Genetic and Protein Biomarkers for
the Diagnosis and Cancer Risk
Assessment of Prostate Cancer

Policy Number: 2.04.33          Last Review: 7/2017
Origination: 7/2014          Next Review: 7/2018

Policy
Blue Cross and Blue Shield of Kansas City (Blue KC) will not provide coverage for Gene-Based Tests for Screening, Detection, and/or Management of Prostate Cancer. This is considered investigational.

Note: This is a type of genetic testing that may be excluded in some contracts. Verify benefits prior to review for Medical Necessity.

When Policy Topic is covered
Not Applicable

When Policy Topic is not covered
The following genetic and protein biomarkers for the diagnosis of prostate cancer are considered investigational:

- Kallikrein markers (eg, 4Kscore™ Test)
- Metabolomic profiles (eg, Prostarix™)
- PCA3 testing
- TMPRSS fusion genes
- Candidate gene panels
- Mitochondrial DNA mutation testing (eg, Prostate Core Mitomics Test™)
- Gene hypermethylation testing (eg, ConfirmMDx®)
- Prostate Health Index (phi)

Single nucleotide polymorphisms (SNPs) testing for cancer risk assessment of prostate cancer is considered investigational.

Description of Procedure or Service

<table>
<thead>
<tr>
<th>Populations</th>
<th>Interventions</th>
<th>Comparators</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Various genetic and protein biomarkers are associated with prostate cancer. These tests have the potential to improve the accuracy of differentiating which men should undergo prostate biopsy or rebiopsy after a prior negative biopsy. This evidence review will address these types of tests for cancer risk assessment. Testing to determine cancer aggressiveness after a tissue diagnosis of cancer is addressed in a separate policy.

For individuals for whom an initial prostate biopsy or a rebiopsy is being considered, the evidence includes systematic reviews and meta-analyses and primarily observational studies. Relevant outcomes are overall survival, disease-specific survival, test accuracy and validity, other test performance measures, resource utilization, hospitalizations, quality of life, and treatment-related mortality and morbidity. The evidence supporting clinical utility varies by test but has not been directly shown for any biomarker test. In general, performance of biomarker testing for predicting biopsy compared with clinical examination, including the ratio of free or unbound prostate-specific antigen (PSA) to total PSA, is lacking. However, procedures for referrals for biopsy based on clinical examination vary, making it difficult to quantify performance characteristics for this comparator. There is considerable variability in biopsy referral practices based on clinical examination alone and many biomarker tests do not have standardized cutoffs to recommend biopsy. Therefore, having prospective, comparative information on how test results are expected to be used or actually being used in practice and the associated effects on outcomes will be needed to determine if the tests improve net health outcomes. Many of the test validation populations have included men with a positive digital rectal exam, prostate-specific antigen (PSA) level outside of the gray zone, or older men for whom the information for test results are less likely to be informative. African Americans have a high burden of morbidity and mortality, but have not been well represented in these study populations. It is unclear how to monitor men with low biomarker risk scores who continue to have symptoms or high or rising PSA levels. Comparative studies of the many biomarkers are lacking and it is unclear how to use the tests in practice, particularly when test results are contradictory. The evidence is insufficient to determine the effects of the technology on health outcomes.
Background

Disease description and epidemiology

Prostate cancer is the second most common cancer in men, with a predicted 181,000 incidence cases and 26,100 deaths expected in United States in 2016.¹

Prostate cancer is a complex, heterogeneous disease, ranging from microscopic tumors unlikely to be life-threatening to aggressive tumors that can metastasize, leading to morbidity or death. Early localized disease can usually be cured with surgery and radiotherapy, although active surveillance may be adopted in men whose cancer is unlikely to cause major health problems during their lifespan or for whom the treatment might be dangerous. In patients with inoperable or metastatic disease, treatment consists of hormonal therapy and possibly chemotherapy. The lifetime risk of being diagnosed with prostate cancer for men in the United States is approximately 16%, but the risk of dying of prostate cancer is 3%.² African-American men have the highest prostate cancer risk in the United States; the incidence of prostate cancer is about 60% higher and the mortality rate is more than 2 to 3 times greater than that of white men.³ Autopsy results have suggested that about 30% of men age 55 and 60% of men age 80 who die of other causes have incidental prostate cancer,⁴ indicating that many cases of cancer are unlikely to pose a threat during a man’s life expectancy.

