Individual prognosis at diagnosis in non-metastatic prostate cancer: Development and external validation of the PREDICT *Prostate* multivariable model

David R Thurtle[1,2*], David C Greenberg[3], Lui S Lee[4], Hong H Huang[4], Paul D Pharoah[5†] & Vincent J Gnanapragasam[1,2,6†*]

1. Academic Urology Group, Department of Surgery, University of Cambridge, Cambridge, UK
2. Department of Urology, Cambridge University Hospitals NHS Foundation Trust, Cambridge, UK
3. National Cancer Registration and Analysis Service [Eastern Region], Fulbourn, Cambridge, UK
4. Department of Urology, Singapore General Hospital, Singapore
5. Centre for Cancer Genetic Epidemiology, Department of Oncology, University of Cambridge, Cambridge, UK
6. Cambridge Urology Translational Research and Clinical Trials, Cambridge, UK

† Joint senior authors

* dt433@cam.ac.uk (DT); * vjg29@cam.ac.uk (VG)

Full title:

Individual prognosis at diagnosis in non-metastatic prostate cancer: Development and external validation of the PREDICT *Prostate* multivariable model

Short title:

PREDICT *Prostate*: An individual prognostic model for prostate cancer

Abstract

Background

Prognostic stratification is the cornerstone of management in non-metastatic prostate cancer (PCa). However, existing prognostic models are inadequate – often using treatment outcomes rather than survival, stratifying by broad heterogeneous groups and using heavily treated cohorts. To address this unmet need, we developed an individualised prognostic model which contextualizes PCa-specific mortality (PCSM) against other cause mortality, and estimates the impact of treatment on survival.

Methods and findings

Using records from the UK National Cancer Registration and Analysis Service, data were collated for 10,089 men diagnosed with non-metastatic PCa between 2000 and 2010 in Eastern England. Median follow-up was 9.8 years with 3,829 deaths (1,202 PCa-specific). 19.8%, 14.1%, 34.6% and 31.5% of men underwent conservative management, prostatectomy, radiotherapy and androgen deprivation monotherapy respectively. 2,546 men diagnosed in Singapore over a similar time period represented an external validation cohort.

Data were randomly split 70:30 into model development and validation cohorts. 15-year PCSM and non-prostate cancer mortality (NPCM) were explored using separate multivariable Cox models within a competing risks framework. Fractional polynomials were utilised to fit continuous variables and baseline hazards. Model accuracy was assessed by discrimination and calibration using Harrell's C-index and chi-squared goodness-of-fit respectively within both validation cohorts.

A multivariable model estimating individualised 10 and 15-year survival outcomes was constructed combining age, PSA, histological grade, biopsy core involvement, stage, and primary treatment which were each independent prognostic factors for PCSM; and age and comorbidity which were prognostic for NPCM. The model demonstrated good discrimination with C-index of 0.84 (95%CI: 0.82-0.86) and 0.84 (95%CI: 0.80-0.87) for 15-year PCSM in the UK and Singapore validation cohorts respectively, comparing favourably to international risk-stratification criteria. Discrimination was maintained for overall mortality with C-index 0.77 (95%CI: 0.75-0.78) and 0.76 (95%CI: 0.73-0.78). The model was well-calibrated with no significant difference between predicted and observed PCa-specific (p=0.19) or overall deaths (p=0.43) in the UK cohort.

Key study limitations were a relatively small external validation cohort, an inability to account for delayed changes to treatment beyond 12 months and an absence of t-stage sub-classifications.

Conclusions

'PREDICT *Prostate*' is an individualised multivariable PCa prognostic model built from baseline diagnostic information and the first to our knowledge which models potential treatment benefits on

overall survival. Prognostic power is high despite using only routinely collected clinico-pathological information.

Author Summary

Why was this study done?

- Among men with non-metastatic prostate cancer a number of treatment options are often appropriate, including surveillance or conservative management.
- Problems of both over-treatment of indolent disease and under-treatment of aggressive disease are both recognised. Many men also suffer lifelong side-effects from a treatment they may not have not needed.
- Estimating prognosis is therefore of crucial importance to inform decision-making on the benefits of treatments at the point of diagnosis. However, existing risk models are inadequate, rarely use survival as an outcome, ignore non-cancer mortality, and often group patients into broad categories. As a result no model is yet to be formally endorsed or widely used in clinical practice.
- In this study we sought to create an individualised model that addresses these gaps and predicts both cancer-specific and overall survival at the point of diagnosis, and which estimates the potential survival benefit of treatment.

What did the researchers do and find?

- We studied a large UK dataset of over 10,000 men diagnosed with non-metastatic prostate cancer and long-term survival information. The dataset was randomly split into model development and validation datasets. An additional dataset of over 2500 men diagnosed in Singapore was used for additional external validation.
- Using Cox regression and fractional polynomials, models were built for 15-year prostate cancer specific mortality and non-prostate cancer mortality using patient and tumour characteristics routinely available at diagnosis. These two models were then adjusted for competing risks to predict overall mortality.
- We found that the new risk model, called 'PREDICT *Prostate*' predicted survival outcomes with a high degree of accuracy in both validation sets with concordance indices up to 0.84.
- We have now incorporated the model into a web-based interface for easy access and utility.

What do these findings mean?

- To our knowledge, we present the first individualised multivariable survival model for non-metastatic prostate cancer built and validated in an unscreened, pre-treatment cohort.
- Our findings need further validation in independent datasets, and may be limited by a relatively small external validation cohort.
- This tool incorporates the impact of radical therapy, which allows comparison to be made against the option of conservative management within the context of an individual's competing risks, to inform decision-making around management.
- The model does not require any additional tests beyond standard of care, and is freely available for use. It's primary application is among men deciding between conservative management and radical treatment – where decision dilemmas are most acute.

- The model has the potential to enable well-informed and standardised decision-making and reduce both over- and under-treatment.

Introduction

Prostate cancer (PCa) is the commonest cancer affecting males and a leading cause of cancer-related morbidity[1]. The vast majority of these new presentations are with localised or locally advanced disease, representing a significant healthcare and economic burden [2]. Treatment decisions are notoriously complex with the risk of cancer related mortality balanced against the potential morbidity associated with treatment as well as competing mortality risks. Estimating prognosis within these contexts is therefore highly important, with over 40,000 consultations for newly diagnosed PCa every year in the UK alone[2]. This importance has been underlined by randomised trial evidence reporting non-inferiority of conservative management compared to radical therapy in many early cancers from the American PIVOT and UK ProtecT trials[3,4].

Despite this importance, there are no high quality individualised prognostic models available for clinical counselling and decision-making. Instead, tiered stratification systems are used that categorise men into different levels of risk. These models are widely endorsed by national and international guideline groups but are often derived using inadequate surrogate endpoints, such as PSA resurgence after treatment, rather than being calibrated against mortality[5,6]. Modern extensions to these models have now sought to validate performance against cancer mortality and have extended the number of sub-classifications[7-10]. Although these extensions add granularity they remain too heterogeneous for modern individualised medicine approaches. More recent attempts at developing survival models have focussed solely on men undergoing radical treatment, and have not been appropriately validated[11,12]. The inadequacies of existing models are evident by the fact that the American Joint Committee on Cancer (AJCC) have not endorsed a single prognostic model for non-metastatic PCa[13].

The objectives of this study were to develop and validate an individualised prognostic model for non-metastatic PCa. Our aim was to produce a model that was able to contextualise the relative PCa-specific and overall survival outcomes for an individual with newly diagnosed disease and allow

modelling of the potential benefit of treatment on these outcomes. Study design and reporting was informed by the AJCC criteria for model adoption and the TRIPOD statement respectively[14,15].

