IMPORTANT REMINDER

Medical Policies are developed to provide guidance for members and providers regarding coverage in accordance with contract terms. Benefit determinations are based in all cases on the applicable contract language. To the extent there may be any conflict between the Medical Policy and contract language, the contract language takes precedence.

PLEASE NOTE: Contracts exclude from coverage, among other things, services or procedures that are considered investigational or cosmetic. Providers may bill members for services or procedures that are considered investigational or cosmetic. Providers are encouraged to inform members before rendering such services that the members are likely to be financially responsible for the cost of these services.

DESCRIPTION

There are a variety of gene-based biomarkers that have been associated with prostate cancer. These tests have the potential to improve the accuracy of risk prediction, diagnosis, staging, or prognosis of prostate cancer.

MEDICAL POLICY CRITERIA

Genetic tests for the screening, detection, and management of prostate cancer are considered **investigational** including but not limited to the following:

A. Single-nucleotide polymorphisms (SNPs) for risk assessment;
B. PCA3 for disease diagnosis;
C. TMPRSS fusion genes for diagnosis and prognosis;
D. Gene hypermethylation for diagnosis and prognosis;
E. Mitochondrial DNA mutation testing for diagnosis;
F. Gene expression analysis
NOTE: A summary of the supporting rationale for the policy criteria is at the end of the policy.

CROSS REFERENCES

1. Genetic and Molecular Diagnostic Testing, Genetic Testing, Policy No. 20
2. Gene Expression-Based Assays for Cancers of Unknown Primary, Genetic Testing, Policy No. 15
3. Gene Expression Profiling for Melanoma, Genetic Testing, Policy No. 29
4. Protein Biomarkers for Screening, Detection, and/or Management of Prostate Cancer, Laboratory, Policy No. 69

BACKGROUND

Prostate cancer is a complex, heterogeneous disease. At the extremes of the spectrum, if left untreated, some prostate cancers behave aggressively, metastasize quickly, and cause mortality, while others are indolent and never progress to cause harm. Current challenges in prostate cancer care are assessing risk; providing early and accurate detection; monitoring low-risk patients undergoing surveillance only; predicting recurrence after initial treatment; detecting recurrence after treatment; and assessing efficacy of treatment for advanced disease.

In response to the need for better biomarkers for risk assessment, diagnosis, and prognosis, a variety of exploratory research is ongoing. Some products of this work have already been translated or are in the process of being translated into commercially available tests, including:

- Single-nucleotide polymorphisms (SNPs) for risk assessment
- The Gen-Probe PROGENSA® PCA3 Assay (PCA3) for diagnosis
- TMPRSS fusion genes for diagnosis and prognosis
- Gene hypermethylation for diagnosis and prognosis
- Mitochondrial DNA mutation testing for diagnosis
- Gene expression analysis for risk assessment and diagnosis

While studies using these tests generate information that may help elucidate the biologic mechanisms of prostate cancer and eventually help design treatments, the above-mentioned tests are currently in a developmental phase, without evidence of clinical utility for diagnosis, prognosis, or risk assessment. Many of the tests listed above have not been submitted to the U.S. Food and Drug Administration (FDA) for marketing clearance but, if available, are offered as laboratory-developed tests by Clinical Laboratory Improvement Amendments (CLIA) licensed laboratories.

SNP testing as part of genome-scanning tests with risk assessment for prostate cancer are offered by a variety of laboratories including Navigenics, LabCorp (23andme), and ARUP (deCode) as laboratory-developed tests. The PCA3 test is offered in the U.S. by a number of reference laboratories including ARUP, Mayo Medical Laboratories, and LabCorp. The reagents used in testing are developed by Gen-Probe. A test for hypermethylation of GSTP1 was available from LabCorp (“Glutathione S-transferase Gene [GSTP1, pi-class] Methylation Assay”); but as of January 2015 this test is no longer offered. Epigenomics AG has entered licensing agreements with one U.S. laboratory (Quest Diagnostics) to establish and commercialize laboratory-developed tests for its proprietary methylation biomarker GSTP1. This test is not yet available, and it is unclear what matrices will be used.
ConfirmMDx is offered from MDxHealth. The tissue-based DNA methylation multigene assay aims to improve stratification of men being considered for repeat prostate biopsy. Hypermethylation of GSTP1, APC, and RASSF1 are assessed in core biopsy samples.

SelectMDx for Prostate Cancer is also offered from MDxHealth. The reverse transcription PCR (RT-PCR) assay is performed on post-DRE (digital rectal examination), first-void urine specimens from patients with clinical risk factors for prostate cancer, who are being considered for biopsy. The test measures the mRNA levels of the DLX1 and HOXC6 biomarkers, using KLK3 expression as internal reference, to aid in patient selection for prostate biopsy.

REGULATORY STATUS

One mitochondrial DNA test, Mitomics (Broomfield, CO), is currently available. Mitomics offers the Prostate Core Mitomics Test which measures mitochondrial DNA mutations in a negative prostate biopsy to determine whether a patient should undergo repeat biopsy. The test is performed on the initial negative prostate biopsy tissue.

The PCA3 Assay was approved by the FDA on February 15, 2012 through the premarket approval process. According to the approval granted by the FDA:[1]

“The PROGENSA PCA3 Assay is indicated for use in conjunction with other patient information to aid in the decision for repeat biopsy in men 50 years of age or older who have had one or more previous negative prostate biopsies and for whom a repeat biopsy would be recommended by a urologist based on current standard of care, before consideration of PROGENSA PCA3 assay results.”

Prolaris®, Oncotype DX® Prostate, and Decipher® gene expression profiling test are available under the auspices of CLIA. Laboratories that offer LDTs must be licensed by CLIA for high-complexity testing. To date, the U.S. Food and Drug Administration (FDA) has chosen not to require any regulatory review of this test.

In November 2015, the FDA’s Office of Public Health Strategy and Analysis published a document on public health evidence for FDA oversight of LDTs.[2] FDA argued that many tests need more FDA oversight than the regulatory requirements of CLIA. CLIA standards relate to laboratory operations, but do not address inaccuracies or unreliability of specific tests. Prolaris is among the 20 case studies in the document cited as needing FDA oversight. The document asserted that patients are potentially receiving inappropriate prostate cancer care because there is no evidence that results from the test meaningfully improve clinical outcomes.

The other tests mentioned in this policy, if available, are offered as laboratory-developed tests under the Clinical Laboratory Improvement Amendments (CLIA) licensed laboratories.

EVIDENCE SUMMARY

Human Genome Variation Society (HGVS) nomenclature[3] is used to describe variants found in DNA and serves as an international standard. It is being implemented for genetic testing medical evidence review updates starting in 2017. According to this nomenclature, the term “variant” is used to describe a change in a DNA or protein sequence, replacing previously-used terms, such as “mutation.” Pathogenic variants are variants associated with disease, while benign variants are not. The majority of genetic changes have unknown effects on human health, and these are referred to as variants of uncertain significance.
In general, the evidence for genetic tests related to prostate cancer screening, detection, and management addresses either preliminary clinical associations between genetic tests and disease states or, in some cases, the clinical validity of these tests i.e., the association of the test result with outcomes of interest, expressed in terms of clinical performance characteristics such as sensitivity, specificity, predictive value, and comparisons to current standards using receiver-operating curve (ROC) analysis and/or logistic regression. There is no published evidence demonstrating clinical utility (i.e., a test will change treatment decisions and improve patient important outcomes).

GENE-BASED TESTS IN GENERAL

A 2009 BlueCross BlueShield Association (BCBSA) TEC Special report of recently published studies on gene-based tests (SNPs, PCA3, TMPRSS, gene panels, and gene hypermethylation) for prostate cancer risk assessment and diagnosis concluded that, in general, research on these tests is still in a “developmental phase, currently without evidence of clinical utility.”[4] This policy was initially based on a 2013 BlueCross BlueShield Association Technology Evaluation Center (TEC) Assessment which was updated in January 2015 with a literature review through September 30, 2014.[5,6] Full-length publications were sought that described the analytic validity (technical performance), clinical validity (prognostic accuracy), and clinical utility (accurately identifying men experiencing improved health outcomes by avoiding treatment or undergoing more appropriate therapies) of Prolaris, Oncotype DX Prostate, and Decipher gene expression profiling. The Blue Cross Blue Shield Association Medical Advisory Panel also reviewed the evidence in September 2017.

SINGLE-NUCLEOTIDE POLYMORPHISMS (SNPS) FOR PROSTATE CANCER RISK ASSESSMENT AND PROGNOSIS

There have been numerous large observational correlational studies focusing on the association of many different SNPs with prostate cancer, an example of which includes the study by Lindstrom of 10,501 cases of prostate cancer and 10,831 controls, which identified 36 SNPs showing association with prostate cancer risk including two (rs2735893 and rs266849) showing differential association with Gleason grade. Per allele odds ratios ranged from 1.07 to 1.44.[7]

Because the SNPs individually provide relatively modest incremental information on both the occurrence of cancer and its behavior, investigators have begun to explore use of algorithms incorporating information from multiple SNPs to increase the clinical value of testing. Several such recent studies focused on the development of testing algorithms incorporating SNPs.[8-11]

Systematic Reviews

A systematic review of multigene panels for prostate cancer risk assessment was published by Little in 2016.[12] The authors included 21 studies that evaluated 18 individual panels. All studies were focused on clinical validity, moderate risk of bias, and had poor discriminative ability for predicting prostate cancer risk and/or distinguishing between aggressive and latent cancers. The authors noted that the current evidence is insufficient to assess analytic validity, and that “at best the panels assessed would add a small and clinically unimportant improvement” to current factors used for risk stratification, like age and family history. Additionally, they found no evidence on the clinical utility of these panels.
A 2012 AHRQ report on multigene panels in prostate cancer risk assessment, also by Little, reviewed the literature on SNP panel tests for assessing risk of prostate cancer.[13] All of the studies included in the review had poor discriminative ability for predicting risk of prostate cancer, had moderate risk of bias, and none of the panels had been evaluated in routine clinical settings. The conclusions of the review were that the evidence on currently available SNP panels does not permit meaningful assessment of analytic validity, the limited evidence on clinical validity is insufficient to conclude that SNP panels would perform adequately as a screening test and that there is no evidence available on the clinical utility of current panels.

