Transcript analysis of commercial prostate cancer risk stratification panels in hard-to-predict grade group 2-4 prostate cancers

Timo-Pekka Lehto

Helsinki 23.11.2021
Tutkielma
timo-pekka.lehto@helsinki.fi
Ohjaajat: dos. Tuomas Mirtti, prof. Antti Rannikko
Helsingin yliopisto
Lääketieteellinen tiedekunta

Tiivistelmä

Tiivistelmä:

Tausta: Eturauhassyövän erilaistumisryhmien (grade group, GG) 2-4 prognoosin arvioinnissa tarvitaan edistysaskeleita ylihoidon minimoimiseksi. Tavoitteemme oli selvittää lähetti-RNA-tasolla kaupallisten paneelien, Decipher, Oncotype DX, Prolaris, ja mutaatiopaneeli MSK-IMPACTin suorituskykyä ennustaa etäpesäkevapaata ja eturauhassyöpäspesifistä selviytymistä GG 2-4 eturauhassyövässä radikaalin eturauhasen poiston jälkeen.

Materiaalit ja metodologia: Retrospektiivinen kohortti koostui radikaalilla eturauhasen poistolla hoidetuista GG 2-4 potilaista (10,4 vuoden mediaaniseuranta-aika). Analysoimme 76 tapausta, joilla oli leikkauksen jälkeinen etäpesäke tai eturauhassyöpäspesifinen kuolema, sekä 84 verrokkia, joilla oli samanveroinen etenemisriski lähtötilanteessa, mutta syöpä ei edennyt. Indeksileesioiden lähetti-RNAt analysoitiin käyttäen NanoString-alustaa. Koulutimme satunnaismetsämalleja (random forest) ennustamaan kliinisiä päätetapahtumia käyttäen paneelien geenikokoelmia. Mittasimme käyrän alla olevaa pinta-alaa (area under the curve, AUC), sensitiivisyyttä, spesifisyyttä, Youden indeksiä sekä NND-lukua (number needed to diagnose). Selviytymistodennäköisyyttä arvioitiin käyttämällä Kaplan-Meier estimaattoria.

Tulokset: Geenikokoelmien suorituskyky oli parempi etäpesäkevapaan ja eturauhassyöpäspesifisen selviytymisen ennustamisessa verrattuna kliinisiin muuttujiin. Paneelien välillä oli kuitenkin merkitseviä eroja. Oncotype DX ennusti metastaaseja heikommin (AUC=0,65) kuin muut paneelit (AUC=0,73-0,74). Decipherin, MSK-IMPACTin ja Prolariksen NND-luvut eivät eronneet merkittävästi (2,83-3,12), kun taas Oncotype DXn NND-luku oli korkein (4,79). Prolaris ennusti eturauhassyöpäspesifistä kuolemaa heikommin (AUC=0,66) kuin MSK-IMPACT tai Decipher (AUC=0,72). Oncotype DXn suorituskyky ei eronnut muista paneeleista (AUC=0,69, p>0,05), mutta sen NND-luku (2,79) oli matalampi kuin muilla paneeleilla (4,22-5,66).

Pohdinta: Kaupallisten riskiluokituspaneelien geenien transkriptianalyysi ennustaa GG 2-4 potilaiden selviytymistä radikaalin eturauhasen poiston jälkeen ja saattaa olla avuksi kliinisessä päätöksenteossa. Paneelien välillä oli merkitseviä eroja, mutta paremmin ennustaville geenikokoelmille on tarvetta. Prospektiivinen tutkimus biopsia-aineistossa on tarpeen.

ORIGINAL ARTICLE

Transcript analysis of commercial prostate cancer risk stratification panels in hard-to-predict grade group 2–4 prostate cancers

Timo-Pekka K. Lehto BM[1] | Carolin Stürenberg MD[1] | Adrian Malén BM[1] | Andrew M. Erickson PhD[2] | Hannu Koistinen DSc[3] | Ian G. Mills PhD[2,4] | Antti Rannikko MD, PhD[5,6] | Tuomas Mirtti MD, PhD[1,6]

[1]Department of Pathology, University of Helsinki and Helsinki University Hospital, Helsinki, Finland

[2]Nuffield Department of Surgical Sciences, University of Oxford, Oxfordshire, UK

[3]Department of Clinical Chemistry and Haematology, University of Helsinki and Helsinki University Hospital, Helsinki, Finland

[4]School of Medicine, Dentistry and Biomedical Sciences, Patrick G Johnston Center for Cancer Research, Queen's University of Belfast, UK

[5]Department of Urology, University of Helsinki and Helsinki University Hospital, Helsinki, Finland

[6]Research Program in Systems Oncology, Faculty of Medicine, University of Helsinki and Helsinki University Hospital, Helsinki, Finland

Correspondence
Tuomas Mirtti, MD, PhD, Department of Pathology, University of Helsinki and Helsinki University Hospital, Haartmaninkatu 3, 00290 Helsinki, Finland.
Email: tuomas.mirtti@helsinki.fi

Funding information
Cancer Foundation Finland; Academy of Finland, Grant/Award Number: 304667; Hospital Disctrict of Helsinki and Uusimaa, Grant/Award Numbers: TYH2018214, TYH2019235; Sigrid Jusélius Foundation

Abstract

Background: Improved prognostication is needed to minimize overtreatment in grade group (GG) 2–4 prostate cancer. Our aim was to determine, at messenger RNA (mRNA) level, the performance of the genes in the commercial panels Decipher, Oncotype DX, Prolaris, and mutational panel MSK-IMPACT to predict metastasis-free and prostate cancer-specific death (PCSD) in patients with GG 2–4 prostate cancer at radical prostatectomy.

Methods: The retrospective cohort consisted of GG 2–4 patients treated with radical prostatectomy (median follow-up 10.4 years). Seventy-six cases with postoperative metastasis or PCSD and 84 controls with similar clinical baseline risk, but without progression, were analyzed. Index lesion mRNA transcripts were analyzed using NanoString technology. Random forest models were trained using panel gene sets to predict clinical endpoints and area under the curve (AUC), sensitivity, specificity, Youden index, and number needed to diagnose (NND) was measured. Survival probability was assessed with Kaplan–Meier estimator.

