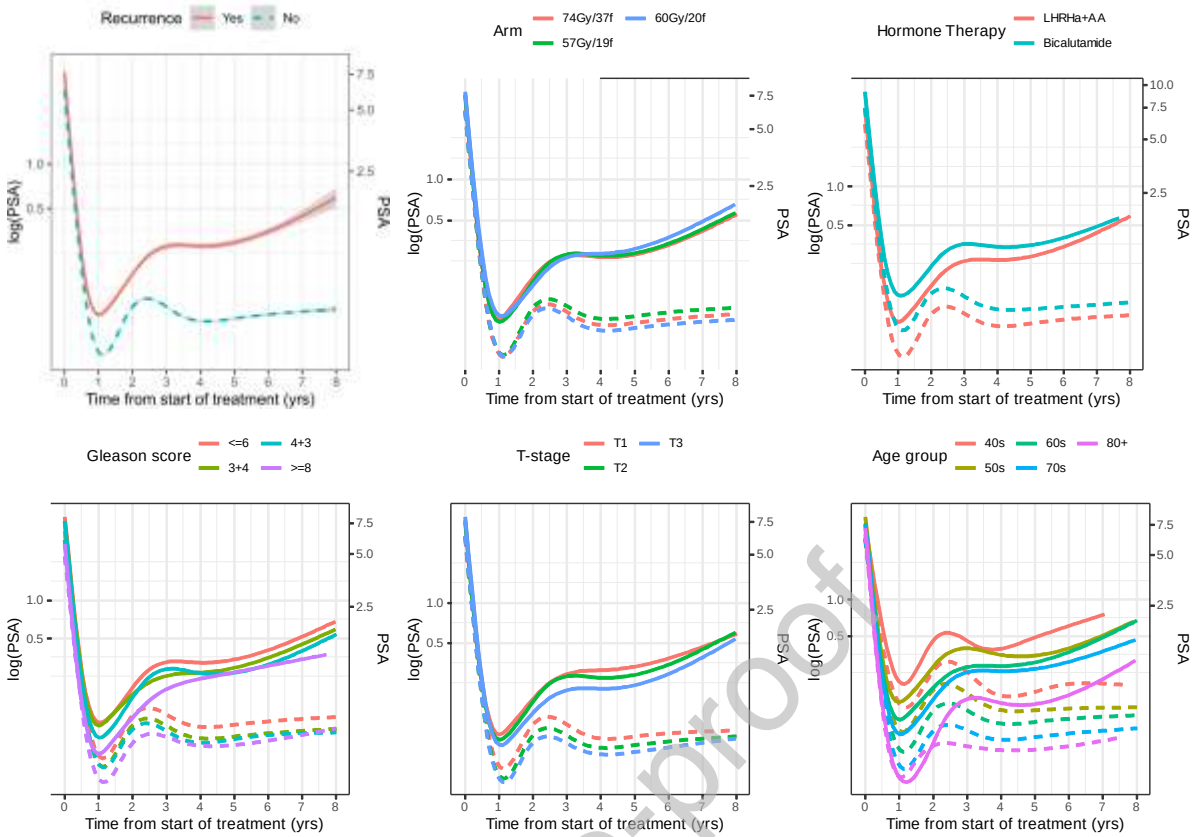
Figure 3 - Dynamic predictions of two patients: A & B, over five panels (V—Z). Patients A & B are ages 63 and 64 respectively and both received the same treatments, with contrasting prognostic factors. The left-hand side of each plot shows their modelled PSA values over time and the right-hand side shows their risk of

