Supplementary materials

Methods

FFPE sample preparation and RNA extraction from CNB samples

The detailed steps of the sample preparation process (tissue sectioning and marking/quantification) using human PCa FFPE samples were published previously[1]. For all patients, marked cancer-cell containing areas were scraped from unstained FFPE sections into DNase/RNase-free microcentrifuge tubes for RNA extraction using a disposables scalpel. A new scalpel was used for each patient to prevent cross-contamination. If FFPE sections from more than one biopsy were used for a patient, the tissue samples from the different biopsies were pooled. For most patients, the cancer area was at least 15 mm^2 with ≥ 50% epithelial cancer cells. From the scraped tissue samples, total RNA was extracted using the commercially available High Pure FFPE RNA Micro Kit (Roche, catalog number: 4823125001) according to the manufacturer's instructions with the following modifications: 1) For specimens collected less than one year prior to RNA extraction, add 350 µl Lysis Buffer and 350 µl Binding Buffer, in the homogenisation step to reduce viscosity; 2) Centrifuge the High Pure Micro filter for 4 minutes instead of 2 minutes at maximum speed before the RNA elution step. Extracted RNA was immediately subjected to the Prostatype® RT-qPCR analysis in the same day without storing.

One-step RT-qPCR reaction and Gene expression

Total RNA extracted from the FFPE samples were used for gene expression analysis using a four-plex one-step RT-qPCR. The expression levels of the three biomarker genes F3, IGFBP3 and VGLL3 as well as the housekeeping gene GAPDH were measured using the commercially available Prostatype® RT-qPCR kit (Prostatype Genomics AB, Solna, Sweden). All measurements were conducted using a Roche LightCycler 480 instrument II (Roche Molecular Systems, Inc.), a qPCR platform on which a colour compensation method was run prior to performing the qPCR analysis. The sequence information of the respective probes and primers has been reported previously[2]. The Prostatype® RT-qPCR kit contains a positive and negative control, which were assayed together with each batch of PCa tissue samples. A batch of RT-qPCR experiments was

considered valid only if positive and negative controls were valid. Samples with GAPDH C(t) values of >28.0 were excluded according to the Prostatype® RT-qPCR kits IFU 0018, revision-7[3]. The expression levels of the genes IGFBP3, F3 and VGLL3 genes in a sample were normalized to that of GAPDH and presented as delta C(t) (ΔC[t]) values. The ΔC(t) value is inversely correlated to the gene expression level. All samples were run in triplicates and the median ΔC(t) value of the triplicates for each gene was used for the following calculations. If a triplicate was invalid, then the mean of the two remaining ΔC(t) values was used.

All steps including FFPE tissue scraping, RNA extraction and Prostatype® RT-qPCR analysis were performed at the Prostatype Genomics AB Quality Control laboratory in Solna, Sweden.