Ishaak reviewed 11 replication studies involving 30 SNPs (19 in men of African descent and 10 in men with familial prostate cancer).[14] Odds ratios were positively associated with prostate cancer, although the magnitude of association was generally small (range, 1.11–2.63).

Al Olama conducted a meta-analysis of 4 GWAS including 5,953 cases of aggressive prostate cancer (PCa) and 11,463 controls (men without PCa).[15] Authors computed association tests for approximately 2.6 million SNPs and followed up the most significant SNPs by genotyping 49,121 samples in 29 studies through the international PRACTICAL and BPC3 consortia. The authors confirmed the association of a PCa susceptibility locus, rs11672691 on chromosome 19, but also showed an association with aggressive PCa [odds ratio = 1.12 (95% confidence interval 1.03-1.21), \( P = 1.4 \times 10^{-8} \)]. The authors concluded their report described a genetic variant which is associated with aggressive PCa, and which is a type of PCa associated with a poorer prognosis.

**Nonrandomized Studies**

A pilot study by Castro tested the use of a 71-SNP panel in 100 men with a family history of prostate cancer.[16] These men underwent a prostate biopsy regardless of PSA level, and 25 were diagnosed with prostate cancer. Age and PSA level were significantly associated with a cancer diagnosis, but the SNP risk score was not. While this study might not have been adequately powered to detect such an association, there was a clear relationship seen for age and PSA level (\( p = 0.00004 \) and \( 0.00037 \), respectively).

Kader evaluated a panel of 33 SNPs identified from GWAS associated with prostate cancer in 1,654 men.[17] Genetic score was a significant (\( p<.001 \)) independent predictor of prostate cancer, with an odds ratio of 1.72 (95% CI, 1.44-2.09) after adjustment for clinical variables and family history. Addition of genetic markers to the classification of prostate cancer risk resulted in 33% of men reclassified into a different risk quartile. Approximately half of these (n=267) were downgraded to a lower risk quartile and the other half (n=265) were upgraded into a higher risk quartile. The net reclassification benefit was 10% (\( p=0.002 \)). The authors concluded that with the additional information of genetic score the same number of cancers could be detected by using 15% fewer biopsies. However, this study includes a limited sample size and there is no clear indication of how clinical management changed when patients were reclassified into lower risk groups.

Ren calculated genetic scores for various combinations of 29 PCa risk-associated SNPs in 667 consecutive patients that underwent prostate biopsy.[18] Performance of these genetic scores for discriminating prostate biopsy outcomes were compared using the area under a receiver operating characteristic curve (AUC). The discriminative performance of genetic score derived from a panel of all 29 SNPs (24 previous and 5 new) was similar to that derived from the 24 previously established SNPs, the AUC of which were 0.60 and 0.61, respectively (\( P = 0.72 \)).
Authors concluded that genetic score based on PCa risk-associated SNPs implicated to date is a significant predictor of biopsy outcome.

Tsuchiya identified 14 SNPs in 6 genes (XRCC4, PMS1, GATA3, IL13, CASP8, and IGF1) that were statistically associated with cancer-specific survival.[19] Using a subset of 6 SNPs, 3 subgroups of men with prostate cancer were defined by the number of SNP’s present (0 to 1, 2 to 3, or 4 to 6). Median cancer-specific survival in these subgroups was 13.3, 7.0, and 3.8 years, respectively (log-rank test, p<0.001).

**Section Summary**

The available evidence generally reports a modest association with future risk for prostate cancer and/or prognosis in patients with prostate cancer. The clinical utility of these tests is uncertain, there is no evidence that the information obtained from SNP testing can be used to change management in ways that will improve outcomes.

**PCA3 FOR PROSTATE CANCER DIAGNOSIS**

PCA3 is overexpressed in prostate cancer and PCA3 mRNA can be detected in urine samples collected after prostate massage. When normalized using PSA to account for the amount of prostate cells released into the urine (PCA3 Score), the test has been proposed for use in discriminating between patients with eventual benign findings on (first or second) biopsies from those with malignant biopsy results. In particular, the test may be especially helpful at identifying patients with elevated PSA levels but negative first biopsy results who need a follow-up biopsy.

**Systematic Reviews**

In 2015, Nicholson published a health technology assessment on behalf of the National Health Service in England and Wales.[20] Publications from 2000 to May 2014 were included in a systematic review. Participants were men suspected of having prostate cancer for whom the results of an initial prostate biopsy were negative or equivocal; and a PCA3 score or phi in combination with existing standard tests, multiparametric MRI and clinical judgement were evaluated for analytic validity and clinical validity. Overall, six studies met inclusion for the analytical validity review; and fifteen studies met inclusion for the clinical validity review. The authors found issues regarding the precision of the PCA3 assay measurements, and insufficient evidence to identify useful clinical thresholds.

In 2013, the Agency for Healthcare Quality and Research (AHRQ) published a comparative effectiveness review entitled, “PCA3 Testing for the Diagnosis and Management of Prostate Cancer.”[21] Literature was searched and updated through May 15, 2012. Forty-three studies were included; all were rated poor quality. In their conclusion, the authors stated, “For diagnostic accuracy, there was a low strength of evidence that PCA3 had better diagnostic accuracy for positive biopsy results than [serum] total PSA elevations, but insufficient evidence that this led to improved intermediate or long-term health outcomes.” This finding appeared to apply to both initial and repeat biopsies. Evidence was insufficient to assess the use of PCA3 in treatment decision-making for men with positive biopsy.

In a 2012 systematic review, authors discuss the potential use of genetic markers to better define groups of men at high risk of developing prostate cancer, to improve screening techniques, discriminate indolent versus aggressive disease, and improve therapeutic strategies in patients with advanced disease.[22] Genetic tests for PCA3 and TMPRSS2-ERG
genes were included. Authors concluded that most markers have not been prospectively validated for providing useful prognostic or predictive information or improvement upon clinicopathologic parameters already in use.

A meta-analysis by Ruiz-Aragon reviewed 14 studies of PCA3 for use in predicting prostate biopsy results.[23] Sensitivity of testing ranged from 47% to 82% and specificity from 56% to 89%. Global results provided a sensitivity of 85% (confidence interval [CI]: 84 to 87) and a specificity of 96% (CI: 96 to 97). No publications on how this information affected decision making or either short- or long-term outcomes has been published.

**Nonrandomized Studies**

Several studies published between 2013 and 2016 reported positive associations between PCA3 levels and prostate cancer diagnosis.[24-32] Predictive value was increased when PCA3 testing was combined with PSA level and other clinical information.[32-34] Other groups reported moderate diagnostic accuracy of PCA3 testing. Among men with PSA level greater than 3 ng/mL, AUC of PCA3 was 0.74.[35] Conversely, in men with PCA3 scores of 100 or greater, positive predictive value was 39%.[36] In a multicenter study of 647 men, sensitivity and specificity were 67% and 72%, respectively; AUC was 0.742.[37] Three studies compared PCA3 to multiparametric MRI; MRI was more accurate than PCA3,[38,39] but the combination was better than either alone.[40]

In 2014, the National Cancer Institute conducted a prospective trial to validate the diagnostic use of PCA3 to complement PSA-based detection of prostate cancer.[41] The target population included men who had been screened for prostate cancer, primarily with a PSA test, some of whom had undergone a previous prostate biopsy. The study included 859 men from 11 centers in the United States. The primary study endpoint was the diagnosis of prostate cancer on biopsy and the secondary study endpoint was diagnosis of high grade prostate cancer, defined as a Gleason score greater than 6. The primary analyses, including PCA3 thresholds, were determined a priori, and statistical power was based on independent analyses of prevalidation data from similar cohorts. Of the men in the study, 562 were presenting for their initial prostate biopsy. Positive predictive value was 80% (95% CI, 72% to 86%), and using a PCA3 score of more than 60, diagnostic sensitivity and specificity of PCA3 was 0.42 (95% CI, 0.36 to 0.48) and 0.91 (95% CI, 0.87 to 0.94), respectively. For patients who underwent a repeat biopsy, the NPV was 88% (95% CI, 81% to 93%), and by using a PCA3 score of less than 20, sensitivity and specificity were 0.76 (95% CI, 0.64 to 0.86) and 0.52 (95% CI, 0.45 to 0.58), respectively. For the detection of high grade cancer, PCA3 performance in combination with Prostate Cancer Prevention Trial’s (PCPT) risk calculator was improved by the addition of PCA3 to the PCPT risk calculator factors with an AUC improvement of 0.74 to 0.78 for initial biopsy and 0.74 to 0.79 on repeat biopsy (p≤.003).

Clinical utility studies using assay results for decision-making for initial biopsy, repeat biopsy, or treatment have not been reported. One group reported potential reductions in unnecessary biopsies of 48-52% with attendant increases in missed prostate cancers of 6-15% using either a PCA3-based nomogram[42] or PCA3 level corrected for prostate volume (PCA3 density).[43] Although both studies were prospective, neither assessed utility of the test for clinical decision-making because all patients underwent biopsy, and recurrence or survival outcomes were not evaluated. Another group evaluated the estimated long-term impact of using the PCA3 score to guide the decision to recommend a repeat biopsy for men with elevated PSA levels.[44] Their models suggested that using a PCA3 score threshold of 25 would result in 55.4% reduction in
repeat biopsies for a base-case patient, while reducing the 10-year survival by 0.93%. However these, estimates have not been validated in real patient populations.

Section Summary

Studies of PCA3 as a diagnostic test for prostate cancer reported sensitivities and specificities in the moderate range. In general, these studies are preliminary and report on clinical performance characteristics in different populations and with various assay cutoff values, reflecting the lack of standardization in performance and interpretation of PCA3 results. Several studies have reported a modest incremental improvement in diagnostic accuracy when PCA3 was tested in combination with PSA level and other clinical findings. The clinical utility of this test is uncertain, as there is insufficient evidence that the use of PCA3 can be used to change management in ways that improve outcomes.

TMPRSS FUSION GENES

TMPRSS2 fusion gene detection has been studied for prognostic value (e.g., to identify aggressive disease or to predict disease recurrence). In prostate cancer, it may be fused to an ETS family transcription factor (ERG, ETV1, ETV4, or ETV5), which modulates transcription of target genes involved in cell growth, transformation, and apoptosis (TMPRSS2:ERG).