Results: All gene sets outperformed clinical parameters and predicted metastasis-free and prostate cancer-specific survival. However, there were significant differences between the panels. In metastasis prediction, the genes in Oncotype DX had inferior performance (area under the curve [AUC] = 0.65) compared to other panels (AUC = 0.73–0.74). Decipher, MSK-IMPACT and Prolaris showed similar NND (2.83–3.12) with Oncotype DX having highest NND (4.79). In PCSD prediction, the Prolaris gene set performed worse (AUC = 0.66) than MSK-IMPACT or Decipher (AUC = 0.72). Oncotype DX performed similarly to other panels (AUC = 0.69, $p > .05$). Oncotype DX demonstrated lowest NND (2.79) compared to other panels (4.22–5.66).

Conclusion: Transcript analysis of genes included in commercial panels is feasible in survival prediction of GG 2-4 patients after radical prostatectomy and may aid in clinical decision making. There were significant differences between the panels, and

overall stronger predictive gene sets are needed. Prospective investigation is warranted in biopsy materials.

KEYWORDS

biomarker, decipher, gene panel, oncotype DX, prolaris

1 | INTRODUCTION

The main parameters in treatment decision-making and risk assessment of prostate cancer are blood concentration of prostate-specific antigen (PSA), cancer stage and grade. Evidence suggests that patients with Gleason score (GS) 3 + 4 and 4 + 3 prostate cancer have different prognosis,[1] which has influenced the development of the grade group (GG) system.[2,3] Patients with GG 1 prostate cancer are classified as harboring low-risk, GG 2–3 as intermediate-risk and GG 4–5 as high-risk disease. Despite the new GG system, there is great demand for novel prognostic factors, especially in intermediate-risk prostate cancer lesions including Gleason pattern 4.

No widely accepted biomarkers of clinically significant prostate cancer for routine clinical practice exist. PSA has been found to be unable to stratify aggressive and indolent prostate cancers.[4] However, several commercial gene expression panels have been suggested to fill this unmet need. While these panels have mainly been proposed for selection of very low- to low-risk patients in active surveillance programs,[5] they have also been found to contribute to cancer recurrence risk stratification after radical prostatectomy (RP).

Prolaris (Myriad Genetics) is a 31-gene panel, which, in biopsy material, can predict biochemical recurrence and the risk of developing metastatic disease after RP.[6] It was originally developed to help stratification of low-risk patients to active surveillance or active treatment. Prolaris cell cycle progression score can also predict prostate cancer-specific death (PCSD) in conservatively managed patients.[7]

Decipher (Decipher Biosciences) is another commercially available gene panel, featuring 22 genes. Decipher genomic classifier was developed and validated to predict the probability of metastatic progression after RP[8] and has been shown to predict PCSD within 10 years after RP.[9] Biopsy Decipher predicts the risk of metastasis 10 years after RP.[10] It has also been shown to predict absence of adverse pathology in very low- to low-risk patients according to the National Comprehensive Cancer Network (NCCN) guidelines.[11]

Oncotype DX (Genomic Health) is a biopsy tissue-based assay featuring 12 cancer-related genes, which predicts the risk of adverse pathology at RP in patients with low- to intermediate-risk prostate cancer at diagnosis.[12] Recently, the ability of Oncotype DX to predict adverse pathology, after adjusting for GG and PSA density, has been questioned.[13] The Oncotype DX genomic prostate score has been shown to predict biochemical recurrence, metastases[14] and PCSD after surgical treatment.[15] Memorial Sloan Kettering Cancer Center (MSK-IMPACT) is a 341-gene next generation sequencing mutational pan-cancer panel. It genetically profiles tumors,[16] without providing a risk-score.

Intra-patient reproducibility of Prolaris, Decipher, and Oncotype DX was recently examined in a small patient cohort with inconsistent results.[17] The commercial panels have been found to be independent predictors of adverse outcomes.[18] However, they have been investigated in mostly low- and high-risk prostate cancer, based on histology, in which GG can stratify patients with reasonably high accuracy. Furthermore, only relatively small amount of intermediate-risk prostate cancer patients have been included in the studies, although this group represents the major clinical need. Notably, Decipher has been studied in a cohort with a major representation of intermediate-risk prostate cancer according to GG, where low-risk and intermediate-risk groups, as predicted by Decipher score, showed similar metastasis-free survival.[19]

By analyzing the transcript levels of the genes included in the commercially available risk stratification panels and MSK-IMPACT, we aimed to evaluate the performance of these panels to stratify patients within a challenging retrospective cohort of localized GG 2–4 prostate cancers treated with RP. The goal was to examine if these gene sets could identify patients harboring prostate cancer with propensity to progress, after RP, to metastatic or lethal disease. Since the commercial panels have been investigated in patient materials mainly including low- or high-risk prostate cancer, as assessed with conventional parameters, it is not known whether they provide additional predictive information in mostly intermediate-risk prostate cancer.

2 | MATERIALS AND METHODS

2.1 | Patients, study design, and ethical considerations

The retrospective patient cohort consisted of 180 men treated with RP at HUS Helsinki University Hospital (HUS) during years 1992–2015. Surgical specimens harboring Gleason pattern 4, that is, GS 3 + 4, 4 + 3 or 4 + 4 and American Joint Committee on Cancer eighth edition histopathological tumor Stage 2–3 foci were curated from a national RP registry, based on availability of comprehensive clinical and follow-up data, histological slides and formalin-fixed paraffin-embedded (FFPE) tissues.

Ninety-one patients with clinical endpoints of metastasis or PCSD during follow-up represented cases. Metastases were confirmed with either bone scan or positron emission tomography-computed tomography in all but one case, where the metastasis was biopsy-confirmed. Eighty-nine patients with matching baseline characteristics, but no endpoint-related events, were assigned as controls. The histological slides were re-evaluated by an expert uropathologist (TM) and the index tumors were annotated for messenger RNA (mRNA) extraction. Blinding and randomization were not relevant to this study setting, since the patients were chosen based on known clinical endpoints. Fifteen cases and five controls

were excluded based on incomplete clinical data, neoadjuvant treatment, or quality control (QC) flags in transcript analysis (Figure 1). The demographics of patients included in analyses are described in Table 1.

The study was approved by the institutional review board of HUS (HUS/1439/2018) and the National Supervisory Agency for Health and Welfare (Dnro V/38176/2018). The data was handled in accordance with the national laws and EU regulations. Since the study was conducted on registry data, no express consent was required from patients, based on the national legislation.