Methods

This study is reported throughout as per the Transparent Reporting of a multivariable Prediction model for individual prognosis or Diagnosis (TRIPOD) guideline (S1 Checklist).

Study population and definition of variables

Fully-anonymised data were retrieved from Public Health England after review by the Office for Data Release(ODR1617/171). Following approvals, Cambridge University Hospitals NHS Trust acted as host institution for data receipt. Information on all men diagnosed with non-metastatic PCa in secondary care in Eastern England, UK, between 2000 and 2010 was collected prospectively by the National Cancer Registration and Analysis Service [NCRAS] Eastern Region. The cohort derivation has been previously described[16]. Men with recorded nodal or metastatic disease at diagnosis were excluded, along with men diagnosed only by endoscopic resection and any remaining men with PSA ≥100ng/ml as a surrogate for occult metastatic disease[17]. Only men with intact information on key candidate predictors – age, PSA (ng/ml), histological grade group, T-stage and primary treatment were included. From a potential cohort of 15,335 men, 5,246 (34.2%) were excluded for missing information in at least one of these variables leaving a final analytic cohort of 10,089. Comorbidity scores, derived from inpatient hospital episode statistics (HES) data were also included. These are based on clinical coding of inpatient episodes in the period between 27 and 3 months before PCa diagnosis, thus excluding PCa from any comorbidity score. Vital status was ascertained at the end of March 2017 with all analyses censored at the end of September 2016 to allow for a lag-time of up to 6 months for non-cancer deaths through the National Health Service Strategic Tracing Service. Death was considered PCa-specific when PCa was listed in 1a, 1b or 1c of the death certificate.

Potential variables entered into the primary model were age, PSA, clinical T-stage, histological grade, ethnicity, comorbidity and primary treatment type. Information from NCRAS was that recorded at the time of diagnosis. T-stage was simplified to T1, T2, T3 or T4 as subcategories were rarely available and have limited impact in determining prognosis[18]. Histological grade groups (1-5) were

used([19]. PSA (ng/ml) refers to the value at diagnosis, prior to biopsy or treatment. Primary treatment refers to the first definitive treatment the patient received in the first 12 months. Here we have used the term conservative management to cover active surveillance and watchful waiting as registry data did not discriminate between the two during this time period. As previously published, the majority of men receiving radiotherapy (RT) in this period were on concomitant hormone therapy which represents current best practice for this treatment modality[20].

Model Development

The primary (UK) cohort was split randomly in a 70:30 ratio into model development (n=7062) and validation cohorts (n=3027) (Table 1). Within the development cohort separate models were built for PCa-specific mortality (PCSM) and non-PCa mortality (NPCM). The general approach to modelling was similar to that used for the PREDICT breast cancer prognosis and treatment benefit model[21]. Cox proportional hazards models were utilised to estimate hazard ratios associated with each candidate predictor. Follow up time was censored at time to death, time to last follow up or 15 years, which ever came first. Each variable was assessed through uni- and multi-variable analysis with the proportional hazards assumption tested. A backwards elimination technique was used for variable selection with a 5% significance level. Risk-relationships between continuous variables were modelled using multivariable fractional polynomials, with continuous data retained wherever possible to maximise predictive information. T-stage, histological grade group, and primary treatment type were modelled as factor variables. Radical treatments (radiotherapy (RT) or radical prostatectomy (RP)) were combined, as explained later. After fitting the multi-variable models, smoothed functions for the baseline hazard of PCSM and NPCM were calculated. The baseline cumulative hazard was estimated for each patient, then the logarithmic value of the baseline hazard was regressed against time using a univariate fractional polynomial function[21].

Competing risks adjustment

Beta coefficients for each prognostic factor in the two Cox models were used to derive a prognostic

index for PCSM (piPCSM) and NPCM (piNPCM) for each patient. The absolute risk (hazard(H)) of PCa death (H_{PCa}) and non-PCa (H_{NPC}) death until time *t*, if there were no competing mortalities, are estimated by the following formulae respectively: $H_{PCa} = 1 – \exp(-\exp(piPCSM)*bhPCSM(t))$ and $H_{NPC} = 1 – \exp(-\exp(piNPCM)*bhNPCM(t))$. Where bhPCSM(*t*) and bhNPCM(*t*) are the cumulative baseline hazards of PCSM or NPCM at time *t* respectively. However, as these risks compete against each other, the cumulative risk (R) of overall mortality (OM) at time *t* is : $R_{OM}(t) = 1 – (1-H_{PCA}(t))*(1-H_{NPC}(t)$. Therefore the formulae for cumulative risk (R) of PCa death and non-PCa death at time *t* are: $R_{PCa}(t) = R_{OM}(t) * (H_{PCa}(t) /(H_{PCa}(t)+H_{NPC}(t))$ and $R_{NPC}(t) = R_{OM}(t) * (H_{NPC}(t) /(H_{NPC}(t)+H_{PCa}(t))$ respectively. The source code for replicating the model's output has been made available online, including this competing risk adjustment.

Model accuracy and comparison to existing models

Model calibration and goodness-of-fit was investigated in the UK validation cohort by comparing observed and predicted deaths within quintiles of predicted mortality and within strata of other prognostic variables. For assessing calibration, we integrated the predicted outcomes across all follow-up times to allow for cases with follow-up of less than 10 or 15 years. Thus the calibration corresponds to a range of different follow-up times. A simplified χ^2 goodness-of-fit (GOF) test was performed using the method of May and Hosmer, whereby a p value of less than 0.05 would suggest a significant difference between the expected and observed number of events, assessed up to 10 years or 15 years[22]. Calibration curves were also visually assessed. Model discrimination was evaluated by estimating 10 and 15-year cumulative mortality risk. Harrell's concordance statistic (C-index) was then calculated for PCa-specific, non-PCa and overall deaths. This accounts for right-censored data, i.e. cases with less than 10 or 15 years follow-up respectively. All analyses were performed using Stata 14 (StataCorp, College Station, TX, USA), with the exception of C-index which was performed using 'rcorr.cens' within the 'Hmisc' package of R[23].

Comparisons against existing models were made by calculating C-indices for 3 well-known tools used at the point of diagnosis internationally – namely the UCSF Cancer of the Prostate Risk Assessment (CAPRA) score, the updated NCCN criteria and the three-tier EAU criteria [7,10,24]. Available information was used to calculate these with no imputation of missing data. Where T stage sub-classification was unknown, integer T stages were used.

External validation

External validation of the model was assessed using a geographically and ethnically independent cohort of men from Singapore General Hospital, diagnosed between 1990 and 2015 which has been previously described[25]. The same inclusion criteria were applied as to the model development dataset. From a potential cohort of 3245, 699 (21.5%) were excluded for missing information. 310 cases had missing data for key candidate predictors, and no follow-up was available for a further 389 men, leaving a final analysable cohort of 2,546 (Table 1). Data amongst this cohort had been recorded on a prospective basis including the same parameters, defined identically as the primary cohort with the addition of biopsy information, but did not include comorbidity information. NPCM estimates therefore assumed the same prevalence of comorbidity as the primary dataset (10.21%) spread evenly across the cohort. Vital status was ascertained via the Singapore Ministry of Home Affairs, using the same definitions for cause of death, with data censored 30th June 2017. Model performance was assessed using the methods described above. Ethics for use of these data is covered by ref. 2009/1053/D approved by the SingHealth Centralised Institutional Review Board.