No studies of clinical utility have been published to date; the evidence consists of correlational studies (association between a fusion gene and prostate cancer).[45-48] However, the results of available studies differ as to the accuracy of TMPRSS2:ERG in improving the ability to predict prostate cancer, and/or the ability to estimate prognosis for this purpose.

Systematic Review

In 2013, Yao published a systematic review with meta-analysis of TMPRSS2:ERG for the detection of prostate cancer.[49] Literature was searched and 32 articles were identified. Pooled sensitivity, specificity, positive likelihood ratio, and negative likelihood ratio were 47% (95% CI, 46 to 49), 93% (95% CI, 92 to 94), 8.9 (95% CI, 5.7 to 14.1), and 0.49 (95% CI, 0.43 to 0.55), respectively. Statistical heterogeneity was high (I2>85%). It was unclear whether studies in screening populations were pooled with enriched patient samples, eg, elevated PSA and/or biopsy-negative. There also was variability in the type of tissue samples analyzed (urine, prostatic secretions, biopsy or surgical specimens); the type of TMPRSS2:ERG assays used (fluorescence in situ hybridization [FISH], immunohistochemistry [IHC], real-time reverse transcriptase polymerase chain reaction [RT-PCR], and transcription-mediated amplification); and in TMPRSS2:ERG threshold cutoff values.

Nonrandomized Studies

Leyton investigated the predictive value of PCA3 and TMPRSS2 as individual biomarkers and as part of a panel in a prospective, multicenter study of 443 men.[50] TMPRSS2 was found to be highly specific (93%) for predicting clinically significant prostate cancer on biopsy. Because of this high specificity, the authors suggested that re-biopsy or magnetic resonance imaging (MRI) be performed in TMPRSS2:ERG-positive patients who do not have prostate cancer detected on initial biopsy. The authors stated that if PCA3 in combination with TMPRSS2 data had been used to select men for prostate biopsy, 35% of biopsies could have been avoided. However, the clinical utility of this test is uncertain, as there are no studies that report the test leads to changes in management that result in improved health outcomes.
Whelan (2014) compared 2 multivariate models to assess up-staging in 216 patients meeting National Comprehensive Cancer Network (NCCN) criteria for active surveillance.\(^{[51]}\) (65) One model included \textit{TMPRSS2:ERG} plus serum PSA; the other model included serum PSA, total RNA in expressed prostatic secretion (EPS, collected by milking the urethra after prostatic massage), and total EPS volume. AUCs were similar (0.80 [95% CI, 0.75 to 0.85] and 0.79 [95% CI, 0.73 to 0.84], respectively). However, the second model was more accurate for detecting patients who were up-staged, or up-staged and up-graded, by NCCN criteria. Specifically, the second model decreased the risk of up-staging in patients with a negative test approximately 8-fold (from 7% to 1%); decreased the risk of up-staging plus up-grading approximately 5-fold (from 5% to 1%); and doubled the prevalence of up-staging in the positive test group. In comparison, the \textit{TMPRSS2:ERG} model decreased up-staging 2.4-fold (from 7% to 3%) and decreased upstaging and upgrading approximately 3-fold (from 5% to 2%).

A modeling study by Merdan estimated that using a \textit{TMPRSS2:ERG} score to guide repeat biopsy decisions in men with elevated PSA could avoid 64.7% of these biopsies, but also reduce the 10-year survival rate by 1.4%.\(^{[44]}\) These estimations have not been validated in real-world trials.

A number of studies have reported positive associations between \textit{TMPRSS2} fusion gene levels and prostate cancer diagnosis.\(^{[28,29,52,53]}\) One study reported a lack of association between \textit{TMPRSS2:ERG} status and biochemical relapse-free rate in 244 men treated with image-guided radiotherapy (IGRT) for prostate cancer.\(^{[54]}\) The authors concluded that “\textit{TMPRSS2-ERG} is therefore unlikely to be a predictive factor for IGRT response.”

**Section Summary**

Limited evidence reports that the measurement of \textit{TMPRSS2:ERG} may improve the ability to predict prostate cancer, and/or the ability to estimate prognosis. However, the results of available studies differ as to the accuracy of \textit{TMPRSS2:ERG} for this purpose. In addition, the clinical utility of this test is uncertain.

**\textit{TMPRSS2:ERG} IN COMBINATION WITH PCA3**

The evidence of increased accuracy following simultaneous detection of \textit{TMPRSS2:ERG} and PCA3 is limited. Three studies that address \textit{TMPRSS2:ERG} and PCA3 are described below.

**Nonrandomized Studies**

Tomlins developed a transcription-mediated amplification assay to measure \textit{TMPRSS2:ERG} fusion transcripts in parallel with PCA3.\(^{[48]}\) Combining results from these 2 tests and incorporating them into the multivariate Prostate Cancer Prevention Trial risk calculator improved the identification of patients with clinically significant cancer by Epstein criteria and high-grade cancer on biopsy. Though the study was large (1312 men at multiple centers), results were confounded by assay modifications during the course of the study, by the use of cross-validation rather than independent validation, and the use of independent training and testing sets. A validation study by the same group evaluated this risk prediction model (termed Mi-Prostate Score or MiPS), which incorporated serum PSA with urine \textit{TMPRSS2:ERG}, and PCA3, in a group of 1244 men presenting for biopsy.\(^{[55]}\) They reported that it improved on PSA alone for predicting prostate cancer and high-grade prostate cancer, but did not assess the clinical utility of this risk score.
A study by Feibus evaluated the clinical use of PCA3 and TMPRSS2:ERG in African-American men undergoing prostate biopsy.[31] This study included 182 African-American and 139 nonAfrican-American patients. They found that PCA3 and TMPRSS2:ERG scores were associated with prostate cancer, and adding PCA3 to a standard of care plus PSA model improved concordance statistics for the detection of any prostate cancer in both groups. However, PCA3 score was only predictive of high grade prostate cancer in African-Americans and not in nonAfrican-Americans, while TMPRSS2:ERG did not improve these measures in either group.

In a pilot study, Salami evaluated 45 men using a multivariable algorithm that included serum PSA plus urine TMPRSS2:ERG and PCA3 from a post-DRE sample.[56] Samples were collected before prostate biopsy at 2 centers. For cancer prediction, sensitivity and specificity were 80% and 90%, respectively. AUC was 0.88. Limitations in this study included the small number of patients and authors used a quantitative RT-PCR assay to measure PCA3 in urine sediment, and evaluation of this model with the commercially available TMA assay of whole urine would help inform broader clinical applicability. Authors concluded a larger validation study could determine whether a multiplex model combining PCA3, TMPRSS2:ERG and serum PSA can predict aggressive versus indolent prostate cancer than any of these biomarkers alone.

Robert retrospectively examined tissue levels of TMPRSS2:ERG and PCA3 in 48 men with benign prostatic hypertrophy, 32 men with normal prostate tissue sampled next to prostate cancer, and 48 men with prostate cancer.[57] Sensitivity, specificity, and positive and negative predictive values for the tests in combination were 94%, 98%, 96%, and 96%, respectively. The limitations of the study included that not all samples were confirmed with a diagnosis, there was a lack of urinary data, and the study results are not generalizable to clinical practice.

Section Summary

Concomitant detection of TMPRSS2:ERG and PCA3 may accurately identify men with prostate cancer. However, estimated accuracy varies across the available studies and clinical utility has not been established.

GENE HYPERMETHYLATION FOR PROSTATE CANCER DIAGNOSIS AND PROGNOSIS

Epigenetic changes, chromatin protein modifications that do not involve changes to the underlying DNA sequence but which can result in changes in gene expression (particularly expression of genes associated with prostate cancer), have been identified in specific genes. An extensive literature reports significant associations between epigenetic DNA modifications and prostate cancer. Studies of genetic associations aim to test whether single-locus alleles or genotype frequencies differ between two groups of individuals (usually diseased subjects and healthy controls). Association studies cannot test causality. The identified studies are primarily small, retrospective pilot evaluations of hypermethylation status of various candidate genes for discriminating prostate cancer from benign conditions (diagnosis) or for predicting disease recurrence and association with clinicopathologic predictors of aggressive disease (prognosis). Research gaps included nonstandardized assays, interpretation criteria, and sample types for measuring potential biomarkers. Consistency and comparison of results across studies is therefore lacking. No published studies were identified that demonstrated the clinical utility of testing.

Several studies reported associations between DNA hypermethylation at various gene loci (RASSF1A, APC, GSTP1, PTGS2, RAR-beta, TIG1, AOX1, C1orf114, GAS6, HAPLN3, KLF8,
and MOB3B) and prostate cancer. In contrast, several studies have not found evidence of an association. Further, Kachakova concluded that HIST1H4K hypermethylation was more likely due to aging than to prostate carcinogenesis. Nevertheless, no standardized assays and interpretation criteria have been established to enable consistency and comparison of results across studies.

Section Summary

Studies reporting the diagnostic accuracy and predictive ability of gene hypermethylation report differing results regarding the accuracy of hypermethylation. These inconsistent results make it difficult to determine whether hypermethylation is a useful parameter for diagnosis and/or prognosis of prostate cancer. Further research is needed to elucidate the clinical validity of this test.

MITOCHONDRIAL DNA MUTATION TESTING

Mutations in the mitochondrial genome (mtDNA) are emerging as tools for the diagnosis of prostate cancer. A growing body of literature is reporting significant associations between both single nucleotide changes and large scale deletions in mtDNA and prostate cancer, however, the identified studies are of small to medium sample size and do not address clinical utility. A laboratory developed assay, offered by Mitomics (formerly Genesis Genomics), called the Prostate Core Mitomics Test™ (PCMT) is a proprietary test which is intended to determine whether a patient has prostate cancer, despite a negative prostate biopsy, by analyzing deletions in mitochondrial DNA by PCR to detect “tumor field effect”. The test is performed on the initial negative prostate biopsy tissue. According to the company website, a negative PCMT result confirms the results of the negative biopsy (ie, the patient doesn’t have prostate cancer) and that the patient can avoid a second biopsy, but that a positive PCMT means that the patient is at high risk of undiagnosed prostate cancer. The company website states that the sensitivity of the test is 85% and has a negative predictive value of 92%.