2.2 | Messenger RNA extraction and transcript analysis

One or two 1 mm diameter punches were extracted from FFPE specimens. After deparaffinization, homogenization, and proteinase K digestion, tissues were transferred to 72-well plates and mRNA was extracted using QIASymphony (QIAGEN) RNA kit according to manufacturer guidelines. Yield and concentration of RNA was assessed using RiboGreen kit (Invitrogen) and the integrity of mRNA
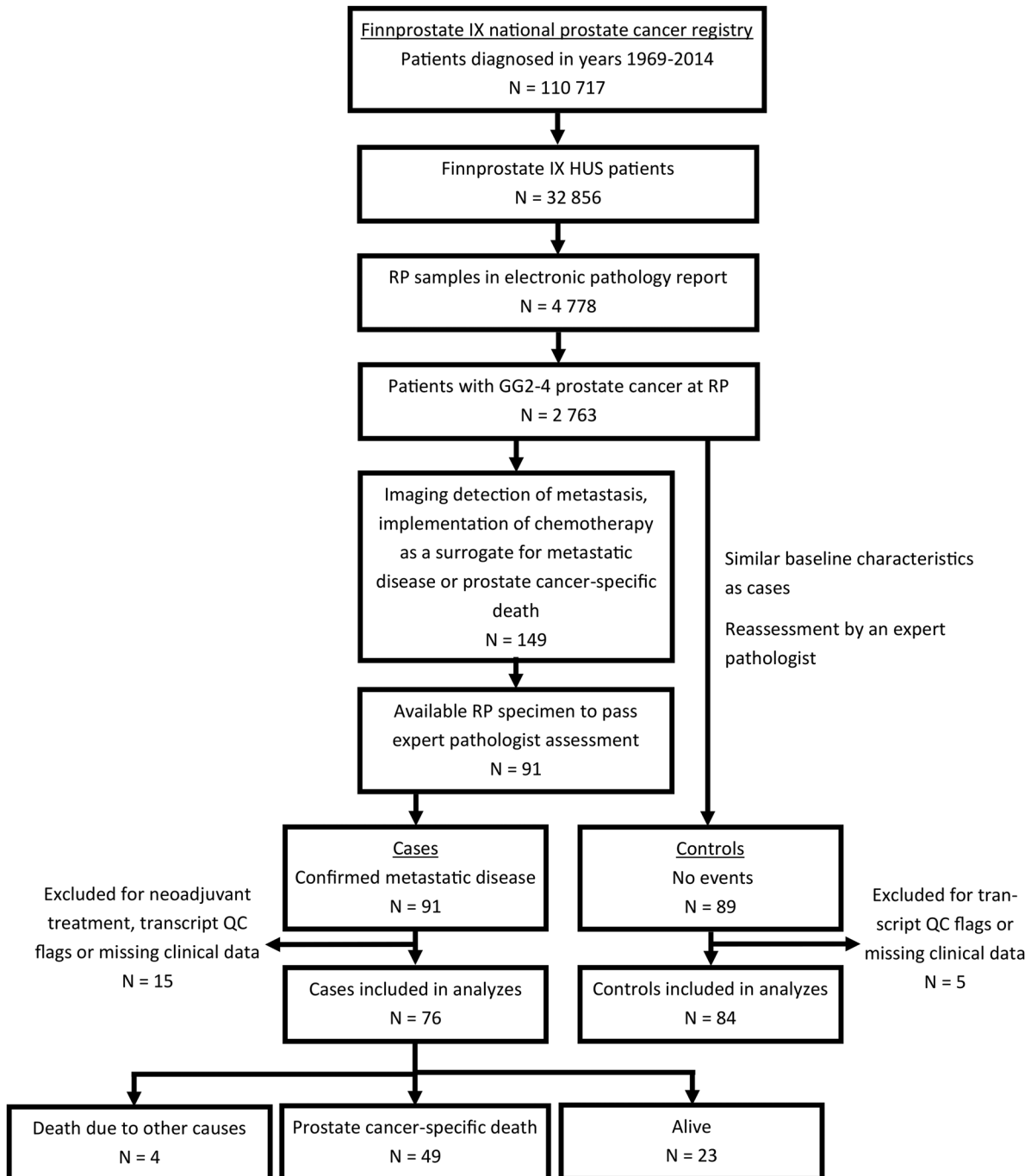
FIGURE 1 REMARK diagram of patients. Cases with clinically relevant endpoints at follow-up after RP and controls with no events were selected from the registry of nationwide Finnprostate IX study cohort. GG, grade group; HUS, HUS Helsinki University Hospital; QC, quality control; RP, radical prostatectomy

TABLE 1 Demographics of the study patients

Characteristics	Case	Control
Primary treatment (N)		
RP	76	84
Age at RP (years)		
Median (IQR)	62 (9.7)	63 (8.0)
PSA at RP		
Median (IQR)	9.5 (6.0)	9.0 (7.0)
NA	0	2
Grade group (GS)		
2 (3 + 4)	22	40
3 (4 + 3)	38	28
4 (8)	16	16
pT stage		
T2	24	32
T3a	23	31
T3b	29	21
Positive surgical margins	38	32
Positive local lymph nodes at RP	10	6
Lymphadenectomy at RP	55	57
Follow-up time (years)		
To metastasis: Median (IQR)	5.6 (4.4)	NA
To PCSD: Median (IQR)	8.6 (5.8)	NA
Total follow-up: Median (IQR)	10.3 (5.7)	11.6 (5.8)
First clinical endpoint		
Metastasis	76	0
No endpoint	0	84
Implementation of chemotherapy	49	0
Vital status		
Alive	23	78
Dead from prostate cancer	49	0
Dead due to other causes	4	6

Note: All pathological stages were recoded according to American Joint Committee on Cancer (AJCC) staging 8th edition for prostate cancer. Total follow-up = the total follow-up defined as time to death for deceased patients, and right censored survival time for alive patients. Abbreviations: GS, Gleason score; IQR, interquartile range; NA, PSA information was not available; PCSD, prostate cancer-specific death; PSA, prostate-specific antigen; pT stage, tumor stage indicated by a pathologist; RP, radical prostatectomy; T2, organ confined prostate cancer; T3a, the cancer has invaded through the capsule surrounding the prostate; T3b, the cancer has invaded the seminal vesicles.

was tested with Agilent Bioanalyzer kit (Agilent Technologies). For the samples included in the final analyses, the mean RNA integrity number was 2.22, with the range of 1.00–5.80.