Inclusion of biopsy information as a variable

Previous risk criteria have included diagnostic biopsy information as a potentially important prognostic variable. To investigate this we undertook an additional sub-cohort analysis on men diagnosed at one hospital within our cohort (n=1451) for whom biopsy characteristics were available. For this we used percentage of positive cores (PPC = number of cores positive for

228 cancer/total number of cores taken). PPC was regressed against PCSM, offset against all parameters
229 within the base model. PPC was modelled continuously and categorically. Likelihood ratio χ^2tests,
230 Akaike(AIC) and Bayesian information criterion(BIC) were used to determine best fit. The eventual
231 parameter was weight-adjusted and incorporated in to the model (Tables F and G in S1 Appendix).
232 Performance of the extended model, including the PPC parameter, was then assessed within the
233 Singaporean cohort using the same methodology as above.

Results

Participants

The model development cohort consisted of 7,063 men; 842 and 1,821 men died from PCa and other causes within 15 years respectively. The UK validation cohort consisted of 3,026 men; 360 and 806 died from PCa and other causes respectively. Median follow-up was 9.8 years for both cohorts with 82,887 person-years of follow-up in total (Table 1). Importantly, the UK cohort included significant numbers of patients who had undergone conservative management (n=1997). Only 114 (5.7%) of these men converted to radical treatment over total study follow-up. Trends across the inclusion period, including increased proportions of T1 disease and increasing uptake of conservative management have been identified previously(16, 20).

		Total UK Cohort		UK Model Development Cohort		UK Validation Cohort		Singapore Validation cohort	
Total Subjects		10,089		7,063		3026		2546	
Time at risk (years)		82,887		58,138		24,750		13,416	
Median follow-up (years)		9.8	Range 0-16	9.8	Range 0-16	9.8	Range 0-16	5.1	Range 0-26
10 year outcomes:			%		%		%		%
PCa deaths		1030	10.2	712	10.1	317	10.5	105	4.1
Non PCa deaths		2246	22.3	1555	22.0	691	22.8	225	8.8
Any-cause death		3276	32.5	2267	32.1	1008	33.3	330	13.0
Observations censored before 10 years		3770	37.4	2667	37.8	1103	36.5	1930	75.8
15-year outcomes:									
PCa deaths		1202	11.9	842	11.9	360	11.9	133	5.2
Non PCa deaths		2627	26.0	1821	25.8	806	26.6	283	11.1
Any-cause death		3829	38.0	2663	37.7	1166	38.5	416	16.3
Observations censored before 15 years		6000	59.5	4212	41.7	1788	59.1	2063	81.0
Crude PCS mortality rate (per patient year)		1.46		1.46		1.46		0.99	
Annual overall mortality rate (per patient year)		4.64		4.6		4.72		3.1	
Age (mean, SD)		69.9	8.30	69.9	8.34	69.9	8.29	66.1	7.96
PSA (mean, SD)		18.4	17.5	18.5	17.5	18.2	17.6	15.7	16.6
Gradegroups			%		%		%		%
	1	3328	33.0	2317	32.8	1011	33.4	1126	44.2
	2	3017	29.9	2125	30.1	892	29.5	723	28.4
	3	1486	14.7	1057	15.0	429	14.2	326	12.8
	4	1032	10.2	710	10.1	322	10.6	170	6.7
	5	1226	12.2	854	12.1	372	12.3	201	7.9
Tumour-stage									
	1	5421	53.7	3761	53.2	1660	54.9	1625	63.8
	2	3213	31.8	2270	32.1	943	31.2	660	25.9
	3	1378	13.7	977	13.8	401	13.3	244	9.6
	4	77	0.8	55	0.8	22	0.7	17	0.7
Primary Treatment									
Radical Prostatectomy		1419	14.1	995	14.1	424	14.0	1012	39.7
Radiotherapy		3495	34.6	2457	34.8	1038	34.3	823	32.3
Hormone Monotherapy		3178	31.5	2226	31.5	952	31.5	164	6.4
Conservative Management		1997	19.8	1385	19.6	612	20.2	538	21.1
Missing		na		na		na		9	0.4
Ethnicity									
White		7804	77.4	5464	77.4	2340	77.3	36	1.4
Missing/unknown		2136	21.2	1491	21.1	641	21.3	0	0.0
Asian		50	0.5	35	0.5	15	0.5	2435	95.6
Other		99	1.0	108	1.5	26	0.9	73	2.9

Table 1 Baseline cohort characteristics in the UK cohort overall, model development and validation cohorts and the external Singapore cohort.

PCa = prostate cancer SD= standard deviation

Model development and specification

Age, PSA, histological grade group, clinical stage and primary treatment type were all independent predictors for PCSM in the development cohort (Table 2). Comorbidity had a predictive effect in relation to NPCM but not PCSM. Age was also independently prognostic for NPCM. In the final model, comorbidity was modelled as a binary variable (0 or ≥1). The hazard ratios and fractional polynomial (FP) functions for prognostic factors in the final model are shown in Table 2. Associated FP functions for age and PSA are plotted in Fig 1. These allow more flexibility in relationships for continuous variables. The estimated baseline survival functions for PCSM and NPCM are recorded in S1 Appendix, and plotted against actual baseline PCSM and NPCM in Fig E in S1 Appendix.

	Prostate Cancer Specific Mortality		
	HR	95%CI	P
Age FP (age/10)^3 -341.16	1.003	1.002-1.003	<0.001
PSA FP ln((psa+1)/100)+1.6364	1.204	1.092-1.328	<0.001
Grade group			
1	1.00	-	-
2	1.32	1.06-1.65	0.014
3	1.73	1.36-2.19	<0.001
4	2.10	1.63-2.69	<0.001
5	3.93	3.15-4.89	<0.001
T stage			
1	1.00	-	-
2	1.18	1.01-1.37	0.042
3	1.49	1.23-1.80	0.000
4	1.88	1.14-3.13	0.014
Primary Treatment			
Conservative management	1.00	-	-
Radical treatment (RP/RT)	0.50	0.38-0.67	<0.001
Hormone monotherapy	2.48	1.92-3.20	<0.001
	Non Prostate Cancer Mortality		
Age FP age-69.87	1.13	1.12-1.14	<0.001
Comorbidity Score			
1+	1.89	1.67-2.14	<0.001

Table 2 The hazard ratios and p values of the variables included in each of the prostate cancer specific mortality and non-prostate cancer mortality models.