Nonrandomized Studies

A trial published in 2016 examined the role of the mitochondrial genome in prostate cancer risk in 4,086 prostate cancer cases and 3,698 controls from the Multiethnic Cohort. In this study, 350 mitochondrial SNPs were tested in five racial/ethnic populations: Asian Americans, Africans, Europeans, Latinos, and Native Hawaiians. No significant associations were found. Published literature from Genesis Genomics on the use of mitochondrial DNA mutations in prostate includes several studies. A 2006 study retrospectively analyzed mitochondrial DNA mutations from three tissue types from 24 prostatectomy specimens: prostate cancer, adjacent benign tissue and benign tissue distant to the tumor (defined as tissue from a nondiseased lobe or at least 10 cell diameters from the tumor if in the same lobe). Prostate needle biopsy tissue from 12 individuals referred for biopsy that were histologically benign were used as controls. Results from the prostatectomy tissue analysis showed that 16 of 24 (66.7%) had mutations in all three tissue types, 22 of 24 (91.7%) had mutations in malignant samples, 19 of 24 (79.2%) in adjacent benign samples and 22 of 24 in distant benign glands. Overall, 273 somatic mutations were observed in this sample set. In the control group, seven (58.3%) patients were found to have between one and five alterations, mainly in non-coding regions. The authors concluded that the mutations found in the malignant group versus the control group were significantly different and that mitochondrial DNA mutations are an indicator of malignant transformation in prostate tissue.
In 2008, Maki reported the discovery and characterization of a 3.4-kb mitochondrial genome deletion and its association with prostate cancer\cite{75}. A pilot study analyzed 38 benign biopsy specimens from 22 patients, 41 malignant biopsy specimens from 24 patients and 29 proximal to malignant (PTM) biopsy specimens from 22 patients. All of the patients with malignant biopsies had a subsequent prostatectomy, and the diagnosis of cancer was confirmed. The PTM biopsy samples were negative for cancer and were from the cohort who underwent prostatectomy. A confirmation study used 98 benign biopsy specimens from 22 patients, 75 malignant biopsy specimens from 65 patients, and 123 PTM biopsy specimens from 96 patients. In the confirmation study, patients were required to have at least 2 successive negative biopsies; the first negative biopsy was used for analyses. For both the pilot and confirmation studies, samples for analysis were selected based upon review of pathology reports. The levels of the mutation were measured by quantitative PCR and using PCR cycle threshold data were used to calculate a score for each biopsy sample. In the pilot study, the scores were statistically significant between benign and malignant (p<.0001) and benign and proximal (p<.003) samples. The PTM samples closely resembled the malignant sample, with no statistical significant resolution between their scores (p<.833), to which the authors attributed as a field cancerization phenomenon. Results from the larger confirmation study were similar. Compared with histopathologic examination of the benign and malignant samples, the sensitivity and specificity were 80% and 71%, respectively, and the area under a ROC curve was 0.83 for the deletion. A blinded, external validation study showed a sensitivity and specificity of 83% and 79% and the area under the ROC curve 0.87.

In 2010, Robinson assessed the clinical value of the 3.4-kb deletion described in the Maki study in predicting re-biopsy outcomes\cite{75}. Levels of the deletion were measured by quantitative PCR in prostate biopsies negative for cancer from 101 patients who underwent repeat biopsy within a year and had known outcomes. Of the 101 first biopsies, the diagnosis was normal in 8, atypical and/or had prostatic intraepithelial neoplasia (PIN) in 50 and hyperplasia or inflammation in 43. Using an empirically established cycle threshold cutoff, the lowest cycle threshold as diagnostic of prostate cancer, and the histopathologic diagnosis on second biopsy, the clinical performance of the deletion was calculated. The final data was based on 94 patients, who on second biopsy had 20 malignant and 74 benign diagnoses. The cycle cutoff gave a sensitivity and specificity of 84% and 54%, respectively, with the area under a ROC curve of 0.75. Negative predictive value was 91%.

**Section Summary**

Studies using the PCMT test for the diagnosis of prostate cancer reported sensitivities and specificities in the moderate range. In general, these studies are preliminary and only report on clinical validity. There is a lack of standardization in methodology and the clinical utility of this test was not addressed.

**GENE EXPRESSION ANALYSIS**

**Oncotype Dx® Prostate**

**Analytic Validity**

In the only study of validity for the Oncotype DX® Prostate test, Knezevic, authors from Genomic Health, Inc., the developer of the test, reported analytical accuracy and reproducibility of Oncotype Dx® Prostate\cite{76}. Estimates of analytic precision and reproducibility were derived from analysis of RNA prepared from 10 microdissected prostate tumor samples.
obtained by needle biopsy. Individual Gleason scores were assigned using the 2005 International Society of Urological Pathology Consensus guidelines.[77]

The results showed that the assay could accurately measure expression of the 12 cancer-related and 5 reference genes over a range of absolute RNA inputs (0.005-320 ng); the limit of detection in a sample was 0.5 ng/uL. The analytic accuracy showed average variation of less than 9.7% across all samples at RNA inputs typical of needle biopsy specimens. The amplification efficiency for the 17 genes in the test ranged from 88% to 100%, with a median of 93% (SD=6%) for all 17 genes in the assay. Analytic precision was assessed by examining variability between replicate results obtained using the same mRNA input. Reproducibility was measured by calculating both within and between mRNA input variation. A low input level of 5 ng mRNA was used to reflect the lowest 2.5 percentile of a tumor sample of 0.023 cm3. When converted to GPS units (unit measure for reporting test results), the standard deviation for analytic precision was 1.86 GPS units (95% confidence interval [CI], 1.60 to 2.20) on the 100-unit scale. The standard deviation for reproducibility was 2.11 GPS units (95% CI, 1.83 to 2.50) on the 100-point scale. This study provides sufficient evidence to establish the analytic validity of Oncotype Dx® Prostate.

**Clinical Validity**

In 2016, Brand combined the Klein (2014) and Cullen (2015) studies using a patient-specific meta-analysis.[78] The GPS was compared to the CAPRA score, NCCN risk group, and AUA/EAU risk group. The authors tested whether the GPS added predictive value for the likelihood of favorable pathology above the clinical risk assessment tools. The model including the GPS and CAPRA score provided the best risk discrimination; the AUC improved from 0.68 to 0.73 by adding the GPS to CAPRA score. The AUC improved from 0.64 to 0.70 by adding the GPS to the NCCN risk group. The improvements were reported to be significant but the confidence intervals for AUC were not provided.

Whalen (2016) prospectively evaluated the correlation of GPS with final pathology at RP in a clinical practice setting. Eligible men were 50 years of age and older with more than 10 years of life expectancy, PSA levels of 20 ng/mL or less, stage cT1c-cT2c newly diagnosed, untreated prostate cancer, and who met NCCN classifications as very low risk, low risk, or low-intermediate risk.[79] Men were enrolled from May 2013 to August 2014 at an academic medical center. After initial review at the institution, Genomic Health further reviewed biopsy samples to assign Gleason score and tumor length. Samples with Gleason grade discrepancy between initial and central review were excluded from analyses. Clinicians were blinded to GPS when counseling patients on management with active surveillance versus definitive treatment. Genomic Health reclassified patients’ cancers as “less favorable,” “consistent with,” or “more favorable” than what would have been predicted by their NCCN risk group. The primary outcome was adverse pathology at RP defined as any pT3 stage and primary Gleason grade of 4 or any pattern 5. Fifty patients had RP pathology and the reclassification results for these participants are discussed here; 21 (42%) met the definition of adverse pathology. The NCCN risk classification categorized 2 (4%) patients as very low risk, 34 (68%) as low risk, and 14 (28%) as low-intermediate risk. Twenty-three (46%) of patients were reclassified using GPS and the percentage with adverse pathology for the reclassification is shown in Table 9 as derived from data provided in the text. Confidence intervals were not provided.

One publication compiled results of three cohorts, two of which were used for test development, of contemporary (1997-2011) patients in a prostatectomy study (N=441;
Cleveland Clinic database, 1987-2004), a biopsy study (N=167; Cleveland Clinic database, 1998-2007), and an independent clinical validation study cohort (N=395; mean age, 58 years; University of California, San Francisco Urologic Oncology Data Base, 1998-2011).[80]

Results from the clinical validation study and prostatectomy study provide information on the potential clinical validity of this test. The cohorts had a mix of low to low-intermediate clinical risk characteristics using National Comprehensive Cancer Network (NCCN) or American Urological Association (AUA) criteria. Patients included in the validation and prostatectomy studies would be considered (a) eligible for active surveillance based on clinical and pathologic findings and (b) representative of patients in contemporary clinical practice. However, all patients elected radical prostatectomy within 6 months of their initial diagnostic biopsies.

The clinical validation study was designed to evaluate the ability of Oncotype Dx® Prostate to predict tumor pathology in needle biopsy specimens. It was prospectively designed, used masked review of prostatectomy pathology results, and as such met the REMARK (Reporting Recommendations for Tumor Marker Prognostic Studies) guidelines for biomarker validation.[81] In the prostatectomy study, all patients with clinical recurrence (local recurrence or distant metastasis) were selected, together with a random sample of those who did not recur, using a stratified cohort sampling method to construct a 1:3 ratio of recurrent to nonrecurrent patients. The prespecified primary end point of the validation study was the ability of the Genomic Prostate Score (GPS) to predict the likelihood of favorable pathology in the needle biopsy specimen. Favorable pathology was defined as freedom from high-grade or non-organ-confined disease. In the prostatectomy study, the ability of the GPS to further stratify patients within AUA groupings was related to clinical recurrence-free interval in regression-to-the-mean estimated survival curves.

The validation study results showed that the GPS could refine stratification of patients within specific NCCN criteria groupings, as summarized in Table 3. The proportions in Table 1 were estimated from a plot of GPS versus the percent likelihood of favorable pathology.[82] These findings suggest that a lower GPS would reclassify the likelihood of favorable pathology (i.e., less biologically aggressive disease) upward (i.e., a potentially lower risk of progression), and vice versa within each clinical stratum. For example, among patients in the cohort classified by NCCN criteria as low risk, the mean likelihood of favorable pathology in a tumor biopsy was about 76%, with 24% then having unfavorable pathology. With the GPS, the estimated likelihood of favorable tumor pathology was broadened, ranging from 55% to 86%, conversely reflecting a 45% to 14% likelihood of adverse pathology, respectively.