NanoString nCounter (NanoString Technologies) analysis was performed at the DNA Sequencing and Genomics Laboratory, Institute of Biotechnology, University of Helsinki, Finland. A total of 100 ng of RNA in 5 µl H_2O was used for each analysis. RNA was incubated overnight with Reporter and Capture CodeSet probes specific to 794 cancer-related and six housekeeping genes in the custom-designed NanoString CodeSet (Table S1), which contains the genes of the investigated panels. Three sequences in Decipher have not been publicly disclosed and were not analyzed in the study. After hybridization, the samples were loaded into nCounter Prep Station where excess probes were removed, sample-probe complexes were captured, aligned and barcodes counted in nCounter Digital Analyzer. The housekeeping genes were *ACTB, ALAS1, CLTC, GUSB, HPRT1,* and *TUBB. HPRT1* was not used in normalization because of inconsistent expression across samples.

2.3 | Statistical analysis

Statistical analyses were performed in nSolver 4.0.70 Analysis Software (NanoString Technologies) and R, v.4.0.2 (R Development Core Team). Detailed methods can be found in supplementary material. Briefly, standard QC procedures were run in nSolver according to NanoString guidelines and the raw mRNA transcript copy count data was exported to R. The counts were variance stabilized and factors of unwanted variations were removed.

Performance of the selected gene panels to predict metastasis and PCSD was analyzed using random forest models (RFMs). The commercial panel genes and the clinical parameters PSA, pathological stage and GG, were considered as feature sets. For the two control samples with missing PSA values, median PSA in the control group was imputed. The parameters mtree and ntry were tuned and three times repeated 10-fold cross-validation was used in training the models as the resampling method for each tested parameter combination. Model accuracy was used to determine the optimal mtry and ntree values. The performance of final models was compared using area under the curve (AUC) for receiver operating characteristic (ROC). DeLong test was used to compare the ROC curves produced by the final RFMs and to calculate confidence intervals for AUC metrics. Heatmaps depicting the expression of genes included in each panel with dendrograms showing hierarchical clustering of genes and samples, were generated. Sensitivities and specificities of panel-based ROC models were calculated. Youden index (Y) and number needed to diagnose (NND = 1/Y) were also computed.

Survival analyses with Kaplan–Meier (KM) curves and log rank statistics were performed for the commercial panel feature sets and clinical parameters. The samples were grouped based on the RFM predictions of metastatic disease and PCSD, respectively. The ground truth observations of metastatic disease or PCSD were used as events in respective analyses. For cases, the time-to-event was defined as days between RP and confirmation of metastases or death respectively. For controls, the right-censored survival time was

defined as days between RP and last laboratory visit or last outpatient clinic control. p Values of the log-rank test and 95% confidence intervals were calculated.

3 | RESULTS

We used RFMs to analyze the effectiveness of the commercial panel genes, as feature sets, in predicting outcomes after RP in our patient material. To establish a baseline for comparison, we also constructed RFMs for metastatic disease and PCSD, using clinical parameters (PSA, pathological stage, and GG) as predictors. The performance of each model was assessed with AUC. As expected, the performance of the studied clinical parameters was poor, since the cohort was selected based on equal baseline characteristics in the most important clinical parameters. All commercial panel gene sets significantly outperformed clinical parameters in predicting metastatic disease and PCSD (p < .005). Decipher, Prolaris, and MSK-IMPACT gene sets showed similar AUC metrics (p > .05 between all panels) in metastasis prediction, while Oncotype DX was a significantly poorer predictor of metastatic disease (Figure 2A). The genes in Prolaris showed significantly worse PCSD prediction performance (Figure 2B) than those in Decipher (p = .046) and MSK-IMPACT (p = .037). Statistical significance was not observed in other ROC comparisons.

Heatmaps depicting the expression of genes in the panels between study groups as well as ROC model predictions are shown in Figure S1. Based on analysis of gene expression, the cases and controls did not strongly separate into clusters. Notably, the clinical parameter-based models were unable to predict PCSD correctly for any patient.

We calculated the sensitivity, specificity, Youden index and NND of each panel in predicting metastasis and PCSD (Table S2). Decipher showed highest sensitivity (0.62) (specificity = 0.70) and Prolaris highest specificity (0.77) (sensitivity = 0,58) in metastasis prediction. In PCSD prediction, the sensitivity of all panels was very low (0.29–0.31) and specificity high (0.89–0.97). Decipher, MSK-IMPACT and Prolaris showed similar NND (2.83–3.12) in metastasis prediction while Oncotype DX showed lower performance (4.79). Conversely, Oncotype DX showed lowest NND (2.79) in PCSD prediction. Again, as expected, all panels outperformed the clinical parameters in terms of NND. NND could not be calculated for clinical parameters when predicting PCSD, since they did not predict any PCSD endpoints correctly.