Handling of clinical parameters

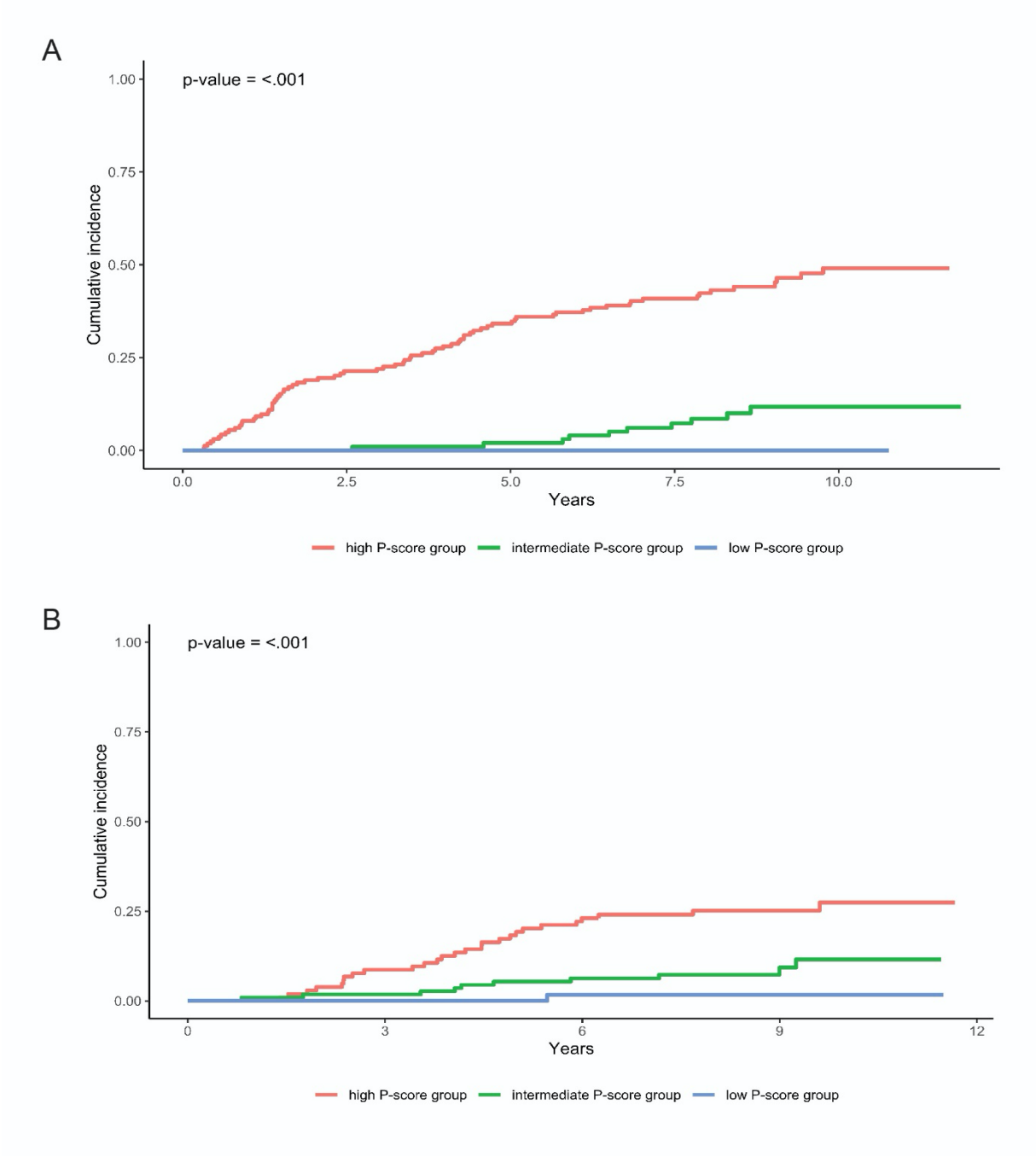
In cases where no substage of the parameter “tumor stage” was available, it was assumed that they were of the lowest substage. For example, if "T2” was recorded, it was assumed that this was “T2a”; if “T3” was recorded, it was assumed that this was “T3a”.