FP = fractional polynomial HR = hazard ratio CI = confidence interval

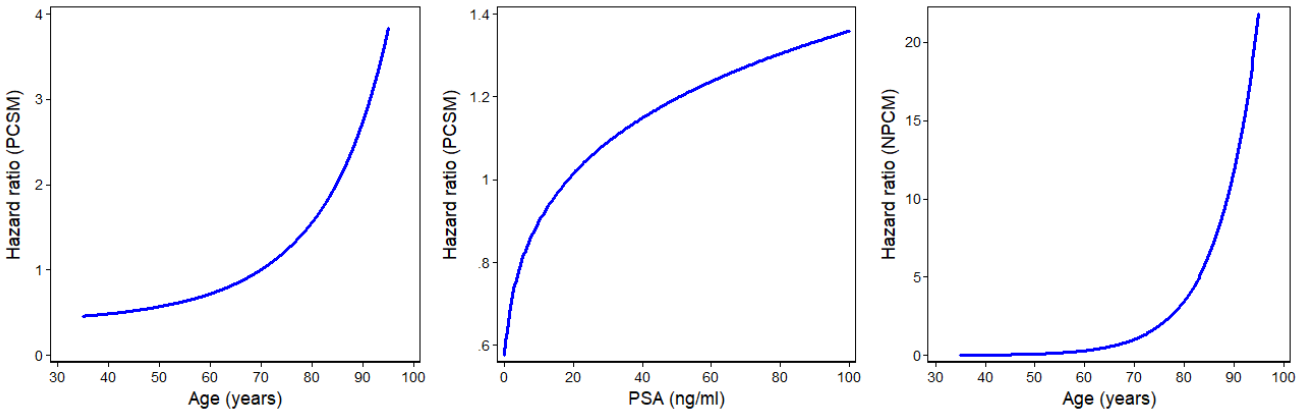


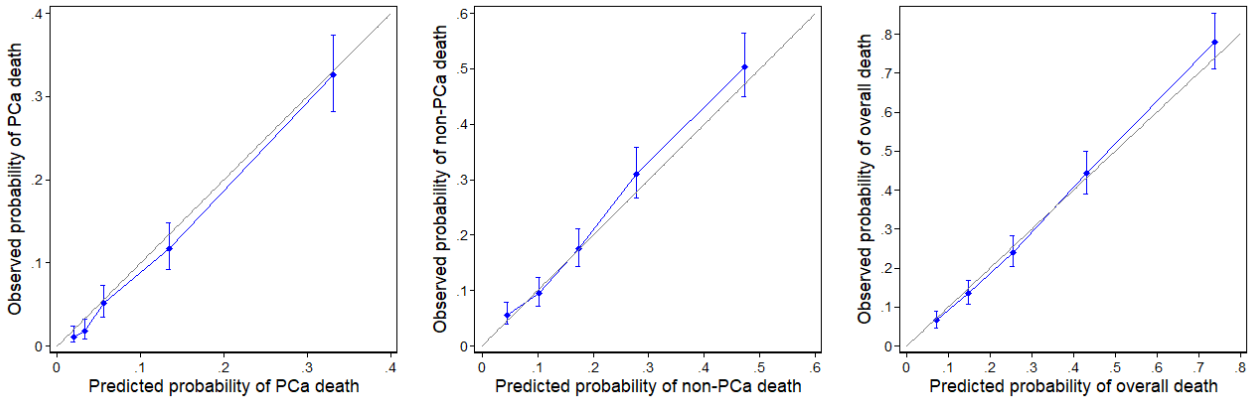
Figure 1 Prostate cancer-specific mortality (PCSM) hazard ratio functions for age (left) and PSA (centre), and non-PCa mortality (NPCM) hazard ratio function for age (right). Each derived from the model development data.

UK validation

The model was well-calibrated within the East of England validation cohort with absolute differences between observed and predicted PCa-specific and overall deaths less than 1% at 10 years (Table 3). The GOF tests suggested the model fitted well across different quintiles of risk, as shown by the calibration curves (Fig 2) with no significant difference in observed and predicted PCa-specific (p=0.19) or overall deaths (p=0.43) over 10 years (Table 3). Model discrimination was good, particularly for PCa-specific mortality, with C-index 0.84 (95%CI 0.82-0.86) and 0.84 (95%CI: 0.82-0.86) over 10 and 15 years follow up respectively (Table 3). Within the UK cohort, model discrimination was superior (p<0.001) to the current EAU, NCCN and CAPRA risk-stratification criteria for both PCSM and overall mortality (Table 4).

	Predicted	Observed	Difference (%)	χ^2 GOF p value	C-index	95%CI
10 years follow-up						
PCa Deaths	343	317	-0.86	0.19	0.84	0.82-0.86
Non-PCa deaths	641	691	1.65	0.19	0.74	0.72-0.77
Overall deaths	986	1008	0.73	0.43	0.77	0.75-0.78
15 years follow-up						
PCa Deaths	413	360	-1.75	0.04	0.84	0.82-0.86
Non-PCa deaths	751	806	1.82	0.02	0.71	0.69-0.72
Overall deaths	1165	1166	0.03	0.63	0.77	0.75-0.78

Table 3 Observed and predicted deaths over 10 and 15 years in the UK validation cohort (n=3026). Goodness of fit (GOF) and C-index are shown for each cause of death.



282

283 **Figure 2** Calibration curves comparing observed and predicted probability of prostate cancer(PCa) (left), non-
284 PCa (centre) and overall (right) deaths at 10 years by quintile of risk within the UK validation cohort.

285

	PCSM			Overall Mortality		
Model	C-index	95% CI	p	C-index	95% CI	p
PREDICT	0.843	0.824-0.862	-	0.766	0.753-0.780	-
EAU	0.688	0.665-0.711	<0.001	0.628	0.613-0.643	<0.001
NCCN	0.720	0.695-0.744	<0.001	0.644	0.628-0.659	<0.001
CAPRA	0.754	0.728-0.779	<0.001	0.656	0.640-0.672	<0.001

Table 4 Discrimination of the model, compared to other existing models amongst the UK validation cohort over 15 years maximum follow-up (n=3026).

EAU = European Association of Urology NCCN = National Comprehensive Cancer Network CAPRA = Cancer of the Prostate Risk Assessment (UCSF)

Calibration remained good across various sub-categories of patients, as demonstrated in Table C in S1 Appendix. Importantly, predictions for both PCa and non-PCa deaths amongst men undergoing either conservative management or radical therapy were within 2%. The GOF tests amongst this treatment sub-cohort continued to demonstrate no significant difference between predicted and observed PCa-specific (p=0.23) or overall deaths (p=0.11) over 10 years.

External Validation

Accuracy of the model, was also assessed using the Singaporean cohort (n=2,546). Here, median follow-up was 5.1 years, with 133 and 283 PCa and non-PCa deaths respectively (Table 1).

Model discrimination amongst this cohort was promising with C-index 0.83 (95%CI: 0.79-0.87) and 0.76 (95%CI 0.73-0.78) for PCSM and overall mortality respectively over 10 years (Table 5). Differences between observed and predicted deaths were less than 1% over 10 and 15-years, albeit within a small cohort (Table 5). GOF analysis showed no significant differences between observed and predicted non-PCa deaths, but the model appeared to slightly underestimate PCSM and overall deaths (Table 5 and Fig F in S1 Appendix). Within this external cohort, our baseline model performed better than the 3 tested comparators in predicting overall mortality (P<0.001) (Table D in S1 Appendix). Discrimination for PCSM was improved compared to the EAU stratification criteria, but not significantly better than the NCCN or CAPRA scores.

	Predicted	Observed	Difference (%)	GOF p value	C-index	95%CI
10 years follow-up						
PCa Deaths	89	105	0.63	0.01	0.83	0.79-0.87
Non-PCa deaths	236	225	-0.43	0.10	0.74	0.70-0.77
Overall deaths	325	330	0.20	0.01	0.76	0.73-0.78
15 years follow-up						
PCa Deaths	112	127	0.59	0.00	0.82	0.78-0.86
Non-PCa deaths	279	273	-0.24	0.08	0.72	0.69-0.76
Overall deaths	391	400	0.35	0.01	0.75	0.72-0.78

310 **Table 5** Observed and predicted deaths over 10 and 15 years in the Singaporean validation cohort (n=2546).
311 Goodness of fit (GOF) and C-index are shown for each cause of death.

312

Model extension and re-testing with the inclusion of diagnostic biopsy information

After assessing multiple categorisations of PPC, PPC was integrated into the model using a dichotomous variable around a cut-off of 50% (Tables E and F in S1 Appendix). PPC <50% or ≥50% were associated with adjusted hazard ratios for PCSM of 0.54 and 1.78 respectively. A hazard ratio of 1.0 is applied if PPC is unknown or to omit the PPC variable (Table G in S1 Appendix).