<table>
<thead>
<tr>
<th>NCCN Risk Level</th>
<th>Estimated Mean Likelihood of Favorable Tumor Pathology</th>
<th>Estimated Corresponding GPS, Range</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>NCCN Criteria, %</td>
<td>GPS + NCCN, % Range</td>
</tr>
<tr>
<td>Very low</td>
<td>≈84</td>
<td>63-91</td>
</tr>
<tr>
<td>Low</td>
<td>≈76</td>
<td>55-86</td>
</tr>
<tr>
<td>Intermediate</td>
<td>≈56</td>
<td>29-75</td>
</tr>
</tbody>
</table>

GPS: Genomic Prostate Score; NCCN: National Comprehensive Cancer Network.

In effect, the risk of adverse tumor pathology indicated by the GPS could be nearly halved (24%-14%) at 1 extreme, or nearly doubled (24%-45%) at the other, but the actual number of
patients correctly or incorrectly reclassified between all 3 categories cannot be ascertained from the data provided. The results suggested that the combination of GPS plus clinical criteria could reclassify patients on an individual basis within established clinical risk categories. However, whether these findings support a conclusion that the GPS could predict the biological aggressiveness of a tumor—hence its propensity to progress—based solely on the level of pathology in a biopsy specimen is unclear. Moreover, extrapolation of this evidence to a true active surveillance population, for which the majority in the study would be otherwise eligible, is difficult because all patients had elective radical prostatectomy within 6 months of diagnostic biopsy.

The prostatectomy study, although used to identify genes to include in the GPS, provided estimates of clinical recurrence rates stratified by AUA criteria (Epstein), compared with rates after further stratification according to the GPS from the validation study. The survival curves for clinical recurrence reached a duration of nearly 18 years based on the dates individuals in the cohort were entered into the database (1987-2004). The restratifications are summarized in Table 2. The GPS groups are defined by tertiles defined in the overall study.

Table 2. Reclassification of Prostate Cancer 10-Year Clinical Recurrence Risk With Oncotype Dx® Prostate

<table>
<thead>
<tr>
<th>Overall 10-Year Risk, % (AUA Risk Level)</th>
<th>10-Year Risk, % (GPS Low Group)</th>
<th>10-Year Risk, % (GPS Intermediate Group)</th>
<th>10-Year Risk, % (GPS High Group)</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.4 Low</td>
<td>2.0</td>
<td>3.4</td>
<td>7.0</td>
</tr>
<tr>
<td>9.6 (Intermediate)</td>
<td>2.8</td>
<td>5.1</td>
<td>14.3</td>
</tr>
<tr>
<td>18.2 (high)</td>
<td>6.2</td>
<td>9.2</td>
<td>28.6</td>
</tr>
</tbody>
</table>

AUA: American Urological Association; GPS: Genomic Prostate Score.

In the NCCN intermediate group, for example, the 10-year recurrence rate among radical prostatectomy patients was 9.6%. When the GPS was used in the analysis, the 10-year recurrence rate fell to as low as 2.0% (71% reduction) among patients in the low GPS group and 5.1% (47% reduction) in the intermediate GPS group, but rose to 14.3% (49% increase) in the high GPS group. These data suggest the GPS can reclassify a patient’s risk of recurrence based on a specimen obtained at biopsy. However, the findings do not necessarily reflect a clinical scenario of predicting disease progression in untreated patients under active surveillance.

In summary, the evidence from Klein on clinical validity for Oncotype Dx® Prostate suggests the GPS can reclassify a patient’s risk of recurrence based on a specimen obtained at biopsy. However, whether these findings support a conclusion that the GPS could predict the biological aggressiveness of a tumor—hence its propensity to progress—based solely on the level of pathology in a biopsy specimen is unclear.

Clinical Utility

Klein also reported a decision-curve analysis that they have proposed reflects the clinical utility of Oncotype Dx® Prostate. The analysis investigated the predictive impact of the GPS in combination with the Cancer of the Prostate Risk Assessment (CAPRA) validated tool versus the CAPRA score alone on the net benefit for the outcomes of patients with high-grade disease (Gleason >4+3), high-stage disease, and combined high-grade and high-stage disease. They reported that, over a range of threshold probabilities for implementing treatment, “incorporation of the GPS would be expected to lead to fewer treatments of patients who have favorable pathology at prostatectomy without increasing the number of patients with adverse
pathology left untreated." For example, at a threshold risk of 40% (eg, a man weighing the harms of prostatectomy versus benefit over active surveillance at 4:6) the test could identify 2 per 100 with highgrade or high-stage disease at a fixed false positive rate compared with using the CAPRA score alone. However, no confidence intervals were presented for the decision curve analysis. Thus, an individual patient could use the findings to assess his balance of benefits and harms (net benefit) when weighing the choice to proceed immediately to curative radical prostatectomy with its attendant adverse sequelae, or deciding to enter an active surveillance program. The latter would have an immediate benefit realized by forgoing radical prostatectomy, but perhaps would be associated with greater downstream risks of disease progression and subsequent therapies.

Klein’s decision-curve analyses suggest a potential ability of the combined GPS and CAPRA data to help patients make decisions based on relative risks associated with immediate treatment or deferred treatment (i.e., active surveillance). This would reflect the clinical utility of the test. However, it is difficult to ascribe possible clinical utility of Oncotype Dx® Prostate in active surveillance because all patients regardless of clinical criteria elected radical prostatectomy within 6 months of diagnostic biopsy. Moreover, the validity of using different degrees of tumor pathology as “markers” to extrapolate the risk of progression of a tumor in vivo is unclear.

Polaris®

Analytic Validity

Although there is no reference standard for gene expression profiling tests, other measures of technical performance are relevant and include reproducibility, tissue-sample adequacy, potential batch effects, and test-set bias. In 2015 Warf[84] evaluated the precision of the Cell Cycle Progression (CCP) score using 6 formalin-fixed, paraffin-embedded (FFPE) biopsy (three replicate scores) and 12 FFPE RP (4-6 replicate scores) specimens. Overall precision was estimated from replicate samples, intended to reflect combined variation from tissue dissection through gene expression. Across replicate samples, the standard deviation of the CCP score was 0.1 (95% confidence interval [CI], 0.98 to 0.13). After eight weeks of sample storage, results were similar. In 2013, Myriad Genetics reported 95.3% of samples were adequate to produce a CCP score.[85]

Information is available on the performance of the TaqMan array platform (Applied Biosystems) used in Prolaris® and Oncotype DX® Prostate from the MicroArray Quality Control (MAQC) project, thus providing an indirect evaluation of the analytical validity of Prolaris®.[86] In the MAQC project, initiated and led by FDA scientists, expression data on 4 titration pools from 2 distinct reference RNA samples were generated at multiple test sites on 7 microarray-based and 3 alternative technology platforms including TaqMan. According to the investigators, the results provided a framework to assess the potential of array technologies as a tool to provide reliable gene expression data for clinical and regulatory purposes. The results showed very similar performance across platforms, with a median coefficient of variation of 5% to 15% for the quantitative signal and 80% to 95% concordance for the qualitative detection call between sample replicates.

Clinical Validity: Needle Biopsy, Conservative Management

In 2015, Cuzick examined three U.K. cancer registries from 1990 to 2003 to identify men with prostate cancer who were conservatively managed following needle biopsy, with follow-up
through December 2012. Men were excluded if they had undergone RP or radiation therapy within 6 months of diagnosis. A combination of the CCP and Cancer of the Prostate Risk Assessment (CAPRA) scores was used to predict prostate cancer death. There were 989 men who fit eligibility criteria; CCP scores were calculable for 761 (77%) and combined CCP and clinical variables were available for 585 (59%). Median age at diagnosis was 70.8 years and median follow-up was 9.5 years. The prostate cancer mortality rate was 17% (n=100), with 29% (n=168) dying from competing causes. Higher CCP scores were associated with increased 10-year risk of prostate cancer mortality: 7% (CCP score <0), 15% (CCP score 0-1), 36% (CCP score 1-2), 59% (CCP score >2). A 1-unit increase in CCP was associated with a crude hazard ratio (HR) for death of 2.08 (95% CI, 1.76 to 2.46) and when adjusted for CAPRA score yielded a HR of 1.76 (95% CI, 1.47 to 2.14). For the combined CAPRA/CCP score, the HR for 10-year prostate cancer mortality increased to 2.17 (95% CI: 1.83 to 2.57). The c-statistic for the CAPRA score was 0.74; adding the CCP score increased the C statistic to 0.78 (no confidence intervals for the AUC were reported). Treatment changes after 6 months were documented in only part of 1 of the 3 cohorts; at 24 months, 45% of the men in this cohort had undergone radiotherapy or prostatectomy. Therefore, the potential effect of treatment changes on prognostic estimates is uncertain.

The original peer-reviewed evidence on the clinical validity of Prolaris® comprises a retrospective cohort (n=349) culled from 6 cancer registries in Great Britain. The investigators assert the CCP score alone was more prognostic than either PSA titer or Gleason score for tumor-specific mortality at 10-year follow-up. Although the patients may be similar to those of a modern U.S. cohort, comparability is unclear from the single publication that is available. Furthermore, the study is limited by the use of archived biopsy specimens, with attendant issues of reproducibility and test reliability.

Cuzick reported this original validation study of Prolaris® to determine its prognostic value for prostate cancer death in a conservatively managed needle biopsy cohort. The authors did not state whether this study adhered to the PRoBE (prospective-specimen-collection, retrospective-blinded evaluation) criteria suggested by Pepe for an adequate biomarker validation study. They noted that the cell cycle expression data were read blind to all other data, which conformed to the criteria; however, patients were identified retrospectively from tumor registries and there were no case-control subjects, which does not conform to the criteria.

Patients with clinically localized prostate cancer diagnosed by needle biopsy between 1990 through 1996 were identified in 6 registries. Additional inclusion criteria included age younger than 76 years at diagnosis, available baseline prostate-specific antigen (PSA) measurement, and conservative management. Exclusion criteria included radical prostatectomy, death, evidence of metastatic disease within six months of diagnosis, or hormone therapy prior to diagnostic biopsy. A cell cycle progression (CCP) score consisting of expression levels of 31 predefined cell cycle progression genes and 15 housekeeper genes was generated using TaqMan low-density arrays. The values of each of the 31 CCP genes were normalized by subtraction of the average of up to 15 nonfailed housekeeper genes for that replicate.