We analyzed metastasis-free and prostate cancer-specific survival by plotting KM. The predictions of RFMs trained using genes in Decipher, MSK-IMPACT, and Prolaris showed statistically significant (p < .05) separation of strata in metastasis-free survival probability (Figure 3). The predictions of all panel based RFMs separated strata significantly in prostate cancer-specific survival (Figure 4). KM plots based on RFMs trained using clinical parameters showed significant separation of strata in metastasis-free survival prediction, but not in prostate cancer-specific survival prediction (Figure S2).
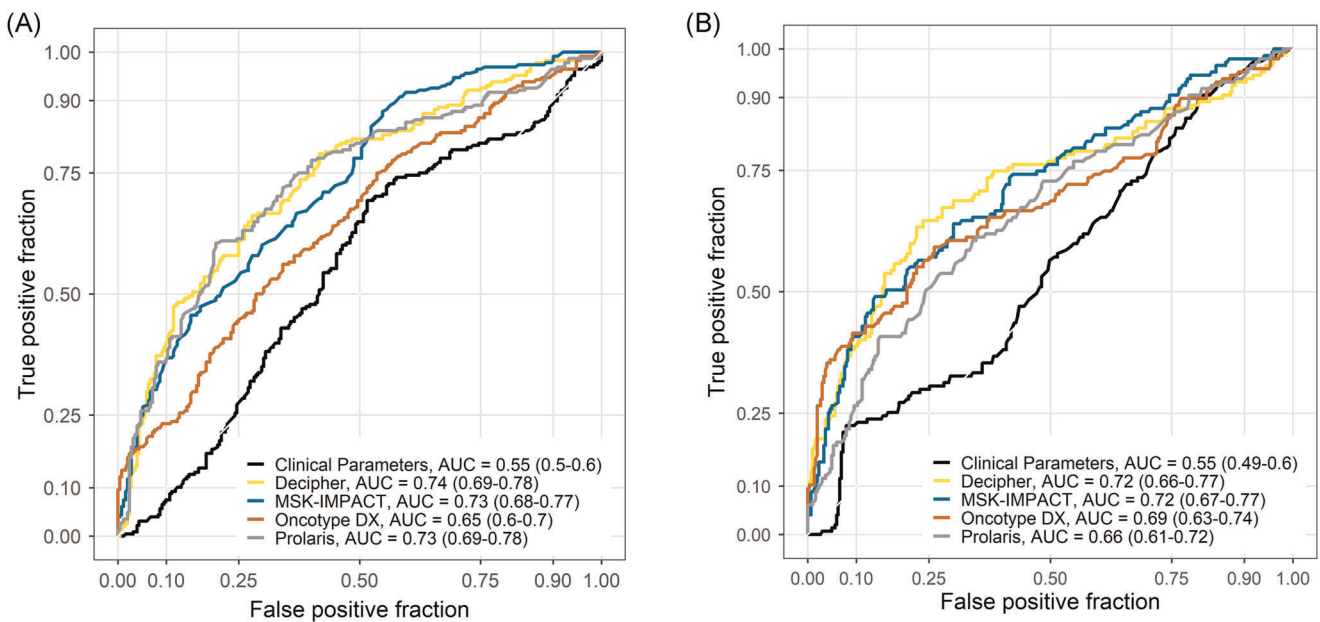
FIGURE 2 Predictive performance of the clinical parameters PSA, cancer stage and grade group, the commercial risk stratification panels and MSK-IMPACT. (A) Receiving operator characteristic (ROC) curves for predicting metastatic disease during follow-up plotted for the clinical parameters, Decipher, MSK-IMPACT, Oncotype DX, and Prolaris. (B) ROC curves for prostate cancer-specific death prediction during follow-up. In parentheses, 95% confidence intervals are shown. AUC, area under the curve; PSA, prostate-specific antigen [Color figure can be viewed at wileyonlinelibrary.com]

4 | DISCUSSION

We evaluated the ability of the genes present in the commercially available prostate cancer risk stratification panels Decipher, Oncotype DX, Prolaris, and the pan-cancer mutational panel MSK-IMPACT to predict clinically relevant outcomes in a cohort of RP treated patients harboring intermediate- to high-risk prostate cancer, as defined by GG, using a case-control approach. Previously, these panels have been shown to have independent prognostic value in cohorts with high number of low- and high-risk prostate cancer and relatively few patients with intermediate-risk disease. While patients in such low- and high-risk groups can be risk stratified with the existing clinical variables, such as PSA, GG, and stage, the cohort of mostly intermediate-risk patients studied here represents the unmet clinical need.

In this study we specifically wanted to address the major clinical need, which is the stratification of patients harboring GG 2–4

prostate cancer. Thus, with the exception of Decipher, the studied endpoints do not directly reflect the surrogate endpoints used in the development of these risk stratification panels. Progression to metastatic disease is the most direct surrogate marker for lethal prostate cancer. Thus, predictive information about lethal outcomes would strongly influence clinical decision-making and are thus important to address when considering possible adoption of new risk stratification tools to clinical practice.