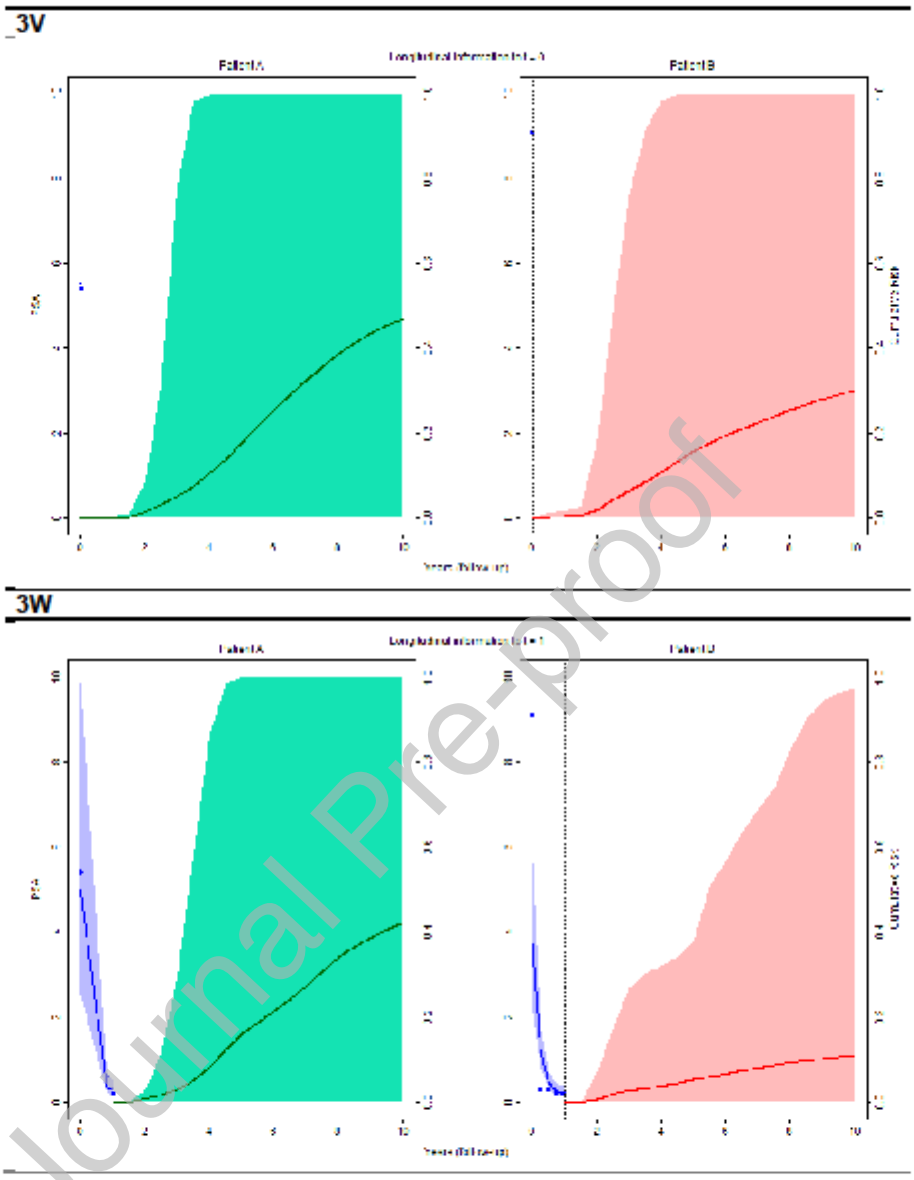
recurrence at particular landmarks by ten years after initiating treatment. The 95% credible intervals are shown (shaded).