Of 776 patients diagnosed by needle biopsy, 349 (79%) produced a CCP score and had complete baseline and follow-up information. The median potential follow-up time was 11.8 years during which a total of 90 deaths from prostate cancer occurred within 2799 person-years of actual follow-up. The main assessment of the study was a univariate analysis of the association between death from prostate cancer and the CCP score. A further predefined
assess assessment of the added prognostic information after adjustment for the baseline variables was also undertaken. The primary end point was time to death from prostate cancer. A number of covariates were evaluated: centrally reviewed Gleason primary grade and score; baseline PSA value; clinical stage; extent of disease (percent of positive cores); age at diagnosis; Ki-67 immunohistochemistry; and initial treatment. The results are shown in Table 3.

Table 3. Univariate and Multivariate Analysis for Death From Prostate Cancer in the Cuzick 2012 Validation Study

<table>
<thead>
<tr>
<th>Variable</th>
<th>N</th>
<th>Univariate Hazard Ratio (95% CI)</th>
<th>Multivariate Hazard Ratio (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-unit increase in CCP score</td>
<td>349</td>
<td>2.02 (1.62 to 2.53)</td>
<td>1.65 (1.31 to 2.09)</td>
</tr>
<tr>
<td><strong>Gleason score</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;7</td>
<td>106</td>
<td>0.46 (0.25 to 0.86)</td>
<td>0.61 (0.32 to 1.16)</td>
</tr>
<tr>
<td>7</td>
<td>152</td>
<td>Referent</td>
<td>Referent</td>
</tr>
<tr>
<td>&gt;7</td>
<td>91</td>
<td>2.70 (1.72 to 4.23)</td>
<td>1.90 (1.18 to 3.07)</td>
</tr>
<tr>
<td><strong>log (1+PSA)/(ng/mL)</strong></td>
<td>349</td>
<td>1.70 (1.31 to 2.20)</td>
<td>1.37 (1.05 to 1.79)</td>
</tr>
<tr>
<td><strong>Proportion of positive cores</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>&lt;50%</td>
<td>69</td>
<td>0.50 (0.22 to 1.12)</td>
<td></td>
</tr>
<tr>
<td>50 to &lt;100%</td>
<td>106</td>
<td>Referent</td>
<td></td>
</tr>
<tr>
<td>100%</td>
<td>160</td>
<td>1.66 (1.01 to 2.73)</td>
<td></td>
</tr>
<tr>
<td><strong>Age at diagnosis (y)</strong></td>
<td>349</td>
<td>1.00 (0.96 to 1.04)</td>
<td></td>
</tr>
<tr>
<td><strong>Clinical stage</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T1</td>
<td>38</td>
<td>0.75 (0.32 to 1.75)</td>
<td></td>
</tr>
<tr>
<td>T2</td>
<td>106</td>
<td>Referent</td>
<td></td>
</tr>
<tr>
<td>T3</td>
<td>43</td>
<td>1.74 (0.90 to 3.38)</td>
<td></td>
</tr>
<tr>
<td><strong>Hormone use</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>No</td>
<td>200</td>
<td>Referent</td>
<td></td>
</tr>
<tr>
<td>Yes</td>
<td>149</td>
<td>1.97 (1.30 to 2.98)</td>
<td></td>
</tr>
</tbody>
</table>

CI: confidence interval

The median CCP score was 1.03 (IQ range, 0.41-1.74). The primary univariate analysis suggested that a 1-unit increase in CCP score was associated with a 2-fold increase in the risk of dying from prostate cancer. In preplanned multivariate analyses, extent of disease, age, clinical stage, and use of hormones had no statistically significant effect on risk; only the Gleason score and PSA remained in the final model. Further exploratory multivariate modeling to produce a combined score, including CCP, Gleason score, and PSA level, suggested a strong, predominant nonlinear influence of the CCP score in predicting the risk of death from prostate cancer (p=0.008).

Cuzick suggested this combined score provided additional discriminatory information to help identify low-risk patients who could be safely managed by active surveillance. For example, among patients with a Gleason score of 6, for whom uncertainty existed as to the appropriate management approach, the predicted 10-year prostate cancer death rate ranged from 5.1% to 20.9% based on Gleason score and PSA; the range when assessed against the combined
CCP, Gleason, and PSA score was 3.5% to 41%. However, the authors cautioned that because death rates were rare in this group, larger cohorts would be required to fully assess the value of the CCP combined score. Measures that would suggest improved discriminatory ability, eg., AUC or reclassification, were not reported. Evidence was not provided that the test could correctly reclassify men initially at high risk to lower risk to avoid overtreatment, or conversely those initially at low risk to high risk to avoid undertreatment.

Table 4 shows Kaplan-Meier analyses of 10-year risk of prostate cancer death stratified by CCP score groupings. Cuzick reported no significance tests for the estimates. Nor did they explain the apparent substantial difference in mortality rates among patients in the $0 \leq \text{CCP} \leq 2$ grouping (range, 19.3-21.1%) and those in the $2 < \text{CCP} \leq 3$ and $> 3$ groupings (range, 48.2-74.9%). The difference may simply reflect clinical criteria, for example, proportions of lower compared with higher Gleason grade cancers, respectively.

Table 4. Kaplan-Meier Estimates of Prostate Cancer Death at 10 Years According to CCP Score Groupings in the Cuzick 2012 Validation Study

<table>
<thead>
<tr>
<th>CCP Score Group</th>
<th>N</th>
<th>10-Year Death Rate (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\text{CCP} \leq 0$</td>
<td>36</td>
<td>19.3</td>
</tr>
<tr>
<td>$0 &lt; \text{CCP} \leq 1$</td>
<td>133</td>
<td>19.8</td>
</tr>
<tr>
<td>$1 &lt; \text{CCP} \leq 2$</td>
<td>114</td>
<td>21.1</td>
</tr>
<tr>
<td>$2 &lt; \text{CCP} \leq 3$</td>
<td>50</td>
<td>48.2</td>
</tr>
<tr>
<td>$&gt; 3$</td>
<td>16</td>
<td>74.9</td>
</tr>
</tbody>
</table>

Clinical Validity: Posttreatment (radical prostatectomy and external beam radiation therapy)

In 2014 Bishoff[90] examined the prognostic ability of the CCP score in 3 cohorts: Martini Clinic (n=283, simulated biopsies from FFPE RP specimen), Durham Veterans Affairs Medical Center (n=176, diagnostic biopsies) and Intermountain Healthcare (n=123, diagnostic biopsies). The combined analysis included all 582 patients. Gleason scores were 7 or lower in 93% of men. In the combined cohorts, a unit increase in the CCP score increased the adjusted HR for BCR by 1.47 (95% CI, 1.23 to 1.76). Metastatic events (n=12) were too few to draw conclusions. Although the CCP score was associated with increased risk of BCR, the analyses do not allow examining whether the CCP score provides improved discrimination over clinicopathologic variables.

In 2013 Myriad-funded study, Freedland evaluated the CCP score’s ability to predict biochemical recurrence (BCR) in a cohort of men treated with external beam radiation therapy (EBRT).[91] The CCP score was derived retrospectively from diagnostic biopsy specimens of men diagnosed with prostate cancer from 1991 to 2006 (n=141). The primary outcome assessed was time from EBRT to BCR. In a multivariable analysis with Gleason score, PSA, percent positive cores, and androgen deprivation therapy, the hazard ratio (HR) was 2.11 for a one-unit increase in CCP score (equivalent to a doubling of gene expression) ($p=0.034$), indicating that CCP provides prognostic information that is not provided by standard clinical parameters. At ten years post-EBRT, the CCP score was associated with prostate cancer specific mortality ($p$-value $= 0.013$). The limitations of this study include small size of the cohort, small number of treatment failures (only 19 patients [13%] had BCR), and short follow-
up time. The authors conceded that “definitive conclusions regarding time dependency will require additional studies”.

In 2013 Cooperberg[80] sought to evaluate the CCP score in a RP cohort and the incremental improvement over the Cancer of the Prostate Risk Assessment Postsurgical (CAPRA-S) score for predicting BCR employing a prospective-retrospective design (conforming to a PRoBE study design). A prognostic model was developed from the RP cohort described by Cuzick (2011).[92] The validation cohort was obtained from patients identified from the University of California, San Francisco (UCSF) Urologic Oncology Database. Tissue sufficient to obtain a CCP score was available for 413 men (69% of the 600 eligible samples). Both UCSF and Myriad Genetics performed statistical analyses. In the validation cohort, 95% had Gleason scores of 7 or lower, 16% of samples had positive margins, 4% had seminal vesicle invasion, and 23% had extracapsular extension. BCR occurred in 82 men (19.9%). The unadjusted HR for BCR increased by 2.1 (95% CI, 1.6 to 2.9) per unit increase in CCP score. A predictive model for the combined CCP/CAPRA-S developed in the Cuzick (2011)[92] RP cohort applied to the UCSF cohort obtained an AUC for BCR with CAPRA-S alone of 0.73 increasing to 0.77 for the combined CCP/CAPRA-S.

Cuzick (2011)[92] examined the potential use of the Prolaris® CCP test combined with a clinical score following RP, using a retrospective cohort and the prospective-retrospective design for archived samples. The study also included a cohort of men with localized prostate cancer detected from specimens obtained during transurethral resection of the prostate, which is not a population of interest here, and so has not been described. Men conservatively managed post RP between 1985 and 1995 were identified from a tumor registry (n=366 with CCP scores, Scott and White Clinic, in Texas). The primary end point was time to biochemical recurrence (BCR) and the secondary end point was prostate cancer death. Myriad Genetics assessed CCP scores blindly. The median age of patients was 68 years and the median follow-up 9.4 years. Gleason scores were 7 or lower in 96%, but margins were positive in 68%. Cancers were clinically staged as T3 in 34%; following RP, 64% was judged pathologic stage T3. CCP score was associated with BCR (adjusted HR=1.77; 95% CI, 1.40 to 2.22). Analyses of prostate cancer deaths in the RP cohort were problematic, owing to only 12 (3%) deaths. The clinical score included PSA, stage, positive surgical margins, and Gleason score. The model was optimized using stepwise variable selection (eg, a development model). The AUC for BCR within 5 years in the RP cohort was 0.825 for the clinical score and 0.842 for the combined clinical/CCP score. The discriminatory ability of the clinical score is of note. Although the CCP increased the AUC by 2%, whether that improvement might be clinically useful is unclear lacking reclassification or examination of net benefit.