References

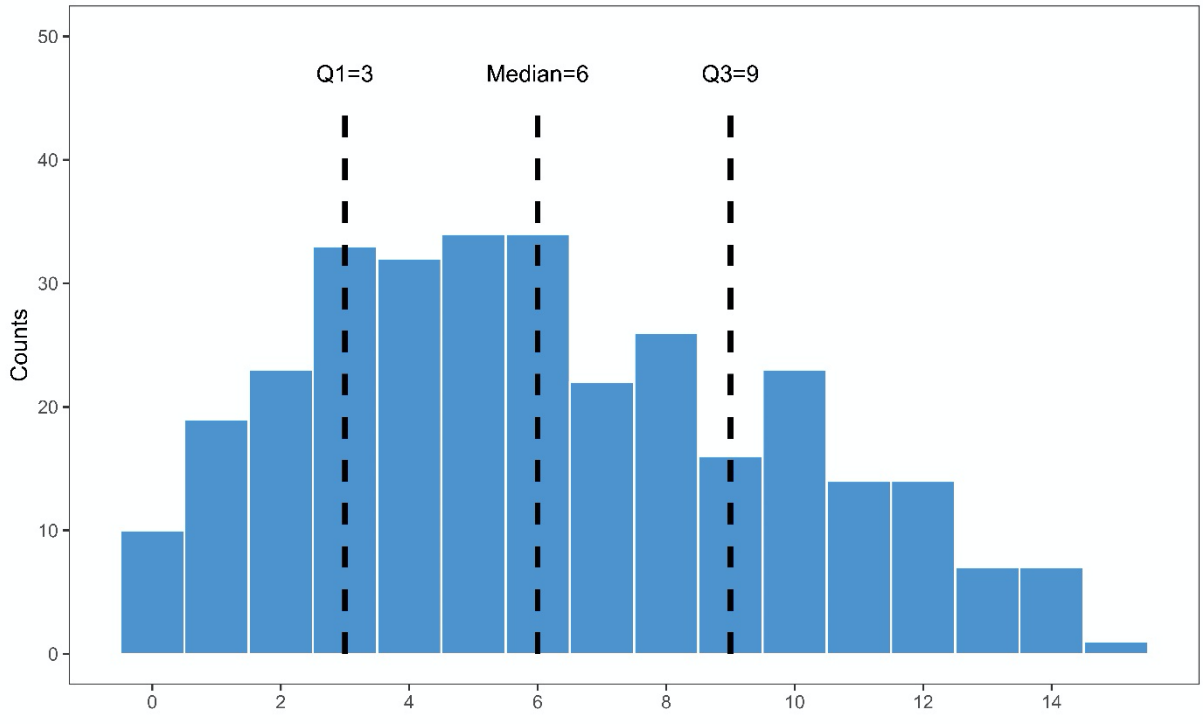
1. Peng Z, Andersson K, Lindholm J, et al. Operator dependent choice of prostate cancer biopsy has limited impact on a gene signature analysis for the highly expressed genes IGFBP3 and F3 in prostate cancer epithelial cells. *PLoS One*. 2014;9(10):e109610.
2. Peng Z, Skoog L, Hellborg H, et al. An expression signature at diagnosis to estimate prostate cancer patients' overall survival. *Prostate Cancer Prostatic Dis*. Mar 2014;17(1):81-90.
3. Prostatype Genomics AB. Prostatype RT qPCR Kit - Instructions for Use. Revision 7. Solna, Sweden. 2021.