Accuracy of the final extended model, incorporating PPC, was re-assessed using the Singaporean cohort (n=2,546). Model discrimination was slightly improved compared to the baseline model with C-index 0.85 (95%CI: 0.82-0.88) and 0.76 (95%CI 0.73-0.79) for PCSM and overall mortality respectively (Table H in S1 Appendix). Calibration was also improved with the incorporation of the PPC variable (Fig K in S1 Appendix). GOF analysis showed no significant difference between observed and predicted PCa-related deaths (p=0.11) although the model still appeared to slightly underestimate PCSM. Calibration within subgroups (Table J in S1 Appendix) suggested the model underestimated PCSM in the context of very high-risk characteristics: grade group 5 (predicted: 30.6, observed: 36), t-stage 4 (predicted: 4.1, observed: 8) and PSA >50ng/ml (predicted: 21, observed: 25).

Next, we compared accuracy of our extended model to existing PCa models within this external cohort. The model continued to out-perform existing models in predicting overall mortality (p<0.001) (Table 6). For PCSM, improved C-indices were observed for PCSM compared to existing models, but again only reached significance compared to the EAU criteria. Finally, we limited the cohort to only men who received conservative management or radical treatment, to model contemporary practice where primary hormone therapy is less commonly used(20). Again, the model generally showed superior discrimination compared to other models (Table K in S1 Appendix).

	PCSM			Overall		
Model	C-index	95% CI	p	C-index	95% CI	p
PREDICT	0.838	0.804-0.872	-	0.756	0.728-0.784	-
EAU	0.763	0.732-0.794	0.001	0.637	0.606-0.667	<0.001
NCCN	0.804	0.767-0.841	0.182	0.649	0.616-0.682	<0.001
CAPRA	0.822	0.785-0.860	0.530	0.671	0.638-0.704	<0.001

Table 6 Discrimination of the extended model, compared to other existing models amongst the Singaporean cohort over 15 years maximum follow-up (n=2546).
EAU = European Association of Urology NCCN = National Comprehensive Cancer Network CAPRA = Cancer of the Prostate Risk Assessment (UCSF)

Proposed clinical utility of the model

To establish utility of the tool for clinicians and patients we have developed a web based interface for free access to the model. We expect that primary utility will be among men for whom conservative management and radical treatment might both be appropriate options. Example outputs from this web tool for 3 hypothetical vignettes are demonstrated in Fig 3. The age and comorbidity status at diagnosis are altered within each case to demonstrate the impact of competing risks on treatment benefit. With increasing age and comorbidity, reductions in PCSM achieved by radical treatment are attenuated by increased rates of NPCM as the risks of PCSM and NPCM compete against one another. For example a 72 year-old with comorbidity and the disease characteristics shown in Case B has an estimated 19.6% 15-year risk of prostate cancer death when conservatively managed. Although the estimated PCSM is reduced to 11.1% by treatment, the overall survival improves by only 3.8%, whereas for a younger man the majority of PCSM benefit translates into overall survival benefit (Fig 3).

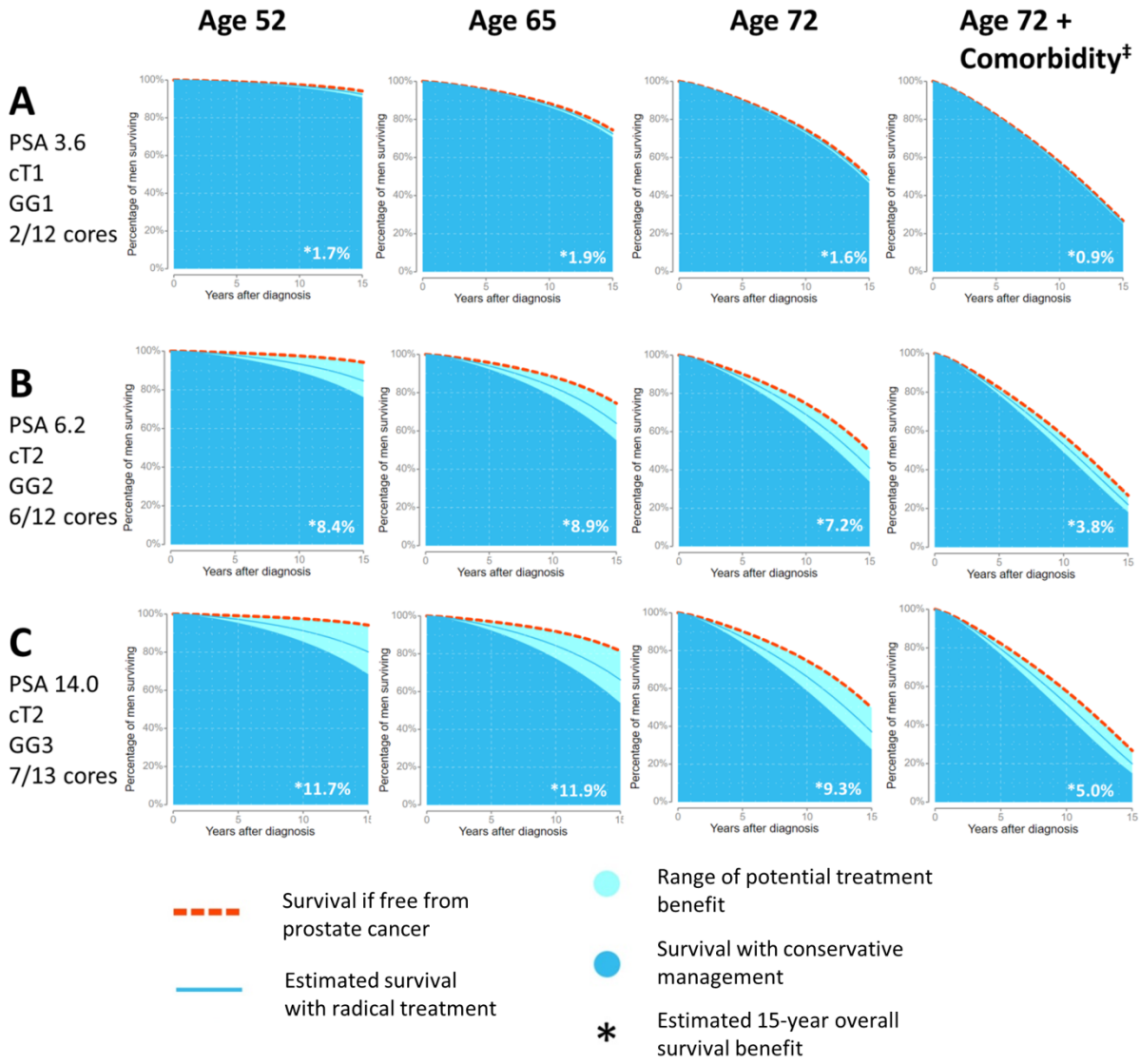


Figure 3 Example model outputs using 15-year overall survival curves for three hypothetical vignettes A, B and C. Only age and comorbidity status has been changed between each column to demonstrate the reduction in benefit from radical treatment when competing risk increases.
PSA = Prostate specific antigen cT = clinical tumour stage GG = histological grade group ‡ Comorbidity refers to a patient with Charlson score of 1 or more who has been admitted to hospital in the 2 years prior to prostate cancer diagnosis.

Discussion

In this study, to our knowledge, we present the first individualised multivariable prognostic model for non-metastatic PCa built and validated in an unscreened, pre-treatment cohort. We show that this model, hereafter referred to as PREDICT *Prostate*, is able to derive predictions for PCa and overall mortality with a high degree of concordance by using routinely available diagnostic clinico-pathological data, and appears to outperform existing models. The model incorporates the impact of radical therapy, which allows comparison to be made against the option of conservative management within the context of an individual's competing risks. Importantly, the model does not require any additional tests or information, but could be refined in the future if additional independent factors with proven prognostic value are established.