One of the proposed indications for the studied commercial panels is stratifying select patients with GG 1 as well as GG 2 prostate cancer into active surveillance and active treatment groups.[5] The emerging consensus is, that these molecular tests should only be used in addition to clinical parameters and only in situations, where they are likely to affect clinical decision-making.[5,13] The results presented in this paper cannot be used to draw conclusions on the comparative performance of the panels in risk stratifying GG 1 patients to active surveillance and active treatment, as no clinically low-risk patients were included in the
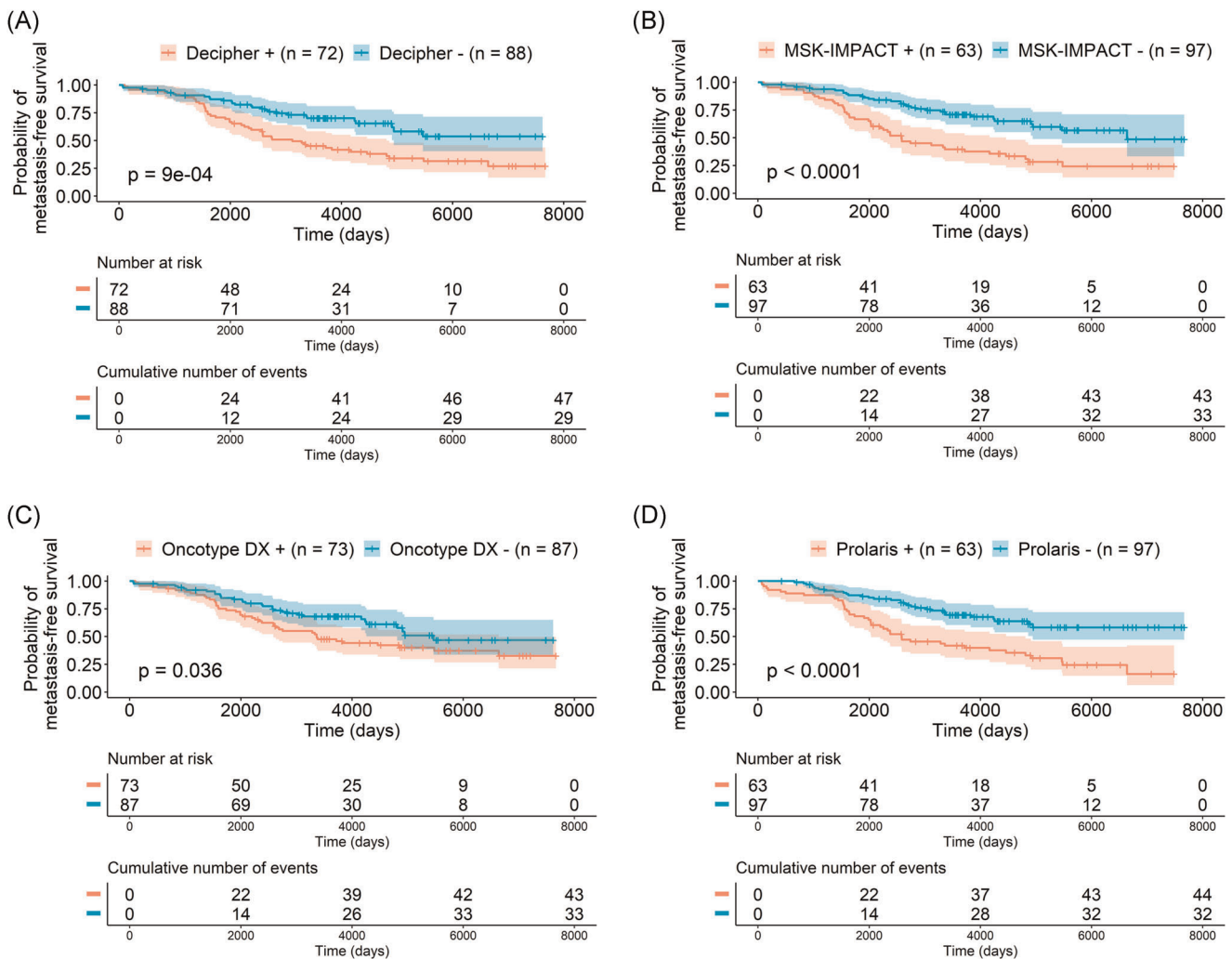
FIGURE 3 Metastasis-free survival probability plotted for the panels (A) Decipher, (B) MSK-IMPACT, (C) Oncotype DX, and (D) Prolaris. Patients were stratified into metastasis positive and negative groups by random forest model prediction. p Values of the log-rank test and 95% confidence intervals are shown. The number of predicted outcomes is displayed in the legend. The number of patients at risk and the cumulative number of events are shown in tables. +, metastatic disease; −, non-metastatic disease [Color figure can be viewed at wileyonlinelibrary.com]
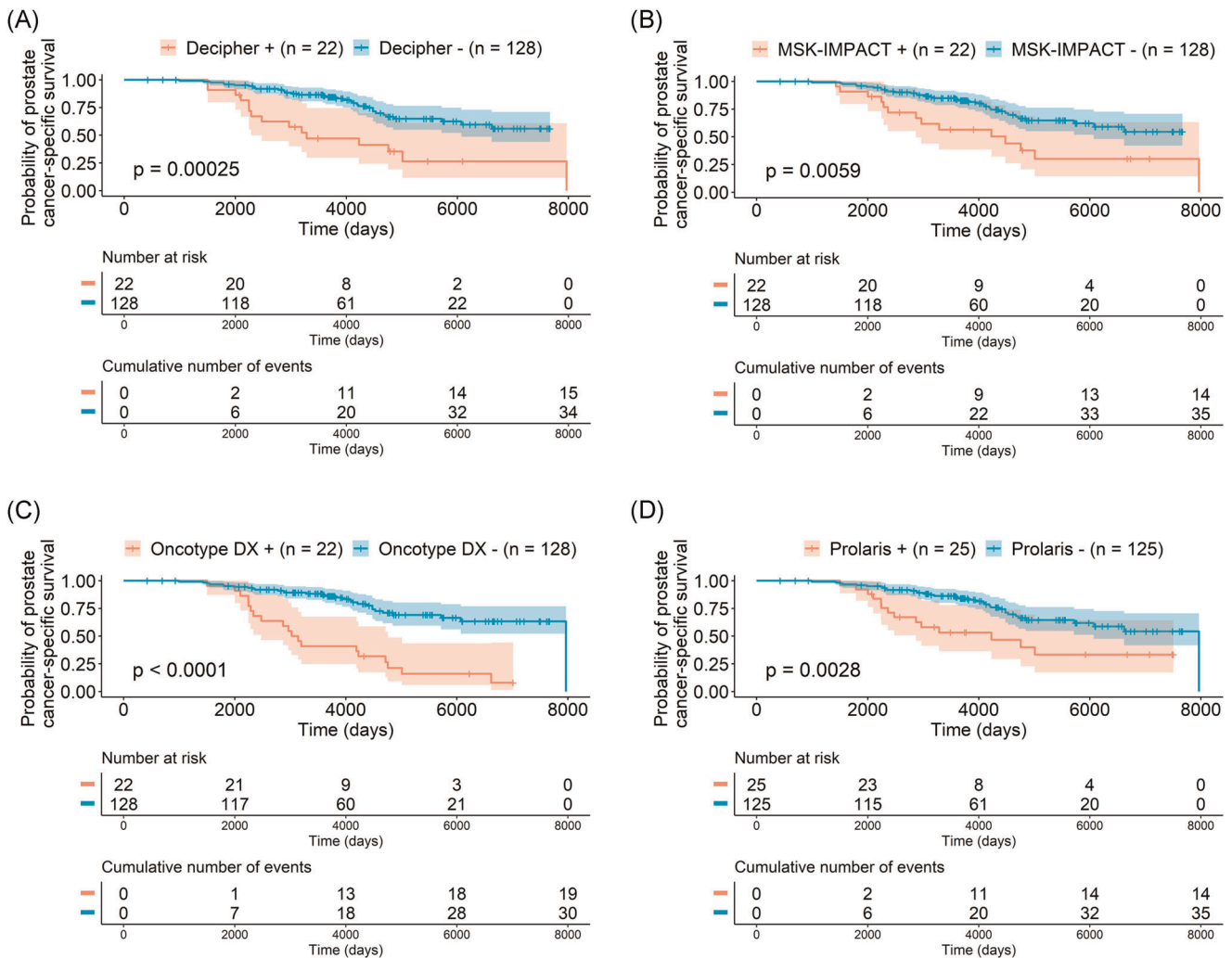
FIGURE 4 Prostate cancer-specific survival probability plotted for the panels (A) Decipher, (B) MSK-IMPACT, (C) Oncotype DX, and (D) Prolaris. Patients were stratified into two groups based on the random forest model prediction of prostate cancer-specific death. *p* Values of the log-rank test and 95% confidence intervals are displayed. The number of predicted outcomes is shown in the legend. The number of patients at risk and the cumulative number of events are shown in the tables. +, prostate cancer-specific death; −, alive patients [Color figure can be viewed at wileyonlinelibrary.com]

study. Additionally, the sample size of patients with GG 2 disease is not large enough to draw significant conclusions in this sub-cohort. The current consensus suggests that patients with GG 3–4 disease should not be managed with active surveillance.[5]