Supplemental figures

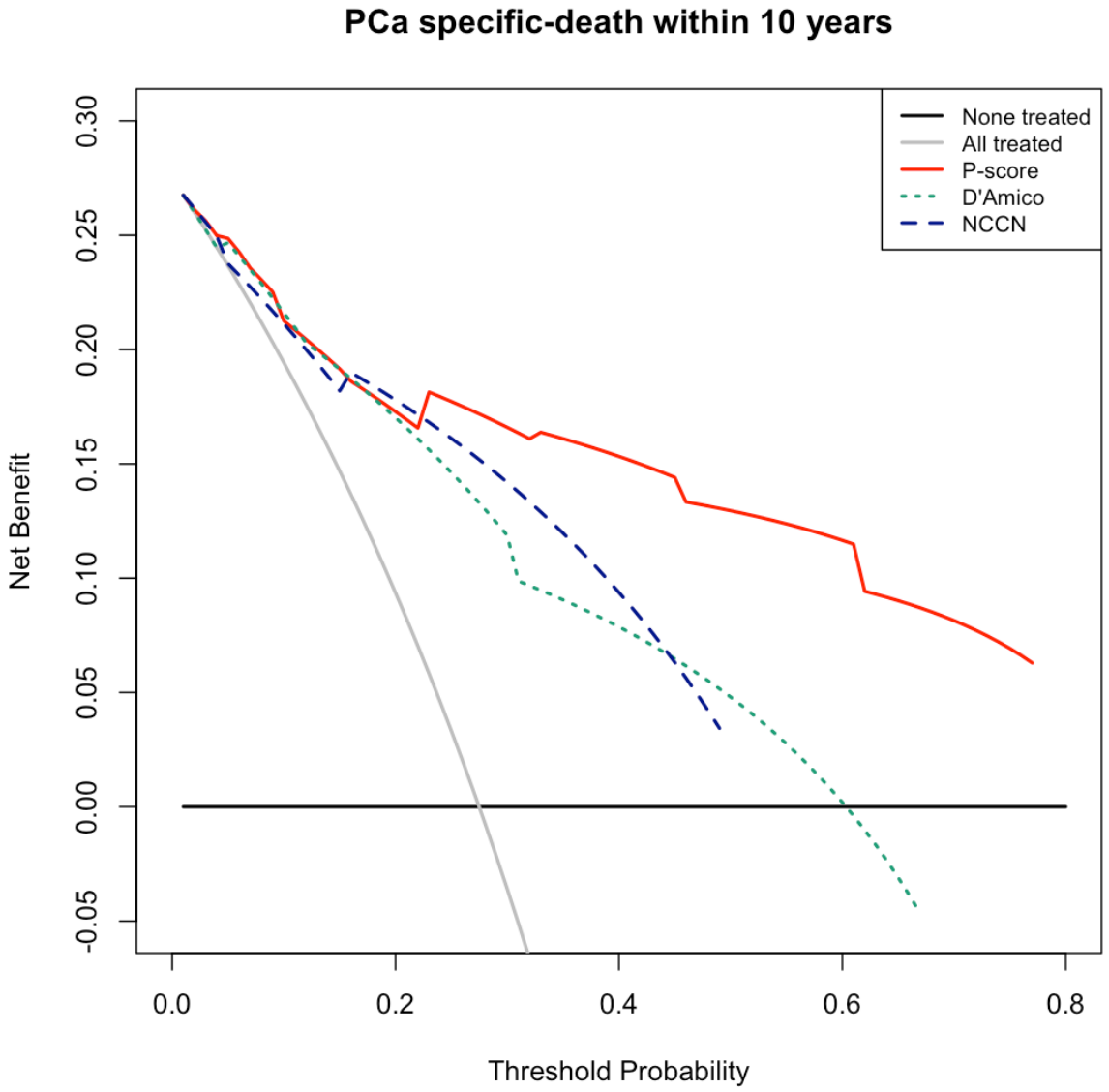
Supplemental Figure S1



Supplemental Figure S2



Supplemental Figure S3



Supplemental figure captions

Supplemental Figure S1:

(A) CIF analysis of the P-score to predict prostate cancer (PCa)-specific mortality versus time in the estimation dataset (n=315).

(B) CIF analysis of the P-score to predict PCa-specific mortality versus time in the validation dataset (n=276).

Supplemental Figure S2: Distribution of P-score in the estimation dataset (n=315).

Supplemental Figure S3: Comparison of P-score, D'Amico, and NCCN score by decision curve analysis for prostate cancer (PCa)-specific mortality within 10 years in 590 patients. Threshold probabilities: the relative harms of false positive and false negatives.