The most widely used grading scheme for prostate cancer is the Gleason system.⁵ It is an architectural grading system ranging from 1 (well differentiated) to 5 (poorly differentiated); the score is the sum of the primary and secondary patterns. A Gleason score of 2 to 5 is regarded as normal prostate tissue; 6 is low-grade prostate cancer that usually grows slowly; 7 is an intermediate grade; 8 to 10 is high-grade cancer that grows more quickly. Ten-year survival rates stratified by Gleason score have been estimated from the Surveillance, Epidemiology, and End Results registry to be about 98% for scores 2 through 6, 92% for a score of 7 with primary pattern 3 and secondary pattern 4 (3+4), 77% for a score of 7 (4+3), and 70% for scores between 8 and 10.⁶

Numerous genetic alterations associated with development or progression of prostate cancer have been described, with the potential for use of these molecular markers to improve selection of men who should undergo prostate biopsy or rebiopsy after an initial negative biopsy.

Clinical context and TEST purpose

The purpose of genetic and protein biomarker testing for prostate cancer is to inform the selection of men who should undergo biopsy or repeat biopsy. Conventional decision-making tools for identifying men for prostate biopsy include digital rectal exam (DRE), serum prostate-specific antigen (PSA), and patient risk factors such as age, race, and family history of prostate cancer.

DRE has relatively low interrater agreement among urologists, with estimated sensitivity, specificity, and positive predictive value (PPV) for diagnosis of prostate cancer of 59%, 94% and 28%, respectively.⁷ DRE might have a higher PPV in the setting of elevated PSA.⁸
The risk of prostate cancer increases with increasing PSA levels; an estimated 15% of men with a PSA level of 4 ng/mL or less and normal DRE, 30% to 35% of men with a PSA level between 4 and 10 ng/mL, and more than 67% of men with a PSA level greater than 10 ng/mL will have biopsy-detectable prostate cancer. Use of PSA levels in screening has improved detection of prostate cancer. The European Randomized Study of Screening for Prostate Cancer (ERSPC) and Goteborg prostate cancer screening trials demonstrated that biennial PSA screening reduces the risk of being diagnosed with metastatic prostate cancer.

However, elevated PSA levels are not specific to prostate cancer; levels can be elevated due to infection, inflammation, trauma, or ejaculation. In addition, there are no clear cutoffs for cancer positivity with PSA. Using a common PSA level cutoff of 4.0 ng/mL, the American Cancer Society (ACS) systematically reviewed the literature and calculated pooled estimates of elevated PSA sensitivity of 21% for detecting any prostate cancer and 5% for detecting high-grade cancers with estimated specificity of 91%.

PSA screening in the general population is controversial. The U.S. Preventive Services Task Force recommended against PSA-based screening (D recommendation) in 2012 while guidelines published by ACS and the American Urological Association (AUA) endorsed consideration of PSA screening based on age, other risk factors, and estimated life expectancy. The utility of PSA screening depends on whether screening can lead to management changes that improve net health outcome. Results from the National Institutes for Health–supported Prostate Testing for Cancer and Treatment (ProtecT) trial demonstrated no difference in prostate cancer mortality rates between the treatment strategies of active monitoring, radical prostatectomy, and external-beam radiotherapy in clinically localized prostate cancer that is detected by PSA testing.

These existing screening tools lead to unnecessary prostate biopsies because of their lack of specificity and inability to discriminate low- from high-risk prostate cancer. More than 1 million prostate biopsies are performed annually in the United States with a resulting cancer diagnosis in 20% to 30%. About one-third of men who undergo prostate biopsy experience transient pain, fever, bleeding, and urinary difficulties. Serious biopsy risks (eg, bleeding or infection requiring hospitalization) are rare, with estimated rates ranging from less than 1% to 4%.

Given the risk, discomfort, and burden of biopsy and the low yield for diagnosis, there is a need for noninvasive tests that distinguish potentially aggressive tumors that should be referred for biopsy from clinically insignificant localized tumors that do not need biopsy or other prostatic conditions with the goal of avoiding low-yield biopsy. The following PICOTS were used to select literature that provides evidence relevant to this review.
Patients
The relevant populations are men for whom an initial prostate biopsy is being considered because of clinical symptoms (eg, difficulty with urination, elevated PSA) or men for whom a rebiopsy is being considered because the results of an initial prostate biopsy were negative or equivocal and other clinical symptoms remain suspicious.

The population for which these tests would potentially be most informative is men in the indeterminate or “gray zone” range of PSA on repeat testing with unsuspicious DRE findings. Repeat testing of PSA is important because results initially reported to be between 4 and 10 ng/mL are frequently normal.\(^\text{22}\) The gray zone for PSA levels is usually between 3 or 4 and 10 ng/mL, but PSA levels varies with age. Age-adjusted normal PSA ranges have been proposed but not standardized or validated.