PCa incidence is rising with an ageing male population and increased testing. In the UK alone, the incidence is projected to rise by 69% by 2030[26]. Over 84% of UK men have non-metastatic disease at presentation with more than half of these classified as low or intermediate-risk using traditional risk criteria[2]. Level 1 evidence shows that many men with these disease characteristics will not benefit from immediate radical therapy, with the randomised ProtecT and PIVOT trials reporting no survival differences in men managed by intervention or conservative management after 10 years of follow up[3,4]. Additionally, radical treatment is associated with risks of significant adverse effects including incontinence, impotence, bowel dysfunction and long-term decisional regret[27,28]. Unsurprisingly, conservative management or active surveillance is therefore becoming increasingly popular in low-risk disease, and emerging evidence also suggests very favourable outcomes in intermediate-risk disease[29].

Identifying men appropriate for initial conservative management and conveying this information to an individual within their own context of competing mortality is currently an imprecise exercise, with a lack of objective data on potential outcomes. Instead, most current prognostication is directed by categorisation of men into risk stratified criteria and discussions with clinicians who may or may not

391 be PCa-specialists and are potentially conflicted by a bias to a treatment they offer [8-10,30].
392 PREDICT *Prostate* was conceived to address this critical gap in clinical need and better inform and
393 standardise the decision-making process. It is built around long-term actual survival data and has
394 been designed to address all AJCC criteria[14]. The model incorporates variables available for almost
395 any man diagnosed around the world and has wide potential applications in informing patient,
396 clinician and multi-disciplinary team decision-making to reduce both over and under-treatment[31].
397 Abundant literature shows that better decision aids contribute to more knowledgeable, informed
398 patients and that this improves clinician-patient communication[32,33]. Therefore, we anticipate our
399 model will be widely acceptable and highly impactful, although formal clinical impact assessments
400 will also be undertaken[34].

401 The parameters used within PREDICT *Prostate* for PCSM are well established independent variables
402 such as Grade group, PSA and T Stage[35-37]. Here, they have been combined in a novel way and by
403 utilising fractional polynomials to maintain as much predictive information as possible. PREDICT
404 *Prostate* is also distinctive in estimating the competing risks of PCSM and NPCM to accurately model
405 overall mortality. The model deliberately uses histological grade groups (1-5) as we standardise
406 practice towards this more-intuitive scale[19]. Biopsy information was integrated as an optional
407 variable in PREDICT *Prostate* as biopsy quantification is accepted as a surrogate for tumour volume.
408 However, no consensus on the best methodology for its assessment yet exists, with few studies
409 exploring its relationship with long-term survival[38]. Hence we used a pragmatic assessment of this
410 by using the simplest common denominator, the number of positive versus overall biopsy cores
411 taken (PPC). Our data showed an independent prognostic impact around the dichotomous cut-off of
412 <50% versus ≥50% PPC. This is the same cut-off reported in two American studies exploring survival,
413 where effect size is comparable. This cut-off has now also been integrated into the latest NCCN risk-
414 criteria[10,39,40]. PPC thus maintains simplicity and facilitates ease of interpretation (although the
415 model can function without biopsy information). During the study period local practice was to
416 perform 12-core systematic trans-rectal biopsy. However, contemporary practice in prostate biopsy

is evolving with the use of more image-targeting[41]. It is unknown how these changes will alter the prognostic value of biopsy involvement. In the meantime, we recommend adherence to the AUA guidelines which suggest any biopsies from a target are considered as a single core if taken as part of a 'target and systematic' biopsy approach[9].

A key question whilst developing PREDICT *Prostate* was whether to use data-derived coefficients for treatment effect or published trial data. Ultimately the data-derived coefficient for the combination of radical treatment types was used, with a hazard ratio of 0.50 (95%CI 0.38-0.67). This is in fact very similar to published randomised controlled trial data of treatment effect e.g. PIVOT (RP vs AS: HR 0.63 95%CI: 0.36-1.09) and ProtecT trials (RT vs active monitoring: HR 0.51 95%CI: 0.15-1.69. RP vs active monitoring: 0.63 95%CI: 0.21-1.93)[3,4]. In the web-based presentation of the model, uncertainty around treatment effect is demonstrated by displaying treatment benefit from 0-100% of PCSM around the estimated survival (Fig 3). Separate presentation of RT and RP outcomes was not explored as no adequate randomised data yet shows a survival difference between the two treatment approaches[4,42]. One caveat in the clinical utility of PREDICT *Prostate* is that primary androgen deprivation, used in a proportion of our study cohorts, is now seldom used as a first line therapy. Indeed, within this cohort the poor prognosis apparently associated with primary androgen deprivation is likely to reflect a selection bias towards men unfit for other treatment options, or with potentially occult metastatic disease. Our model however is primarily for use among men deciding between conservative management and radical treatment – where decision dilemmas are most acute. Indeed, as shown in Table C in S1 Appendix, calibration of the model was best amongst men with low to intermediate-risk features where this model would be most useful and appropriate in clinical decision-making. Using disease status information from the National Prostate Cancer Audit, this may represent up to 47% of all newly diagnosed prostate cancers[2].

Particular strengths of PREDICT *Prostate* include the derivation from a large cohort from a geographical area straddling 2 academic centres and 9 general hospitals. These data were collected

prospectively by an independent cancer registry with accurate death certificate notification, avoiding many potential biases associated with single-centre studies. The accuracy of UK PCa cause of death reporting is also known to be very reliable[43]. However, we do acknowledge limitations in the model. We do not have data on MRI-defined lesions or radiological stage. However, it is yet unknown if these data will improve prognostic ability with MRI primarily used to guide biopsies rather than offer prognostic information. Indeed, the additional value of MRI in detecting missed cancers is debatable given that men with a missed cancer using non-imaging approaches have extremely low rates of PCa death[44]. The model also does not currently integrate genomic tests or molecular markers. However, the most established tools such as Prolaris CCP and Oncotype DX GPS have predominantly been tested against shorter-term outcomes in very selected groups, particularly in the post-treatment setting[45,46]. When these expensive tools have been assessed against PCSM, concordance is very similar to our model. For example the Decipher genomic classifier alongside CAPRA showed an AUC of 0.78 (95%CI 0.68-0.87) for 10-year PCSM following prostatectomy[47]. We agree with others, that good data should be sought as to whether any such marker truly adds independent prognostic information beyond a gold-standard multivariable model[48]. As with MRI, if one or more marker does show independent prognostic value in the future it can be included in future refinements to PREDICT *Prostate*[49]. By using real world data, our treatment categories were based upon actual treatments received as opposed to assigned treatments as is often problematic in randomised trials[4]. However, our analysis cannot account for the impact of delayed conversions to treatment beyond 1 year, albeit the number of men switching from conservative management was very small (5.7%). A final potential limitation of the model is the lack of t-stage sub-classifications. However, it is accepted that t stage is often inaccurately assigned in localised disease[18].