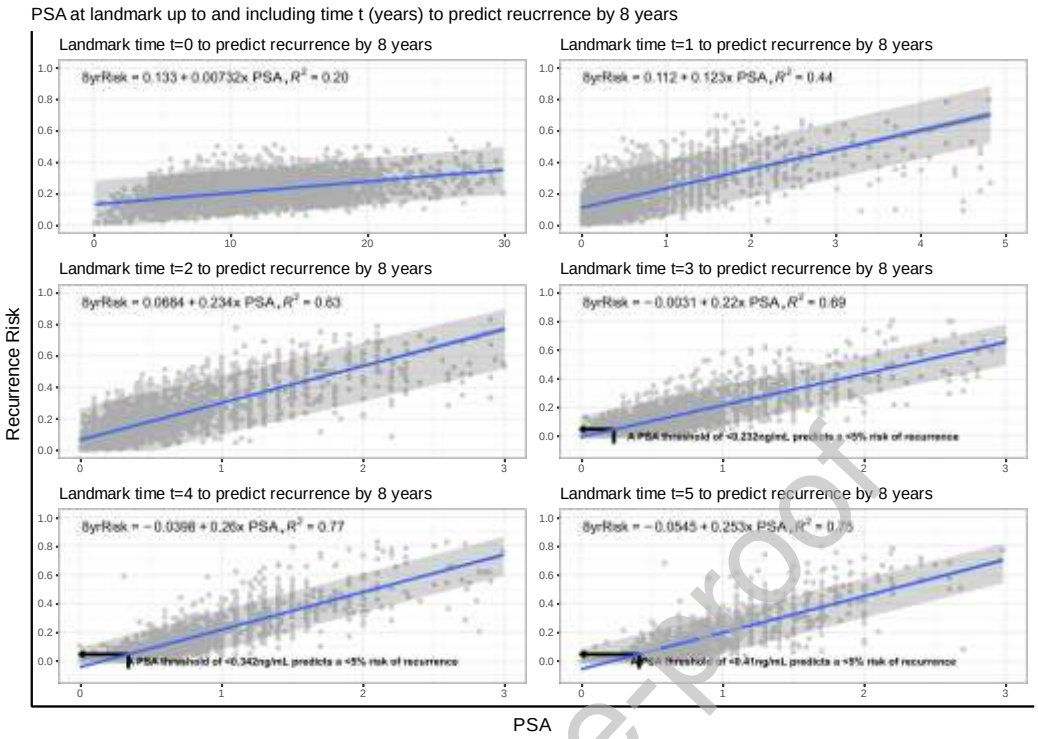
Scatter plots of PSA predicting risk

Figure 4 – Scatter plots of PSA predicting prognosis/recurrence risk by 8 years (horizon), each panel represents landmarks 0 – 5 years. Each grey dot indicates a patient's PSA (nearest to that landmark time) and risk at each landmark time. PSAs ≤ 3ng/mL are considered after t=1. The blue line indicates regression fit with the corresponding equation and R^2 labelled in each panel, with 95% confidence intervals. The wider grey bands indicate 95% prediction intervals. At the intercept (or less) indicates the predicted recurrence risk for a nil PSA; for the regression lines at t=3,4,5, each PSA threshold is labelled that predicts a <5% risk.









Tables

Table 1 - Baseline characteristics, follow-up time (N=3071) considered in model development. LHRHa – Luteinizing-Hormone-Releasing-Hormone analogue + possible anti-androgen.

Baseline Factors	**N = 3071[1]** **[1] n (%); Median (IQR)**
Allocated fractionation group	
74Gy/37f	1017 (33%)
57Gy/19f	1025 (33%)
60Gy/20f	1029 (34%)
Gleason score	
≤ 6	1022 (33%)
$3+4$	1354 (44%)
$4+3$	598 (19%)
≥ 8	97 (3%)
Clinical T-stage	
T1	1088 (35%)
T2	1713 (56%)
T3	270 (9%)
Hormone Therapy	
LHRHa	2668 (87%)
150mg Bicalutamide	403 (13%)
Age (years)	69.1 (64.5, 73.2)
Baseline/presenting PSA (ng/mL)	10.3 (7.3, 14.6)

Table 2 – optimism-corrected model metrics from landmark times t=0—7 predicting at a horizon time of eight years. Discrimination – AUC (area under the curve); calibration – ICI (integrated calibration index); overall predictive performance – (Brier score). Mean, [95% bootstrapped CIs] refers to the bootstrapped replications. The ICI & Brier are loss functions (where lower is better), with higher AUC measures indicating better discrimination. Ns are patients remaining at risk at development.

		Optimism-corrected metrics		
Landmark t_{years} for prediction interval [t, 8]	**N still at risk**	**AUC**	**ICI**	**Brier**
$t = 0$ ***(baseline)***	3071	0.525 [0.500—0.553]	0.056 [0.043—0.068]	0.16 [0.154—0.166]
$t = 1$	3039	0.58 [0.556—0.6]	0.06 [0.045—0.072]	0.153 [0.147—0.16]
$t = 2$	2947	0.612 [0.583—0.644]	0.083 [0.069—0.098]	0.153 [0.145—0.16]
$t = 3$	2823	0.651 [0.632—0.677]	0.061 [0.049—0.069]	0.123 [0.113—0.132]
$t = 4$	2705	0.748 [0.728—0.767]	0.045 [0.036—0.052]	0.097 [0.089—0.106]
$t = 5$	2528	0.797 [0.767—0.821]	0.038 [0.031—0.048]	0.068 [0.062—0.075]
$t = 6$	2357	0.838 [0.807—0.868]	0.024 [0.019—0.029]	0.047 [0.039—0.054]
$t = 7$	2176	0.806 [0.756—0.873]	0.016 [0.013—0.019]	0.027 [0.022—0.033]