Screening of men with a life expectancy of less than 10 years is unlikely to be useful because most prostate cancer progresses slowly. However, the age range for which screening is most useful is controversial. The ERSPC and Goteborg trials observed benefits of screening only in men up to about 70 years old.

Interventions
For assessing future prostate cancer risk, numerous studies have demonstrated the association between many genetic and protein biomarker tests and prostate cancer. Commercially available tests include those described in Table 1.

| Table 1. Commercially Available Tests to Determine Candidate for Prostate Biopsy or Repeat Biopsy |
|-----------------|-----------------|-----------------|
| Test            | Manufacturer   | Description                                              |
| 4Kscore®        | OPKO lab       | Blood test that measures 4 prostate-specific kallikreins, which are combined into an algorithm to produce a score |
| Prostarix™      | Metabolon/Bostwick Laboratories | Urine test that measures several metabolites, which are combined with an algorithm to produce a score |
| Progensa®       | Hologic Gen-Probe | Urine test that measures PCA3 mRNA |
|                 | Many labs offer PCA3 tests (eg, ARUP Laboratories, Mayo Medical Laboratories, and LabCorp) |
| ConfirmMDx®     | MDxHealth      | Measures hypermethylation of 3 genes in tissue sample |
| Prostate Health Index™ (phi) | Beckman Coulter | Blood test that combines several components of PSA with an algorithm to produce a score |
| Prostate Core Mitomics Test™ (PCMT) | Mitomics (formerly Genesis Genomics) | Measures deletions in mitochondrial DNA by polymerase chain reaction in tissue sample |

PSA: prostate specific antigen.
In addition to commercially available tests, single-nucleotide polymorphism testing as part of genome-scanning tests for prostate cancer risk assessment are offered by a variety of laboratories, such as Navigenics (now Life Technologies), LabCorp (23andme), and ARUP Laboratories (deCODE), as laboratory-developed tests.

**Comparators**

Standard clinical examination for determining who goes to biopsy might include DRE, review of history of PSA values, along with consideration of risk factors such as age, race, and family history. The ratio of free or unbound PSA to total PSA (%fPSA) is lower in men who have prostate cancer than in those who do not. A %fPSA cutoff of 25% has been shown to have a sensitivity and specificity of 95% and 20%, respectively, for a group of men with total PSA values between 4.0 and 10.0 ng/mL. 23

The best way to combine all of the risk information to determine who should go to biopsy is not standardized. Risk algorithms have been developed that incorporate clinical risk factors into a risk score or probability. Two examples are the Prostate Cancer Prevention Trial (PCPT) predictive model 24 and the Rotterdam Prostate Cancer risk calculator (also known as the European Research Screening Prostate Cancer Risk Calculator 4 [ERSPC-RC]). 25 The AUA and the Society of Abdominal Radiology recently recommended that high-quality prostate magnetic resonance imaging, if available, should be strongly considered in any patient with a prior negative biopsy who has persistent clinical suspicion for prostate cancer and who is under evaluation for a possible repeat biopsy. 26

**Outcomes**

Outcomes of interest are overall survival, disease-specific survival, test accuracy and validity, other test performance measures, resource utilization, hospitalizations, quality of life, and treatment-related mortality and morbidity.

The beneficial outcome of the test is to avoid a biopsy that would be negative for prostate cancer. A harmful outcome is failure to undergo a biopsy that would be positive for prostate cancer, especially when disease is advanced or aggressive. Thus the relevant measures of clinical validity are the sensitivity and negative predictive value. The appropriate reference standard is biopsy. Prostate biopsies are not perfect for diagnosis. Biopsies can miss cancers and repeat biopsies are sometimes needed to confirm diagnosis. Detection rates vary by method used for biopsy and patient characteristics, with published estimates between 14% and 22% for the initial biopsy, 10% and 28% for a second biopsy, and 5% and 10% for a third biopsy. 27,28

**Time**

The timeframe of interest for calculating performance characteristics is time to biopsy result. Men who forgo biopsy based on test results could miss or delay diagnosis of cancer. Longer follow-up would be necessary to determine effects on overall survival.
Setting
Initial screening using PSA levels and DRE may be performed in primary care with referral to specialty (urologist) care for suspicious findings and biopsy. Clinical practice regarding screening methods and frequency vary widely.