We are among the first independent groups to compare these panels in patients with primarily intermediate-risk prostate cancer using clinically relevant endpoints. While the genes in the studied panels were found to independently contribute to risk stratification, their performance here was modest compared to earlier publications.[12,18–21] However, it should be noted, that the studied cohort was extremely challenging, as it was selected based on similar baseline clinical characteristics in cases and controls. Thus, the clinical characteristics were not able to predict the study endpoints and any additional predictive ability, as shown here by the panel transcripts, may be clinically useful in this group of mostly intermediate-risk prostate cancer patients.

The RF models we trained based on the gene expression of each gene panel showed low sensitivity to predict either progression to metastasis or PCSD. However, their specificity, especially in PCSD prediction was high. Decipher, MSK-IMPACT, and Prolaris demonstrated comparable and lower NND in metastasis prediction than Oncotype DX. Conversely, Oncotype DX showed the lowest NND in PCSD prediction. Sensitivity, specificity and NND are not dependent on the prevalence of the studied endpoints. Further studies should be carried out to reproduce these results in contemporary cohorts. Additionally, research with larger, multicenter intermediate-risk cohorts and possibly full commercial assays should be done to analyze the value of these genes in guiding post-prostatectomy treatment decisions, such as early adjuvant therapy or salvage radiation therapy. Although we did not utilize the full commercial assays, our results suggest that the panel gene sets do have predictive value. Thus, a more intensive post-prostatectomy treatment regimen might be

considered for patients with clinically intermediate-risk prostate cancer and high Decipher genomic classifier, Oncotype DX genomic prostate score, or Prolaris cell cycle progression score.

The Oncotype DX genomic prostate score has already been found to be cost-effective when incorporated into treatment decision-making for patients with very low- to low-risk disease according to NCCN guidelines.[22] Decipher has been found to be more cost effective than treating all patients with adjuvant therapy, but more expensive than the usual adjuvant therapy rates, whilst improving treatment results.[23] Prolaris has been found to largely increase overall costs with little savings from active surveillance protocol improvements.[24] More multicenter studies on the cost-effectiveness of all commercial panels are critical before adoption to routine clinical use.[5]

By study design, we controlled the effect of clinical variables and interdependency of studied genes with independent clinical features of preoperative PSA, GG, and pathological stage. We used RP samples, instead of diagnostic biopsies. This may generate bias towards aggressive disease, although we controlled clinical and pathological variables. However, using RP samples allowed us to extract mRNA from the index tumors. Because of this, we could not account for intra-tumoral or inter-tumoral genetic heterogeneity, which can lead to misinterpretation of genomic classifiers.[25] It should also be noted that the proportion of metastases and PCSD endpoints is not reflective of the full patient population harboring GG 2–4 disease. A larger sample size of metastatic and lethal cancer was included in the study to provide statistical power.

Extensive clinicopathological data strengthens the findings of this study. Information on the pre- and postoperative treatments were collected for all specimens. Since none of the patients received neoadjuvant treatment, our results may be projected to what is detectable in diagnostic biopsies, apart from the potential selection bias for performing RP. The retrospective study cohort, despite its limited size, is extremely well-characterized and represents the clinically most challenging to predict patients, for whom new prognostic markers and risk stratifications are truly needed to improve quality of life and prognosis as well as to minimize over- and undertreatment.

Each of the proprietary panels have their own laboratory protocols, housekeeping genes and, apart from MSK-IMPACT mutational panel, an output score. We did not use any proprietary protocols or methodology specific to any of the studied panels. Since we did not use the full commercial assays, but only the publicly available gene sets of the panels, we could not use the proprietary weighting algorithms of individual genes. We also used the NanoString nCounter platform for mRNA analysis. This inter-technology variation may decrease the impact of our comparative transcript-level study, yet it reduces the possible confounding introduced by proprietary or panel-specific methods.

Despite the use of a different laboratory protocol, analysis methodology, and partially endpoints, the prognostic ability of these gene sets was demonstrated in our study. This suggests that the gene sets themselves are valuable, although not exceptional in predicting clinically significant endpoints in prostate cancer independent of a given test methodology. Since the studied panels share almost no genes and have independent predictive value, novel combinations of these genes, as well as others, might offer even better performance especially in difficult to stratify patient materials.

Prostate cancer diagnostics are shifting towards magnetic resonance imaging and targeted biopsies that are known to better sample clinically significant prostate cancer.[26] The index tumors used here may represent, better than sextant biopsies, tissues typically sampled with targeted biopsies.[27] Whether the performance of any biomarker test done using targeted biopsies outperforms the same analysis done on tissue from sextant biopsies can only be studied once the cohorts with targeted biopsies are mature enough to study clinically relevant endpoints.

5 | CONCLUSION

The genes present in the commercial risk stratification panels did not perform particularly well in this clinically significant and challenging intermediate-risk group of patients, irrespective of the number of genes in the panels. However, they significantly outperformed the clinical parameters PSA, cancer stage, and grade, which may have implications in guiding post-prostatectomy treatment decisions in localized intermediate-risk prostate cancer. The commercial panels should be prospectively validated in a contemporary cohort, preferably combined with targeted biopsies. Novel, more predictive gene sets, should also be investigated in intermediate-risk prostate cancer.

CONFLICT OF INTERESTS
The authors declare that there are no conflict of interests.

DATA AVAILABILITY STATEMENT
Data is available from the corresponding author (Mirtti T.) upon reasonable request.The data was handled in accordance with the national laws and EU regulations.

ORCID

Timo-Pekka K. Lehto http://orcid.org/0000-0003-3262-2333

Carolin Stürenberg http://orcid.org/0000-0001-7033-7344

Andrew M. Erickson http://orcid.org/0000-0002-4850-4086

Hannu Koistinen http://orcid.org/0000-0003-0926-3109

Ian G. Mills http://orcid.org/0000-0001-5347-5083

Antti Rannikko http://orcid.org/0000-0002-4261-3484

Tuomas Mirtti http://orcid.org/0000-0003-0455-9891

REFERENCES

1. Stark JR, Perner S, Stampfer MJ, et al. Gleason score and lethal prostate cancer: does 3 + 4 = 4 + 3? *J Clin Oncol.* 2009;27(21):3459-3464.