In terms of statistical approach, we recognise that more complex flexible parametric survival modelling frameworks exist. For example, there are several penalized regression approaches such as LASSO, ridge-regression and random forests which could have been applied. However, we have used an established methodology, which in other tumour types could not be improved upon by more

complex approaches[50]. Our approach also has the advantages of allowing straight-forward external validation and the incorporation of additional parameters should sufficient evidence support their inclusion, as demonstrated by updates to the PREDICT breast cancer model[51]. We also appreciate that our external validation cohort was relatively small, and different from our model development dataset. Gaining access to well-annotated cohorts with long term follow-up outcomes is difficult, this dataset represented the best independent cohort available to us. Applying the model in this cohort of differing case-mix and ethnicity was considered a good test of the generalisability of the tool. The similar discriminatory performance herein, may suggest ethnicity is not a key determinant of prognosis. However, we recognise that follow-up duration in the Singaporean cohort is short, and the model remains untested among many other healthcare, geographic and ethnic contexts. Finally, our comparisons to the EAU, NCCN and CAPRA stratification criteria are pragmatic but potentially unfair. These models are intended to delineate patients into groups of risk, rather than offering predictions of 10- or 15-year risk. However, these are widely used clinical models such that these comparisons may be of interest to PCa specialists, particularly in the absence of equivalent models to compare against.

In conclusion, we have developed an individualised prognostication and decision-making tool for use at the point of PCa diagnosis. For the first time to our knowledge, this simultaneously presents individualised estimates of cancer-specific and overall survival outcomes and can model the impact of treatment on these outcomes. The accuracy of the model is promising across populations, and provides encouraging levels of discrimination in two validation cohorts. This model underpins a proposed new web-based tool and decision-aid to inform the decision-making process for patients and clinicians. Further external validation of the model should be established to explore accuracy and generalisability across other contexts – particularly testing validity amongst non-Caucasians and those detected through screening.

Acknowledgements

Data for this study is based on information collected and quality assured by the PHE National Cancer Registration and Analysis Service. Access to the data was facilitated by the PHE Office for Data Release.

We thank our colleagues from The Winton Centre for Risk and Evidence Communication, Cambridge, particularly David Spiegelhalter, Alex Freeman and Mike Pearson who have provided invaluable insight into this project and important web-development and design expertise.

References

1. Cancer Research UK, Prostate cancer statistics, Available from: http://www.cancerresearchuk.org/cancer-info/cancerstats/types/prostate/ .
2. National Prostate Cancer Audit - Annual Report 2017. Available from https://www.npca.org.uk/reports/npca-annual-report-2017/.
3. Wilt TJ, Brawer MK, Jones KM, Barry MJ, Aronson WJ, Fox S, et al. Radical prostatectomy versus observation for localized prostate cancer. N Engl J Med. 2012;367(3):203-13.
4. Hamdy FC, Donovan JL, Lane JA, Mason M, Metcalfe C, Holding P, et al. 10-Year Outcomes after Monitoring, Surgery, or Radiotherapy for Localized Prostate Cancer. N Engl J Med. 2016.
5. D'Amico AV, Whittington R, Malkowicz SB, Schultz D, Blank K, Broderick GA, et al. Biochemical outcome after radical prostatectomy, external beam radiation therapy, or interstitial radiation therapy for clinically localized prostate cancer. JAMA. 1998;280(11):969-74.
6. Jhaveri FM, Zippe CD, Klein EA, Kupelian PA. Biochemical failure does not predict overall survival after radical prostatectomy for localized prostate cancer: 10-year results. Urology. 1999;54(5):884-90.
7. Mottet N, Bellmunt J, Bolla M, Briers E, Cumberbatch MG, De Santis M, et al. EAU-ESTRO-SIOG Guidelines on Prostate Cancer. Part 1: Screening, Diagnosis, and Local Treatment with Curative Intent. European Urology. 2017;71(4):618-29.
8. NICE. National Institute for Health and Care Excellence NICE Guidelines [CG175] Prostate cancer: diagnosis and treatment. January 2014.
9. Sanda MG, Cadeddu JA, Kirkby E, Chen RC, Crispino T, Fontanarosa J, et al. Clinically Localized Prostate Cancer: AUA/ASTRO/SUO Guideline. Part I: Risk Stratification, Shared Decision Making, and Care Options. J Urol. 2018;199(3):683-690.
10. National Comprehensive Cancer Network (NCCN) Clinical Practice Guidelines in Oncology: Prostate Cancer. Version 2. 2018 Available from: https://www.nccn.org/professionals/physician_gls/pdf/prostate.pdf.
11. Kerkmeijer LGW, Monninkhof EM, van Oort IM, van der Poel HG, de Meerleer G, van Vulpen M. PREDICT: model for prediction of survival in localized prostate cancer. World Journal of Urology. 2016;34(6):789-95.
12. Kutikov A, Cooperberg MR, Paciorek AT, Uzzo RG, Carroll PR, Boorjian SA. Evaluating prostate cancer mortality and competing risks of death in patients with localized prostate cancer using a comprehensive nomogram. Prostate Cancer and Prostatic Diseases. 2012;15(4):374-9.
13. American Joint Committee on Cancer. AJCC Cancer Staging Manual, 8th Edition. 2016. p. 1-7.
14. Kattan MW, Hess KR, Amin MB, Lu Y, Moons KG, Gershenwald JE, et al. American Joint Committee on Cancer acceptance criteria for inclusion of risk models for individualized prognosis in the practice of precision medicine. CA Cancer J Clin. 2016;66(5):370-4.

15. Collins GS, Reitsma JB, Altman DG, Moons KG. Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): the TRIPOD statement. BMJ. 2015;350:g7594.
16. Greenberg DC, Wright KA, Lophathanon A, Muir KR, Gnanapragasam VJ. Changing presentation of prostate cancer in a UK population--10 year trends in prostate cancer risk profiles in the East of England. Br J Cancer. 2013;109(8):2115-20.
17. Buzzoni C, Auvinen A, Roobol MJ, Carlsson S, Moss SM, Puliti D, et al. Metastatic Prostate Cancer Incidence and Prostate-specific Antigen Testing: New Insights from the European Randomized Study of Screening for Prostate Cancer. Eur Urol. 2015;68(5):885-90.
18. Reese AC, Sadetsky N, Carroll PR, Cooperberg MR. Inaccuracies in assignment of clinical stage for localized prostate cancer. Cancer. 2011;117(2):283-9.
19. Epstein JI, Zelefsky MJ, Sjoberg DD, Nelson JB, Egevad L, Magi-Galluzzi C, et al. A Contemporary Prostate Cancer Grading System: A Validated Alternative to the Gleason Score. Eur Urol. 2016;69(3):428-35.
20. Greenberg DC, Lophatananon A, Wright KA, Muir KR, Gnanapragasam VJ. Trends and outcome from radical therapy for primary non-metastatic prostate cancer in a UK population. PLoS One. 2015;10(3):e0119494.
21. Candido Dos Reis FJ, Wishart GC, Dicks EM, Greenberg D, Rashbass J, Schmidt MK, et al. An updated PREDICT breast cancer prognostication and treatment benefit prediction model with independent validation. Breast Cancer Res. 2017;19(1):58.
22. May S, Hosmer DW. A simplified method of calculating an overall goodness-of-fit test for the Cox proportional hazards model. Lifetime Data Anal. 1998;4(2):109-20.
23. Harrell F. Package 'Hmisc'. In: Dupont C, editor. CRAN2018. p. 235-6.
24. Cooperberg MR, Pasta DJ, Elkin EP, Litwin MS, Latini DM, Du Chane J, et al. The University of California, San Francisco Cancer of the Prostate Risk Assessment score: a straightforward and reliable preoperative predictor of disease recurrence after radical prostatectomy. J Urol. 2005;173(6):1938-42.
25. Gnanapragasam VJ, Bratt O, Muir K, Lee LS, Huang HH, Stattin P, et al. The Cambridge Prognostic Groups for improved prediction of disease mortality at diagnosis in primary non-metastatic prostate cancer: a validation study. BMC Med. 2018;16(1):31.
26. Mistry M, Parkin DM, Ahmad AS, Sasieni P. Cancer incidence in the United Kingdom: projections to the year 2030. Br J Cancer. 2011;105(11):1795-803.
27. Donovan JL, Hamdy FC, Lane JA, Mason M, Metcalfe C, Walsh E, et al. Patient-Reported Outcomes after Monitoring, Surgery, or Radiotherapy for Prostate Cancer. N Engl J Med. 2016;375(15):1425-37.
28. Hoffman RM, Lo M, Clark JA, Albertsen PC, Barry MJ, Goodman M, et al. Treatment Decision Regret Among Long-Term Survivors of Localized Prostate Cancer: Results From the Prostate Cancer Outcomes Study. Journal of Clinical Oncology. 2017;35(20):2306-2314.
29. Klotz L, Zhang LY, Lam A, Nam R, Mamedov A, Loblaw A. Clinical Results of Long-Term Follow-Up of a Large, Active Surveillance Cohort With Localized Prostate Cancer. Journal of Clinical Oncology. 2010;28(1):126-31.
30. Kim SP, Gross CP, Nguyen PL, Nguyen PY, Smaldone MC, Thompson RH, et al. Specialty bias in treatment recommendations and quality of life among radiation oncologists and urologists for localized prostate cancer. Prostate Cancer Prostatic Dis. 2014;17(2):163-9.
31. JLA. James Lind Alliance Priority setting partnerships. Prostate Cancer Top 10:1. How can overtreatment for prostate cancer be prevented by identifying and excluding the treatment of harmless tumours? . 2016. Available from: http://www.jla.nihr.ac.uk/priority-setting-partnerships/prostate-cancer/top-10-priorities/
32. O'Connor AM, Rostom A, Fiset V, Tetroe J, Entwistle V, Llewellyn-Thomas H, et al. Decision aids for patients facing health treatment or screening decisions: systematic review. BMJ. 1999;319(7212):731-4.