Regulatory Status
Clinical laboratories may develop and validate tests in-house and market them as a laboratory service; laboratory-developed tests (LDTs) must meet the general regulatory standards of the Clinical Laboratory Improvement Amendments (CLIA). Laboratories that offer LDTs must be licensed by CLIA for high-complexity testing. BioReference Laboratories and GenPath Diagnostics (subsidiaries of OPKO Health; 4Kscore®), Metabolon (Prostarix™), ARUP Laboratories, Mayo Medical Laboratories, LabCorp, BioVantra, others (PCA3 assay), Clinical Research Laboratory (Prostate Core Mitomic Test™), MDx Health (ConfirMDx), and Innovative Diagnostics (phi™), are CLIA-certified. To date, the U.S. Food and Drug Administration (FDA) has chosen not to require any regulatory review of this test.

In February 2012, the Progensa® PCA3 Assay (Gen-Probe; now Hologic) was approved by FDA through premarket approval process. According to the company’s press release, this 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 1 or more previous negative prostate biopsies and for whom a repeat biopsy would be recommended by a urologist based on the current standard of care, before consideration of Progensa PCA3 assay results.” FDA product code: OYM.

In June 2012, pro-PSA, a blood test used to calculate the Prostate Health Index (phi; Beckman Coulter) was approved by FDA through the premarket approval process. The phi test is indicated as an aid in distinguishing prostate cancer from benign prostatic condition in men ages 50 and older with prostate-specific antigen levels of 4 to 10 ng/mL and with digital rectal exam findings that are not suspicious. According to the manufacturer, the test reduces the number of prostate biopsies. FDA product code: OYA.

Rationale
This evidence review was created in April 2004 and has been updated periodically with literature search of the MEDLINE database. The most recent literature review covers the period through August 26, 2016. In 2009, this evidence review was extensively updated and broadened in scope, primarily based on a 2008 TEC Special Report on prostate cancer genetics and genetic testing. Genetic and protein biomarker tests are best evaluated within the framework of a diagnostic or prognostic test, because such frameworks provide diagnostic and prognostic information that assists in clinical management decisions. Assessment of a diagnostic or prognostic tool typically focuses on 3 categories of evidence: (1) analytic validity (ability of the test to accurately and reliably measure the marker of interest); (2) clinical validity (ie, statistically significant association between the
test result and health outcomes); and (3) clinical utility (ie, demonstration that use of the diagnostic or prognostic information clinically can improve net health outcome compared with patient management without use of the tool). Because these tests are used as an adjunct to the usual diagnostic workup, it is important to evaluate whether the tests provide incremental information above the standard workup to determine if the tests have utility in clinical practice.

This review evaluates evidence for genetic and protein biomarkers for the purpose of guiding decision making about biopsy or rebiopsy (see Appendix Table 1 for genetic testing categories).

**prostate-specific antigen–Related Biomarkers**

**Kallikreins Biomarkers and 4Kscore Test**
The 4Kscore test (OPKO Lab) uses results from a blood test to generates a risk score estimating the probability of finding high-grade prostate cancer (defined as a Gleason score ≥7) if a prostate biopsy were performed. The intended use of the test is to aid in a decision whether to proceed with a prostate biopsy. A kallikrein is a subgroup of enzymes that cleaves peptide bonds in proteins. The intact prostate-specific antigen (iPSA) and human kallikrein 2 (hK2) tests are immunoassays that employ distinct mouse monoclonal antibodies. The score combines the measurement of 4 prostate-specific kallikreins (total PSA [tPSA], free PSA [fPSA], iPSA, hK2), with an algorithm including patient age, digital rectal exam (DRE; nodules or no nodules), and a prior negative prostate biopsy.

The manufacturer’s website states that the ideal patient for the 4Kscore is one whose other test results are equivocal. The test is not intended for patients with a previous diagnosis of prostate cancer, who have had a DRE in the previous 4 days, who have received 5-alpha reductase inhibitor therapy in the previous 6 months, or who have undergone treatment for symptomatic benign prostatic hypertrophy in the previous 6 months.30

**Analytic Validity**
Measures of analytic validity include sensitivity (detection rate), specificity (1 – false-positive rate); reliability (repeatability of test results), and assay robustness (resistance to small changes in preanalytic or analytic variables). As described above, the 4Kscore combines 4 blood biomarkers. Total and free PSA are measured using Food and Drug Administration (FDA)–approved kits from Roche Diagnostics. Intact PSA and hK2 are proprietary assays validated by OPKO. Only 1 published study was found describing any components of analytic validity for a test of kallikrein biomarkers. In 2006, Vaisanen et al reported on results of a new method to reduce false high results by eliminating assay interference in measurement of intact fPSA and free hK2.31 Using 1092 female heparin plasma samples as controls and 957 male samples, they optimized the protocol for immunoassays by replacing monoclonal capture or tracer antibodies with F(ab’)2 or recombinant Fab fragments. They tested the new method on another set of 444 samples and found that the optimized assay eliminated 70% to 85% of the falsely elevated results. Other measures of analytic validity were not found in the
literature or the 4Kscore website. The laboratories that perform the analyses for 4Kscore are certified under the Clinical Laboratory Improvement Amendments (CLIA).