Clinical Utility

One large prospective registry study, funded by Myriad, was recently published that evaluated the impact of the CCP test on treatment decision making for patients newly diagnosed with prostate cancer.[93] Patients (n=1206) with newly diagnosed prostate adenocarcinoma had the CCP test performed on initial prostate biopsy tissue. Changes in treatment decision making was tracked using the answers provided by physicians in sequential surveys relative to initial therapy recommendations (before cell cycle progression). The CCP test caused a change in actual treatment in 47.8% of patients, 72.1% of which were reductions and 26.9% of which were increases in treatment. For each clinical risk category there was a significant change in treatment modality (intervention vs nonintervention) before vs after CCP testing (p=0.0002). This study did not report any changes in patient-important outcomes, such as biochemical
recurrence, cancer-specific survival or long term survival. Although this study reported a change of management in a modest percentage of patients, there was no evidence that these changes in management lead to clinically important improvements in health outcomes.

Two retrospective survey studies that assessed the potential impact of Prolaris® on physicians’ treatment decisions.[94,95] The authors of each study have suggested their findings support the “clinical utility” of the test, based on whether the results would lead to a change in treatment. Although this information may be useful in assessing the potential test uptake by urologists, it does not demonstrate clinical utility in clinical settings. In a decision-curve analysis, Cooperberg[80] found the CAPRA-S score superior to CCP alone (as well as treat-none or treat-all strategies) in men postprostatectomy. A combined CCP/CAPRA-S predictor appeared only slightly better than CAPRA-S alone for thresholds of approximately 30% or more. For example, at a threshold of 30% (ie, meaning a man would value the harm-to-benefit of treatment such as radiotherapy as 3:7), the combined CCP/CAPRA-S would detect about two more men per 100 likely to experience BCR if the false-positive rate was fixed. However, the lack of confidence intervals for the decision-curve analysis, together with the small difference, is consistent with an uncertain net benefit obtained by adding CCP to the CAPRA-S score.

Systematic Review

In 2016, results of a systematic review and meta-analysis supported by the manufacturer were reported.[96] Published and unpublished studies of prognostic validity or clinical utility of CCP testing were eligible for inclusion. Seven published studies were identified; 5 were clinical validity studies. Two were reviewed in the previous paragraphs and the remaining validity studies will be reviewed in a subsequent section on post-RP management. The other 2 “utility” studies are discussed in the following section. Two validity studies reported outcomes for disease-specific mortality[87,92] but of the 2 only the Cuzick (2012) included newly diagnosed patients, so the pooled outcome is not of relevance in this section.

Decipher®

Analytic Validity

Published data on the analytic validity of the Decipher test consists of one study, which was performed on surgical resection specimens from patients with prostate cancer identified to be in a postsurgery high-risk population. The Decipher test platform was performed in formalin-fixed, paraffin-embedded (FFPE) tissue to assess the differential expression in the discovery, validation and clinical application.[97] Matched FFPE and unfixed fresh-frozen specimens from paired tumor and normal samples from kidney, lung and colon were compared and the microarray signals derived from the degraded RNA extracted from FFPE specimens was found to be highly analogous to the signals from the RNA in the fresh frozen specimens.

Clinical Validity

The clinical validity of the Decipher test genomic classifier (GC) has been reported in studies to predict metastasis, mortality or BCR after RP in patients with postoperative high-risk features like pathologic stage T2 with positive margins, pathologic stage T3 disease or a rising PSA.[98-106] All studies were conducted from registry data. The development study was a nested case-control design,[104] four were case-cohort studies, and four used retrospective cohorts. Owing to apparent overlap in samples, the number of unique patients in the studies is difficult to
ascertain. Seven studies were supported by GenomeDx, which offers the Decipher test; all studies identified multiple authors as company employees.

Four studies,[101-104] including the test (validation) sample from the development study, examined men observed following radical prostatectomy and undergoing adjuvant or salvage radiotherapy. Median follow-up periods ranged from 6.4 to 16.9 years. The distributions of Gleason scores in the studies varied—from 24.3% to 49.3% with 8 or higher and 0.4% to 15.1% with 6 or lower. Extracapsular extension of the tumor ranged from 42.7% and 72.3% of men across the studies.

In 2016, Klein evaluated the ability of the Decipher genomic classifier in predicting metastasis from the prostate needle biopsy diagnostic tumor tissue from 56 men.[106] Median follow-up time was eight years. In that time, eight patients metastasized and three died of PCa. Decipher plus NCCN model had an improved c-index of 0.88 (95% confidence interval [CI] 0.77-0.96) compared to NCCN alone (c-index 0.75; 95% CI 0.64-0.87). Using the Cox multivariable analysis, Decipher was the only significant predictor of metastasis when adjusting for age, preoperative PSA and biopsy Gleason score, with a hazard ratio of 1.72 per 10% increase; 95% CI 1.07-2.81; p= 0.02).

In 2015, Ross[107] assessed the prognostic accuracy for metastasis through 10 years, excluding men receiving any adjuvant therapy following radical prostatectomy over median follow-up periods of 7.8 and nine years. The investigators reported a 6.5% 5-year cumulative incidence of metastases in men with GC scores of 0.45 or lower, compared with 30.3% in those with scores higher than 0.60. The AUCs for development of metastases was 0.76 for the GC. In addition, it was found that combining the GC with the best clinicopathologic tool improve the AUC. The study did not include a “standard” reclassification table, but did report 10-year cumulative incidence of metastases stratified by GC and CAPRA-S. The GC appeared to discriminate within CAPRA-S categories, but appeared to add little to a score greater than 5.

In 2015, Den reported on the use of the Decipher genomic classifier (GC) to provide prognostic and predictive information into the development of metastases in men receiving post-RP RT (either 3-dimensional conformal or IMRT).[99] Genomic classifier scores were calculated from 188 men who were identified within the GenomeDx prostate cancer database with pathologic stage T3 or margin-positive prostate cancer and had received post-RP RT at 1 of 2 academic centers between 1990 and 2009. The primary endpoint was metastatic disease (regional or distant) documented on computed tomography or bone scan. Adjuvant versus salvage RT was defined by PSA levels of 0.2 ng/mL or less and more than 0.2 ng/mL before initiation of RT. The clinical characteristics of eligible patients included 72% of men with extraprostatic extension, 35% with seminal vesicle invasion, and 78% with positive surgical margins. Twenty-one percent of patients had a Gleason score of 8 or more. Fifty-one percent of patients received adjuvant RT (89% within 12 months of RP) and overall, patients received RT at a median of 5 months after RP (range, 1-160 months). Thirty percent of patients received hormonal therapy with RT. Median follow-up after RP and RT was 10 and 8 years, respectively. Cumulative incidence of metastatic disease at 5 years after RT for low, average and high GC scores was 0%, 9% and 29% (p=0.002). In a multivariate analysis, GC and pre-RP PSA were independent predictors of metastasis (both p<0.01). In the low GC score group (score <0.4) there was no difference in cumulative incidence of metastasis compared with patients who received adjuvant or salvage RT (p=0.79), however, for patients with higher GC scores (≥0.4), the cumulative incidence of metastasis at 5 years was 6% for patients treated with adjuvant RT compared to 23% treated with salvage RT (p<0.01). The authors concluded
that patients with low GC scores are best treated with salvage RT and those with high GC scores with adjuvant RT.

In 2014 Klein evaluated whether use of the Decipher genomic classifier (GC) test improved accuracy in predicting metastasis within 5 years following radical prostatectomy (rapid metastasis [RM]).[98] Participants included 169 patients who underwent radical prostatectomy between 1987 and 2008, of which 15 were RM and 154 were non-RM controls. Metastasis developed between 1.7 and 3.3 years (median 2.3 years). Test performance was evaluated both individually and in combination with clinical risk factors. After adjusting for clinical factors, Decipher was a significant predictor of RM (OR 1.48; p=0.018). Compared to the Stephenson model, the CAPRA-S, and previously reported biomarkers, Decipher had the highest concordance index (c-index), with the highest c-index achieved with integration of Decipher into the Stephenson nomogram.

Karnes prospectively created a randomly selected subcohort from the same initial 1010 post-prostatectomy patients in the Cooperberg study.[103] Patients with metastasis at diagnosis or with any prior treatment for prostate cancer were excluded. A randomly selected subcohort was created, with genomic data was available for 219 patients. Following radical prostatectomy the rates of biochemical recurrence (BCR) at 3 years was 35% and metastasis at 5 years was 6%. Median genomic classifier scores were consistently higher in patients with metastases at last follow-up (mean 6.7 years). Median genomic classifier scores also increased with higher Gleason scores. The authors concluded that the higher net benefit of genomic-based classifiers suggested increased specificity (i.e., lower false positives) compared with clinical-only risk models. Because patients with intermediate risk tumors may progress to advanced disease, the authors recommended further study of genomic classifiers in randomized datasets to determine whether genomic classifier scores from diagnostic biopsy specimens can predict metastasis as well as postoperative specimens. A possible limitation of this study was that nearly 15% of patients were node-positive and 45% received adjuvant therapy. Whether the genomic classifier predicted benefit from local (i.e., radiation) or systemic (e.g., hormone) therapies could not be determined because patients were not randomized to these treatments.

In 2014, Den reported that in within a Decipher low-risk group that was treated post-RP with RT, there was no difference in oncologic outcomes (either biochemical failure or metastasis) whether they received adjuvant or salvage RT.[100] For the men classified as high-risk by Decipher, a median 4-year PSA-free survival advantage was observed in the patients that received adjuvant versus salvage RT. Of these men classified as high-risk by GC, those who received adjuvant radiation had a 3% cumulative incidence of metastases as compared with 23% incidence of metastasis by 8 years in those who delayed treatment and received salvage radiation.

Clinical Utility

Several studies have compared physician’s treatment recommendations before and after receiving GC results.[108-111] Because the studies did not include information on outcomes and clinical validity has not been established, it is not known whether these treatment decisions represent a clinical improvement in management.