586 33. Lin GA, Aaronson DS, Knight SJ, Carroll PR, Dudley RA. Patient decision aids for prostate
587 cancer treatment: a systematic review of the literature. CA Cancer J Clin. 2009;59(6):379-90.
588 34. Moons KG, Altman DG, Vergouwe Y, Royston P. Prognosis and prognostic research:
589 application and impact of prognostic models in clinical practice. BMJ. 2009;338:b606.
590 35. Grogan J, Gupta R, Mahon KL, Stricker PD, Haynes AM, Delprado W, et al. Predictive value of
591 the 2014 International Society of Urological Pathology grading system for prostate cancer in patients
592 undergoing radical prostatectomy with long-term follow-up. Bju Int. 2017;120(5):651-8.
593 36. Bostwick DG, Foster CS. Predictive factors in prostate cancer: current concepts from the
594 1999 College of American Pathologists Conference on Solid Tumor Prognostic Factors and the 1999
595 World Health Organization Second International Consultation on Prostate Cancer. Semin Urol Oncol.
596 1999;17(4):222-72.
597 37. Partin AW, Steinberg GD, Pitcock RV, Wu L, Piantadosi S, Coffey DS, et al. Use of nuclear
598 morphometry, Gleason histologic scoring, clinical stage, and age to predict disease-free survival
599 among patients with prostate cancer. Cancer. 1992;70(1):161-8.
600 38. Vollmer RT. Tumor length in prostate cancer. American Journal of Clinical Pathology.
601 2008;130(1):77-82.
602 39. Huang JY, Vicini FA, Williams SG, Ye H, McGrath S, Ghilezan M, et al. Percentage of Positive
603 Biopsy Cores: A Better Risk Stratification Model for Prostate Cancer? International Journal of
604 Radiation Oncology Biology Physics. 2012;83(4):1141-8.
605 40. D'Amico AC, Renshaw AA, Cote K, Hurwitz M, Beard C, Loffredo M, et al. Impact of the
606 percentage of positive prostate cores on prostate cancer-specific mortality for patients with low or
607 favorable intermediate-risk disease. Journal of Clinical Oncology. 2004;22(18):3726-32.
608 41. Kasivisvanathan V, Rannikko AS, Borghi M, Panebianco V, Mynderse LA, Vaarala MH, et al.
609 MRI-Targeted or Standard Biopsy for Prostate-Cancer Diagnosis. N Engl J Med. 2018.
610 42. Roach M, Lizarraga TLC, Lazar AA. Radical Prostatectomy Versus Radiation and Androgen
611 Deprivation Therapy for Clinically Localized Prostate Cancer: How Good Is the Evidence?
612 International Journal of Radiation Oncology Biology Physics. 2015;93(5):1064-70.
613 43. Turner EL, Metcalfe C, Donovan JL, Noble S, Sterne JA, Lane JA, et al. Contemporary accuracy
614 of death certificates for coding prostate cancer as a cause of death: Is reliance on death certification
615 good enough? A comparison with blinded review by an independent cause of death evaluation
616 committee. Br J Cancer. 2016;115(1):90-4.
617 44. Klemann N, Roder MA, Helgstrand JT, Brasso K, Toft BG, Vainer B, et al. Risk of prostate
618 cancer diagnosis and mortality in men with a benign initial transrectal ultrasound-guided biopsy set:
619 a population-based study. Lancet Oncology. 2017;18(2):221-9.
620 45. Ontario HQ. Prolaris Cell Cycle Progression Test for Localized Prostate Cancer: A Health
621 Technology Assessment. Ont Health Technol Assess Ser. 2017;17(6):1-75.
622 46. Cucchiara V, Cooperberg MR, Dall'Era M, Lin DW, Montorsi F, Schalken JA, et al. Genomic
623 Markers in Prostate Cancer Decision Making. Eur Urol. 2018;73(4):572-82.
624 47. Cooperberg MR, Davicioni E, Crisan A, Jenkins RB, Ghadessi M, Karnes RJ. Combined Value of
625 Validated Clinical and Genomic Risk Stratification Tools for Predicting Prostate Cancer Mortality in a
626 High-risk Prostatectomy Cohort. European Urology. 2015;67(2):326-33.
627 48. Herlemann A, Washington SL, Eapen RS, Cooperberg MR. Whom to Treat: Postdiagnostic
628 Risk Assessment with Gleason Score, Risk Models, and Genomic Classifier. Urol Clin North Am.
629 2017;44(4):547-55.
630 49. Wishart GC, Bajdik CD, Dicks E, Provenzano E, Schmidt MK, Sherman M, et al. PREDICT Plus:
631 development and validation of a prognostic model for early breast cancer that includes HER2. Br J
632 Cancer. 2012;107(5):800-7.
633 50. Karapanagiotis S, Pharoah PDP, Jackson CH, Newcombe PJ. Development and External
634 Validation of Prediction Models for 10-Year Survival of Invasive Breast Cancer. Comparison with
635 PREDICT and CancerMath. Clin Cancer Res. 2018;24(9):2110-5.

636 51. Wishart GC, Rakha E, Green A, Ellis I, Ali HR, Provenzano E, et al. Inclusion of KI67
637 significantly improves performance of the PREDICT prognostication and prediction model for early
638 breast cancer. BMC Cancer. 2014;14:908.

639

640 **Legends**

641 **S1 Checklist**. Transparent reporting of a multivariable prediction model for individual prognosis or
642 diagnosis (TRIPOD) Checklist.

643 **S1 Proposal**. Prospective research proposal for doctoral project on the development and
644 implementation of a risk prediction model for non-metastatic prostate cancer.

645 **S1 Appendix**. Technical appendix to the manuscript, including additional text, tables and figures.