Clinical Validity
At least 13 retrospective studies\textsuperscript{32-44} and 1 prospective study\textsuperscript{45} have estimated the performance characteristics of a risk score (KLK) derived from 4 kallikrein biomarkers. Many studies appear to be developmental work for the currently marketed version of the test. In general, the comparators used in these studies were other risk calculators or models that included terms for age, total PSA, and occasionally other risk factors. The reference standard was usually biopsy. Some studies performed in Sweden had long-term follow-up from a national registry of prostate cancer. The eligibility criteria included a lower limit of PSA (2 or 3 ng/mL) in most studies with no upper limit, and men with and without positive DRE. The studies included a mix of men who had or did not have previous PSA testing or biopsies. Mathematical methods used to calculate the KLK risk score varied across studies with respect to whether kallikrein values derived from plasma or serum measurements, the additional risk factors included in the model (age, DRE, biopsies, other risk factors), and how kallikrein marker values were entered into the model (linearly, with splines or cubic splines). The area under the receiver operating characteristic (AUC ROC) curve, or a similar metric, was calculated in all studies. The estimated AUC for the KLK model ranged from 0.72 to 0.90 and was numerically higher than the comparator in all studies except Carlsson et al, who compared the KLK model to a clinical model including length of benign tissue.\textsuperscript{36} However, the confidence intervals (CIs) for AUC of the KLK model frequently overlapped with those of the comparator. A few studies provided results for the KLK model calculated with and without each of the 4 kallikreins. In many cases, the addition of iPSA and hK2 did not significantly improve the model. In Bryant et al, the confidence intervals of the AUC for 4 kallikreins model overlapped considerably with a model that included age, total and free PSA for any grade and high-grade cancer.\textsuperscript{44} Nordstrom et al included a comparison to another biomarker test (phi) and found both tests had very similar AUCs.\textsuperscript{34}

Review of the clinical validity of the 4Kscore only includes studies that stated use of the marketed 4Kscore version of the KLK model. The marketed version of the test appears to have been used in 3 studies.\textsuperscript{32,33,45} Cutoffs for categorizing risk into low, medium, or high levels were only given in Konety et al and therefore sensitivity and negative predictive value (NPV) have not generally been calculated.\textsuperscript{33} Results of the studies are summarized in Table 2. Two studies were conducted in the United States\textsuperscript{33,45} and more detail on them is available in the following paragraphs.

<table>
<thead>
<tr>
<th>Study</th>
<th>Population</th>
<th>Reference Standard</th>
<th>Blinded Comparison to Reference</th>
<th>Performance Characteristics (95% CI)</th>
</tr>
</thead>
</table>

Table 2. Clinical Validity Studies of the 4Kscore for Diagnosing High-Grade Prostate Cancer
### Performance of the 4Kscore Test

Performance of the 4Kscore test was validated in 1012 patients enrolled in a blinded, prospective study at 26 urology centers in the United States. Enrollment was open to all men scheduled for a prostate biopsy, regardless of age, PSA level, DRE, or prior prostate biopsy. Each patient underwent a transrectal ultrasound (TRUS)–guided prostate biopsy of at least 10 cores. A blinded blood sample collected before biopsy was sent to OPKO Lab for the 4 kallikrein markers. Results of the kallikrein markers, prostate biopsy histopathology, patient age, DRE, and prior biopsy status were unblinded and analyzed.

Most participants (86%) were white; 85 (8%) African-American men were included. At baseline, 247 (24%) men had an abnormal DRE, 348 (34%) had a
PSA level less than 4 ng/mL, and 104 (10%) had PSA level greater than 10 ng/mL. Approximately 25% of the men were older than 70 years. Biopsies were negative in 54% (n=542) of cases, and showed low-grade (all Gleason grade 6) prostatic cancer in 24% (n=239) and high-grade cancer in 23% (n=231). Statistical analysis of 4Kscore test clinical data had AUC ROC of 0.82 (95% CI, 0.79 to 0.85) for the detection of high-grade prostate cancer; the AUC ROC for the PCPT risk calculator model was 0.74, but a precision estimate was not given.