Supplemental tables

Supplemental Table S1

Dataset	Patients, n (%)	PCa death, n (%)	HR (95% CI)	P-value
All patients	591 (100)	123 (100)	1.48 (1.40 – 1.57)	<0.001
Hormone treatment	202 (34)	90 (73)	1.35 (1.24 – 1.48)	<0.001
Radical treatment	196 (33)	19 (15)	1.47 (1.24 – 1.75)	<0.001
Active surveillance or Watchful waiting	193 (33)	14 (11)	1.60 (1.27 – 2.01)	<0.001

Supplemental Table S2

	Patients, n	PCa death, n	Scores	C-index (95% CI)	AUC (95% CI)
All	590	123	P-score	0.82 (0.78 – 0.85)	0.83 (0.78 – 0.88)
			D'Amico	0.74 (0.71 – 0.78)	0.76 (0.71 – 0.80)
			NCCN	0.75 (0.72 – 0.79)	0.76 (0.71 – 0.81)
Estimation dataset	314	85	P-score	0.83 (0.79– 0.87)	0.85 (0.80 – 0.90)
			D'Amico	0.75 (0.71 – 0.79)	0.78 (0.72 – 0.83)
			NCCN	0.76 (0.72 – 0.80)	0.78 (0.72 – 0.84)

Supplemental Table S3

		Patients, n	**PCa death, n**	**Scores**	**C-index (95% CI)**	**AUC (95% CI)**
GS≤7	Estimation dataset	235	37	P-score	0. 80 (0.73 – 0.87)	0.82 (0.73 – 0.90)
				D'Amico	0.74 (0.67 – 0.81)	0.76 (0.68 – 0.84)
				NCCN	0.75 (0.67 – 0.82)	0.75 (0.67 – 0.84)
	Validation dataset	233	23	P-score	0.72 (0.60 – 0.83)	0.71 (0.57 – 0.80)
				D'Amico	0.66 (0.57 – 0.76)	0.65 (0.54 – 0.77)
				NCCN	0.67 (0.58 – 0.76)	0.65 (0.54 – 0.76)
GS>7	Estimation dataset	79	48	P-score	0.70 (0.63 – 0.78)	0.758 (0.65 – 0.87)
				D'Amico	0.59 (0.51 – 0.67)	0.60 (0.48 – 0.71)
				NCCN	0.58 (0.52 – 0.64)	0.62 (0.53 – 0.70)
	Validation dataset	43	15	P-score	0.68 (0.55 – 0.81)	0.77 (0.69 – 0.93)
				D'Amico	0.64 (0.50 – 0.76)	0.67 (0.51 – 0.83)
				NCCN	0.53 (0.44 – 0.58)	0.56 (0.45 – 0.67)

Supplemental Table S4

Variables	**C-index**
T-stage	0.69
T-stage + IGFBP3 + F3 + VGLL3	0.74
T-stage + PSA	0.75
T-stage + PSA + IGFBP3 + F3 + VGLL3	0.76
T-stage + PSA + Gleason	0.78
IGFBP3 + F3 + VGLL3	0.71
D'Amico	0.74
D'Amico + IGFBP3 + F3 + VGLL3	0.78
NCCN	0.75
NCCN + IGFBP3 + F3 + VGLL3	0.80
P-score	0.82

Supplemental table captions

Supplemental Table S1. Hazard ratio (HR) of the P-score by treatment group (full dataset, N = 591). 95% confidence interval (CI) and p-value are presented.

Supplemental Table S2. Prediction performance comparison of P-score, D'Amico, and NCCN in the full dataset (N=590) and estimation dataset (N=314). Both concordance index (C-index) and area under the curve (AUC) evaluated at 10 years follow-up time are presented. 95% confidence intervals (CI) are shown.

Supplemental Table S3: Prediction performance comparison of P-score, D'Amico, and NCCN in subgroups with Gleason score (GS) ≤7 and GS >7 in the estimation dataset (N=314) and validation dataset (N=276). Both concordance index (C-index) and area under the curve (AUC) evaluated at 10 years follow-up time are presented. 95% confidence intervals (CI) are shown.

Supplemental Table S4: Prediction performance comparison of combinations of different prognostic factors in the full dataset. Concordance index (C-index) calculated by competing risk model.