2. Epstein JI, Egevad L, Amin MB, et al. The 2014 International Society of Urological Pathology (ISUP) Consensus Conference on Gleason grading of prostatic carcinoma: definition of grading patterns and proposal for a new grading system. *Am J Surg Pathol.* 2016;40(2):244-252.

3. Pierorazio PM, Walsh PC, Partin AW, Epstein JI. Prognostic Gleason grade grouping: data based on the modified Gleason scoring system. *BJU Int.* 2013;111(5):753-760.

4. Romero Otero J, Garcia Gomez B, Campos Juanatey F, Touijer KA. Prostate cancer biomarkers: an update. *Urol Oncol.* 2014;32(3):252-260.

5. Eggener SE, Rumble RB, Armstrong AJ, et al. Molecular biomarkers in localized prostate cancer: ASCO guideline. *J Clin Oncol.* 2020;38(13):1474-1494.

6. Bishoff JT, Freedland SJ, Gerber L, et al. Prognostic utility of the cell cycle progression score generated from biopsy in men treated with prostatectomy. *J Urol.* 2014;192(2):409-414.

7. Cuzick J, Berney DM, Fisher G, et al. Prognostic value of a cell cycle progression signature for prostate cancer death in a conservatively managed needle biopsy cohort. *Br J Cancer.* 2012;106(6):1095-1099.

8. Spratt DE, Yousefi K, Deheshi S, et al. Individual patient-level meta-analysis of the performance of the decipher genomic classifier in high-risk men after prostatectomy to predict development of metastatic disease. *J Clin Oncol.* 2017;35(18):1991-1998.

9. Cooperberg MR, Davicioni E, Crisan A, Jenkins RB, Ghadessi M, Karnes RJ. Combined value of validated clinical and genomic risk stratification tools for predicting prostate cancer mortality in a high-risk prostatectomy cohort. *Eur Urol.* 2015;67(2):326-333.

10. Klein EA, Haddad Z, Yousefi K, et al. Decipher genomic classifier measured on prostate biopsy predicts metastasis risk. *Urology.* 2016;90:148-152.

11. Kim HL, Li P, Huang HC, et al. Validation of the decipher test for predicting adverse pathology in candidates for prostate cancer active surveillance. *Prostate Cancer Prostatic Dis.* 2019;22(3):399-405.

12. Klein EA, Cooperberg MR, Magi-Galluzzi C, et al. A 17-gene assay to predict prostate cancer aggressiveness in the context of Gleason grade heterogeneity, tumor multifocality, and biopsy undersampling. *Eur Urol.* 2014;66(3):550-560.

13. Lin DW, Zheng Y, McKenney JK, et al. 17-Gene Genomic prostate score test results in the canary prostate active surveillance study (PASS) cohort. *J Clin Oncol.* 2020;38(14):1549-1557.

14. Cullen J, Rosner IL, Brand TC, et al. A biopsy-based 17-gene genomic prostate score predicts recurrence after radical prostatectomy and adverse surgical pathology in a racially diverse population of men with clinically low- and intermediate-risk prostate cancer. *Eur Urol.* 2015;68(1):123-131.

15. Van Den Eeden SK, Lu R, Zhang N, et al. A biopsy-based 17-gene genomic prostate score as a predictor of metastases and prostate cancer death in surgically treated men with clinically localized disease. *Eur Urol.* 2018;73(1):129-138.

16. Abida W, Armenia J, Gopalan A, et al. Prospective genomic profiling of prostate cancer across disease states reveals germline and somatic alterations that may affect clinical decision making. *JCO Precis Oncol.* 2017:2017-16.

17. Alam S, Tortora J, Staff I, McLaughlin T, Wagner J. Prostate cancer genomics: comparing results from three molecular assays. *Can J Urol.* 2019;26(3):9758-9762.

18. Cucchiara V, Cooperberg MR, Dall'Era M, et al. Genomic markers in prostate cancer decision making. *Eur Urol.* 2018;73(4):572-582.

19. Spratt DE, Dai DLY, Den RB, et al. Performance of a prostate cancer genomic classifier in predicting metastasis in men with prostate-specific antigen persistence postprostatectomy. *Eur Urol.* 2018;74(1):107-114.

20. Erho N, Crisan A, Vergara IA, et al. Discovery and validation of a prostate cancer genomic classifier that predicts early metastasis following radical prostatectomy. *PLOS One.* 2013;8(6):e66855.

21. Fine ND, LaPolla F, Epstein M, Loeb S, Dani H. Genomic classifiers for treatment selection in newly diagnosed prostate cancer. *BJU Int.* 2019;124:578-586.

22. Albala D, Kemeter MJ, Febbo PG, et al. Health economic impact and prospective clinical utility of oncotype DX(R) genomic prostate score. *Rev Urol.* 2016;18(3):123-132.

23. Lobo JM, Trifiletti DM, Sturz VN, et al. Cost-effectiveness of the decipher genomic classifier to guide individualized decisions for early radiation therapy after prostatectomy for prostate cancer. *Clin Genitourin Cancer.* 2017;15(3):e299-e309.

24. Health Quality O. Prolaris cell cycle progression test for localized prostate cancer: a health technology assessment. *Ont Health Technol Assess Ser.* 2017;17(6):1-75.

25. Wei L, Wang J, Lampert E, et al. Intratumoral and intertumoral genomic heterogeneity of multifocal localized prostate cancer impacts molecular classifications and genomic prognosticators. *Eur Urol.* 2017;71(2):183-192.

26. Kasivisvanathan V, Stabile A, Neves JB, et al. Magnetic resonance imaging-targeted biopsy versus systematic biopsy in the detection of prostate cancer: a systematic review and meta-analysis. *Eur Urol.* 2019;76(3):284-303.

27. Erickson AM, Lokman U, Lahdensuo K, et al. PTEN and ERG expression in MRI-ultrasound guided fusion biopsy correlated with radical prostatectomy findings in men with prostate cancer. *Prostate.* 2020;80:1118-1127.

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.