Section Summary: Clinical Validity
The intended use population is not well defined. In addition, there is uncertainty regarding clinical performance characteristics such as sensitivity, specificity, and predictive value due to lack of standardization of cutoffs to recommend biopsy, study populations including men with low (<4 ng/mL) and high (>10 ng/mL) baseline PSA levels and positive DRE results who are likely outside of the intended use population, and lack of comparison to models using information from standard clinical examination. African Americans have a high burden of morbidity and mortality but were not well represented in the study populations. The evidence needed to draw conclusions on clinical validity is insufficient. Longer term data on incidence of prostate cancer in men who did not have a biopsy following testing with the marketed version of 4Kscore are not available. However, the Stattin et al case-control study that was nested in a cohort study of more than 17,000 Swedish men estimated that, for men ages 60 with PSA levels of 3 or higher and a KLK risk score less than 10%, the risk of metastasis at 20 years was 1.95% (95% CI, 0.64% to 4.66%).

Clinical Utility
No studies reporting direct evidence of utility for clinical outcomes were found. Various cutoffs for the KLK probability score were used in decision curve analyses to estimate the number of biopsies versus cancers missed. Parekh et al estimated that 307 biopsies could have been avoided and 24 cancer diagnoses would have been delayed with a 9% 4Kscore cutoff for biopsy and 591 biopsies would have been avoided with 48 diagnoses delayed with a 15% cutoff. However, inferences on clinical utility cannot be made due to deficiencies in estimating the clinical validity that are described in the previous section.

Konety et al reported results for a survey of 35 U.S. urologists identified through the 4Kscore database at OPKO Lab as belonging to practices that were large users of the test. All 611 patients of participating urologists who were referred for abnormal PSA or DRE and had a 4Kscore test were included. Six percent of the men had an abnormal DRE; the distribution of PSA levels was not reported. Urologists, who received the 4Kscore as a continuous risk percentage, were retrospectively asked about their plans for biopsy before and after receiving the test results and whether the 4Kscore test results influenced their decisions. Scores were grouped into 3 risk categories: less than 7.5%, low risk; 7.5% to 19.9%, intermediate risk; and 20% or more, high risk. The physicians reported that the 4Kscore results influenced decisions in 89% of men and that the test led to a 64.6% reduction in prostate biopsies. The 4Kscore risk categories correlated highly (p<0.001) with biopsy outcomes in 171 men with biopsy results. Calculated
performance characteristics are shown in Table 2. No other risk calculators were included as comparators.

Absent direct evidence of clinical utility, an indirect chain might be constructed. The 4Kscore test is associated with diagnosis of aggressive prostate cancer. The incremental value of the 4Kscore with respect to clinical examination and risk calculators in the intended use population is unknown due to deficiencies in estimating clinical validity described in the previous section. There is no prospective evidence that use of 4Kscore changes management decisions. The indirect chain is incomplete.

**Section Summary: Kallikreins Biomarkers and 4Kscore Test**

Published data on most components of the analytic validity of the 4Kscore test are lacking. At least 13 studies have reported on clinical validity of the kallikreins biomarkers but only 3 clearly used the marketed version of the 4Kscore test. The eligibility criteria for these studies generally had a lower limit for screening PSA but no upper limit. Given that the test manufacturer’s website states that the test is for men with inconclusive results, the inclusion of men with PSA levels greater than 10 ng/mL and positive DRE in the validation studies is likely not reflective of the intended use population. Studies that provide data on the incremental value of the components of the test show only small improvements with the iPSA and hKA components (components specific to the 4Kscore). The 2 studies performed in U.S. men did not provide estimates (with confidence intervals) of validity compared to a standard clinical examination with ratio of free or unbound PSA to total PSA (%fPSA). Very few data are available on longer term clinical outcomes of men who were not biopsied based on 4Kscore results. No direct evidence supports the clinical utility of the test and the indirect chain of evidence is incomplete due to the limitations in estimates of clinical validity and utility.

**pro-PSA and Prostate Health Index**

The Prostate Health Index (phi; Beckman Coulter) is an assay combining results of 3 blood serum immunoassays (tPSA, fPSA, [-2]proPSA [p2PSA]) numerically to produce a “phi score.” This score is calculated in a routine laboratory using Beckman Coulter equipment and software with the phi algorithm incorporated in the software using the following formula: \([-2\text{proPSA}/fPSA] \times \sqrt{tPSA}\). It has been suggested that the PSA isoform [-2]proPSA (or p2PSA) might better distinguish between prostate cancer and benign prostatic conditions.

The phi score has been approved by FDA for distinguishing prostate cancer from benign prostatic conditions in men 50 years and older with above-normal tPSA readings between 4.0 and 10 ng/mL who have had a negative DRE. The manufacturer’s website states that the test gives men “accurate information on what an elevated PSA level might mean and the probability of finding cancer on biopsy” and when “combined with family and patient history, the phi results can be used to determine the best individualized patient management decisions.”