Ross (2016) reported results of a retrospective, comparative study of RT after RP for 422 men with pT3 disease or positive margins.[112] The men were from 4 cohorts previously described (Karnes 2013; Den 2014; Ross 2016; Freedland 2016). The 4 treatment groups were adjuvant
RT (n=111), minimal residual disease salvage RT (n=70), salvage RT (n=83), and no RT (n=157). The primary end point was metastasis. Thirty-seven men developed metastasis and the median follow-up was 8 years. Both CAPRA-S (HR=1.39; 95% CI, 1.18 to 1.62) and Decipher (HR=1.28; 95% CI, 1.08 to 1.52) were independently associated with metastasis in multivariable analysis. There was no evidence that treatment effect was dependent on genomic risk (interaction p=0.16 for CAPRA-S, p=0.39 for Decipher). Men with low CAPRA-S or low Decipher scores had a low risk of metastatic events regardless of treatment selection and men with high CAPRA-S or Decipher scores benefitted from adjuvant RT compared to the other treatments.

Lobo (2015) reported an individualized decision analysis comparing the GC to “usual care” using data from the cohorts in Karnes (2013) and Den (2014). The usual care probabilities of receiving each treatment were derived from the published literature. A 6% threshold for the GC score was used for GC-based treatment. Using the cohort from Karnes (2013), the estimated 10-year probability of metastasis or death was 0.32 (95% CI, 0.32 to 0.33) for usual care compared to 0.31 (95% CI, 0.30 to 0.32) for GC-based treatment. In the cohort from Den (2014), the estimated 10-year probability of metastasis or death was 0.28 (95% CI, 0.27 to 0.29) for usual care compared to 0.26 (95% CI, 0.25 to 0.27) for GC-based treatment.

In a retrospective review of 110 patients with pT3 disease (71%) or positive surgical margins (PSMs) (63%) following radical prostatectomy. US board certified urologists (n=107) were invited to provide adjuvant treatment recommendations for 10 cases randomly drawn from the pool of case histories. These recommendations were based on clinical variable only; GC test results were not provided. After these recommendations were completed, the urologists were asked to make recommendations with GC test results provided for the same patients, with case histories randomly re-ordered to minimize recall bias. Of the 107 urologists invited to participate, 51 took part in the study, providing 530 recommendations without GC test results. The overall change in recommendations when GC test results were added to clinical information was 31%). Of the 303 patients recommended for observation based on clinical information only, 38 (13%; 95% CI 9-17%) were recommended for radiation therapy and 9 (3%; 95% CI 1-6%) for radiation plus hormone therapies when GC results were made available. Of 193 patients recommended for radiation therapy initially, when GC results were available 77 (40%; 95% CI: 33-47%) were changed to observation. Therapy intensity was also highly correlated with higher risk according to the GC test. This study had a number of limitations which the authors noted as typical for early-stage evaluations of biomarker technologies. These limitations included the recommendations being based on case histories from chart reviews, the 51 participating urologists may not be representative of all urologists treating prostate cancer, patient health status and pathological information was not available to the urologists, and the effect on health outcomes of any changes in treatment were not studied.

In 2014, Michalopoulos assessed the effect of the GC test on urologists’ decisions regarding treatment of men with high-risk disease post-RP. Participating urologists were from 15 community practices who had ordered the GC test for 146 prostate cancer patients with either pathologic stage T3 or positive surgical margins post-RP. The urologists were asked to provide their treatment recommendations before and after receiving the GC test report. Prior to availability of the GC test result treatment recommendations were based on Gleason score and CAPRA-S risk: 40 (27.4%) were recommended to undergo adjuvant therapy, 102 (69.9%) close observation and 4 (2.7%) “other.” Using the GC risk score, 61.6% and 38.4% of patients were identified as low- and high-risk of metastasis, respectively. More than 60% of high-risk
patients were reclassified as low risk after the GC test results. Overall, adjuvant treatment recommendations were modified for 30.8% (95% CI, 23% to 39%) of patients. With the GC test results, 42.5% of patients who were initially recommended for adjuvant therapy were subsequently recommended observation. The GC test score also influenced the intensity of the treatment recommendation, with about 40% of patients classified as high-risk by GC score recommended more intense therapy versus 1.1% of those deemed low-risk by GC score. Limitations to the study included that treatment recommendations were submitted electronically and did not track the actual treatment administered, it was not possible to assess patient influence on the decision-making process, the association between GC test results and treatment recommendations was determined using “early adopters” of the test, and all participants were community-based physicians whose treatment recommendations may differ from those of academic centers.

PRACTICE GUIDELINE SUMMARY

AMERICAN UROLOGICAL ASSOCIATION (AUA)

In 2013, the AUA published guidelines for the early detection of prostate cancer.[114] Based on a systematic review of the literature, the guideline panel recognized that novel urinary markers, such as PCA3 and TMPRSS2:ERG, may be “used as adjuncts for informing decisions about the need for a prostate biopsy – or repeat biopsy – after PSA screening,” but emphasized the lack of evidence “that these tests will increase the ratio of benefit to harm.”

EVALUATION OF GENOMIC APPLICATIONS IN PRACTICE AND PREVENTION (EGAPP)

In 2013, the EGAPP Working Group published the following recommendations for PCA3 testing in prostate cancer, based on an AHRQ comparative effectiveness systematic review summarized above:[115]

- Evidence was insufficient to recommend PCA3 testing to inform decisions for when to re-biopsy previously biopsy-negative patients for prostate cancer, or to inform decisions to conduct initial biopsies for prostate cancer in at-risk men (e.g., previous elevated PSA or suspicious digital rectal examination).
- Evidence was insufficient to recommend PCA3 testing in men with cancer-positive biopsies to determine if the disease is indolent or aggressive in order to develop an optimal treatment plan.
- The overall certainty of clinical validity to predict the diagnosis of prostate cancer using PCA3 is deemed “low.” Clinical use for diagnosis is discouraged unless further evidence supports improved clinical validity.
- The overall certainty of net health benefit is deemed “low.” Clinical use is discouraged unless further evidence supports improved clinical outcomes.

NATIONAL COMPREHENSIVE CANCER NETWORK

The National Comprehensive Cancer Network (NCCN) guidelines for Prostate Cancer Early Detection (2.2017) suggest considering tests that improve specificity in higher risk patients after negative biopsy, including 4Kscore, percent free PSA, PHI, PCA3, and ConfirmMDx.[116] PCA3 is not recommended for use in the initial biopsy setting. Guideline authors note:

“Biomarkers that improve the specificity of detection are not recommended as frontline screening tests. However, there may be some patients who meet PSA standards for
consideration of prostate biopsy, but for whom the patient and/or the physician wish to further define the probability of high-grade cancer. A percent free PSA <10%, PHI >35 or 4Kscore (which provides an estimate of the probability of high-grade prostate cancer) are potentially informative in patients who have never undergone biopsy or after a negative biopsy; a PCA3 score >35 is potentially informative after a negative biopsy."

The NCCN guidelines for prostate cancer (1.2018) do not include gene expression profile analysis in their recommendations, though tests are listed in the discussion on tissue-based molecular assays. The discussion section notes that the clinical utility has not been established in prospective RCTs or comparative effectiveness studies.[117]

U.S. PREVENTIVE SERVICES TASK FORCE RECOMMENDATIONS

The U.S. Preventive Services Task Force published recommendations for Prostate Cancer Screening on May 2012. Genetic tests addressed in this policy, including PCA3, were not mentioned.

SUMMARY

There is not enough research to recommend using gene-based tests for prostate cancer screening, detection and management, as many important characteristics of these tests have not yet been determined. Some research shows that they might help predict the diagnosis or prognosis of prostate cancer, but it is not yet known how much information they add to currently available tests. More research is needed to demonstrate how these tests can improve outcomes for patients. Therefore, use of gene-based testing for screening, detection, and management of prostate cancer is considered investigational.

REFERENCES


### CODES

<table>
<thead>
<tr>
<th>Codes</th>
<th>Number</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPT</td>
<td>0005U</td>
<td>Oncology (prostate) gene expression profile by real-time RT-PCR of 3 genes (ERG, PCA3, and SPDEF), urine, algorithm reported as risk score</td>
</tr>
<tr>
<td></td>
<td>0011M</td>
<td>Oncology, prostate cancer, mRNA expression assay of 12 genes (10 content and 2 housekeeping), RT-PCR test utilizing blood plasma and/or urine, algorithms to predict high-grade prostate cancer risk</td>
</tr>
<tr>
<td></td>
<td>0012M</td>
<td>Oncology (urothelial), mRNA, expression profiling by real-time quantitative PCR of five genes (MDK, HOXA13, CDC2 [CDK1], IGFBP5, and XCR2), utilizing urine, algorithm reported as a risk score for having urothelial carcinoma</td>
</tr>
<tr>
<td></td>
<td>0013M</td>
<td>Oncology (urothelial), mRNA, gene expression profiling by real-time quantitative PCR of five genes (MDK, HOXA13, CDC2 [CDK1], IGFBP5, and CXCR2), utilizing urine, algorithm reported as a risk score for having recurrent urothelial carcinoma</td>
</tr>
<tr>
<td></td>
<td>81313</td>
<td>PCA3/KLK3 (prostate cancer antigen 3 [non-protein coding]/kallikrein-related peptidase 3 [prostate specific antigen]) ratio (eg, prostate cancer)</td>
</tr>
<tr>
<td></td>
<td>81479</td>
<td>Unlisted molecular pathology procedure</td>
</tr>
<tr>
<td></td>
<td>81541</td>
<td>Oncology (prostate), mRNA gene expression profiling by real-time RT-PCR of 46 genes (31 content and 15 housekeeping), utilizing formalin-fixed paraffin-embedded tissue, algorithm reported as a disease-specific mortality risk score</td>
</tr>
<tr>
<td></td>
<td>81551</td>
<td>Oncology (prostate), promoter methylation profiling by real-time PCR of 3 genes (GSTP1, APC, RASSF1), utilizing formalin-fixed paraffin-embedded tissue, algorithm reported as a likelihood of prostate cancer detection on repeat biopsy</td>
</tr>
<tr>
<td></td>
<td>81599</td>
<td>Unlisted multianalyte assay with algorithmic analysis</td>
</tr>
<tr>
<td>HCPCS</td>
<td>None</td>
<td></td>
</tr>
</tbody>
</table>

*Date of Origin: October 2012*