---

46 Genetic and Protein Biomarkers for the Diagnosis and Cancer Risk Assessment of Prostate Cancer 2.04.33
Analytic Validity

The FDA Summary of Safety and Effectiveness Data (SSED) provides data on the analytic validity of the assay. The analytic validity was also reviewed by the National Institute for Health and Care Excellence (NICE). The limit of blank of p2PSA was 0.5 pg/mL, the limit of detection was 0.7 pg/mL, and the limit of quantification was 3.23 pg/mL. Accuracy was calculated by the percentage recovery of measured p2PSA pg/mL in 6 male serum samples containing different known amounts of purified p2PSA. One hundred percent of samples fell within 100%±15%. Mean recovery was 93% (range, 90%-96%). The within- and between-run imprecision at internal and external sites demonstrated acceptable performance. Dilution recovery and linearity were tested in 12 samples. Eleven of 12 had a slope of 1.0±0.15.

Clinical Validity

Systematic Reviews

Several systematic reviews and meta-analyses have described the clinical validity of pro-PSA and phi. The characteristics of the reviews are shown in Table 3. The reviews cover studies reported between 1990 and 2014. All primary studies were observational and most were retrospective. All reviews included studies of men with a positive, negative, or inconclusive DRE; only 2 of the 5 reviews restricted eligibility to studies including PSA levels between 2 and 10 ng/mL. The Wang et al review included only studies that had sufficient information to distinguish aggressive from indolent prostate cancer. The 2 most recent reviews (Pecoraro et al, Nicholson et al) included most of the studies covered in the older reviews; we review them in more detail below.

Pecoraro et al performed a search of MEDLINE, EMBASE, Web of Science, Scopus, and the Cochrane Register of Diagnostic Test Accuracy Studies for studies including men with PSA between 2 and 10 ng/mL. The quality of each study was assessed using Quality Assessment of Diagnostic Accuracy Studies (QUADAS-2) checklist and the evidence was evaluated using the GRADE approach. Random-effects bivariate models were used to calculate pooled estimates.

Nicholson et al performed a systematic review and health technology assessment (HTA) commissioned to support development of NICE guidance for diagnosing prostate cancer with PCA3 and phi. The search included the databases MEDLINE, EMBASE, the Cochrane Library, Web of Science, Medion, Aggressive Research Intelligence Facility database, ClinicalTrials.gov, International Standard Randomised Controlled Trial Number Register, and World Health Organization International Clinical Trials Registry Platform. The review included studies of men whose initial prostate biopsies were negative or equivocal and studies using clinical examination or clinical examination plus magnetic resonance imaging as comparators. Pooled estimates were not reported due to heterogeneity.

Table 3. Characteristics of Systematic Reviews of the Clinical Validity of phi for Diagnosing Prostate Cancer

<table>
<thead>
<tr>
<th>Study</th>
<th>Dates</th>
<th>Key Inclusion Criteria</th>
<th>Design</th>
</tr>
</thead>
</table>
Results of the systematic reviews and meta-analyses are shown in Table 4. Pecoraro et al included 17 studies with 6912 men. They rated most of the primary studies as low quality due to the design (most were retrospective), lack of blinding of outcome assessors to reference standard results, lack of clear cutoffs for diagnosis, and lack of explicit diagnostic question. Pooled estimates had high heterogeneity across studies but with generally low specificity of phi at 90% sensitivity.

Nicholson et al included 4 studies with 767 men that included estimates of clinical assessment alone versus clinical assessment plus phi. They concluded that the implication of adding phi to clinical assessment was unclear. Due to heterogeneity in cutoffs used in the primary studies, it was not possible to identify thresholds to use in a clinical setting and the clinical relevance of many reported outcomes was unclear.

Table 4. Results of Systematic Reviews of the Clinical Validity of phi for Diagnosing Prostate Cancer

<table>
<thead>
<tr>
<th>Study</th>
<th>Studies</th>
<th>N (Range)</th>
<th>Outcome</th>
<th>Results (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pecoraro (2016)</td>
<td>17</td>
<td>6912 (63-1091)</td>
<td>Diagnostic performance, any prostate cancer</td>
<td>Pooled specificity at 90% sensitivity:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>phi=0.31 (0.29 to 0.33)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>tPSA=0.25 (0.23 to 0.27)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>AUC range:</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>phi plus clinical assessment, 0.65-0.81</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Clinical assessment alone, 0.62-0.75</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Derived sensitivity (at given specificity)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>phi plus clinical assessment, 42% (at 80%), 25% (at 90%), 19% (at 95%)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Clinical assessment alone, 48%</td>
</tr>
</tbody>
</table>
Bruzzese (2014)  
8  
3173 (64-896)  
Diagnostic performance, any prostate cancer  
- Significant heterogeneity  
- Sensitivity range: phi, 0.31-0.90  
  %fPSA, 0.12-0.90  
- Specificity range: phi, 0.30-0.90  
  %fPSA, 0.11-0.90  
- AUC: phi, 0.74 (90.70 to 0.77)  
  %fPSA=0.63 (0.58 to 0.67)

Wang (2014)  
12  
Diagnostic performance, high-grade (Gleason score ≥7) prostate cancer  
- Pooled sensitivity, 0.90 (0.87 to 0.92)  
- Pooled specificity, 0.17 (0.14 to 0.19)  
- Diagnostic OR=3.06 (1.61 to 5.84)  
- Pooled AUC=0.67 (0.57 to 0.77)

Filella (2013)  
13  
3928 (63-1091)  
Diagnostic performance, any prostate cancer  
- Significant heterogeneity  
- Pooled specificity at 90% sensitivity, 0.32 (0.29 to 0.34)

Other Relevant Clinical Studies
The pivotal study described in the FDA SSED included men 50 years and older with nonsuspicious DRE and PSA levels between 4 and 10 ng/mL who had a histologically confirmed diagnosis. The null hypothesis was that the phi specificity at 95% sensitivity would be no greater than the specificity of percent free PSA (%fPSA). Seven sites in the United States enrolled 658 men between 2008 and 2009 (97% enrolled prospectively, 3% enrolled retrospectively). Eighty-one percent of participants were white, 5% were African American, and 1% were Asian. At 95% sensitivity, using a phi cutoff of 22.1, the specificity was 14.1% (precision not reported) for phi compared to 9.9% for %fPSA. AUC was 0.71 (95% CI, 0.67 to 0.75) for phi compared to 0.65 (95% CI, 0.61 to 0.69) for %fPSA.

Additional studies have been published since the systematic reviews. In 2015 Fossati et al conducted a case-control study with 1036 European men younger than 60 years of age. They reported that phi had a higher AUC than tPSA in men younger than 60 years (0.70 [95% CI, 0.64 to 0.76] vs 0.55 [95% CI, 0.48 to 0.61]) for detecting any prostate cancer. At 91% sensitivity, phi and tPSA had similar specificity (11.1% [95% CI, 6.8% to 16.8%] vs 10.5% [95% CI, 6.4% to 16.1%]) and NPV (76.0% [95% CI, 59.3% to 92.7%] vs 75.0% [95% CI, 57.7% to 92.3%]). At the best combination of sensitivity and specificity (phi cutoff ≥41.2, tPSA cutoff ≥5.72), phi had a sensitivity of 64.2% (95% CI, 51.5% to 75.5%), a specificity of 63.2% (95% CI, 55.5% to 70.4%), and an NPV of 81.8% (95% CI, 75.2% to 88.4%) while tPSA had a sensitivity of 52.2% (95% CI, 39.7% to 64.6%), a specificity of 52.0% (95% CI, 44.3% to 59.7%) and an NPV of 73.6% (95% CI, 65.7% to 81.4%). A decision curve analysis found that using a model
with age, prostate volume, tPSA, fPSA, %fPSA, and phi with a probability cutoff of 10% would avoid 13% of biopsies while missing 0% of cancers; a cutoff of 20% would avoid 51% of biopsies while missing 18% of cancers; and a cutoff of 50% would avoid 94% of biopsies while missing 66% of cancers.

In 2016 Boegemann et al reported results of a study of 769 European men ages 65 years and younger scheduled for initial or repeat prostate biopsy who were prospectively and retrospectively enrolled. The investigators compared phi to other PSA measures for detecting clinically significant versus insignificant cancer (PRIAS-criteria: T-stage T1c/T2; Gleason score ≤6; number of positive cores per biopsies ≤2; tPSA ≤10 ng/mL; PSA density <0.2 ng/mL). The AUC for phi (0.72; 95% CI, 0.68 to 0.76) was higher than that for tPSA (0.62; 95% CI, 0.58 to 0.66) or %fPSA (0.64; 95% CI, 0.60 to 0.68).

Morote et al reported numerically higher but not statistically significantly higher AUC for phi compared to tPSA or %fPSA for detecting aggressive prostate cancer in 357 men with PSA levels between 3 and 10 ng/mL scheduled for first biopsy in a retrospective study in Spain. Similarly, Yu et al reported numerically but not statistically significantly higher AUC for phi versus tPSA in 114 men in China with PSA levels between 2 and 10 ng/mL and negative DRE.

Section Summary: Clinical Validity
Many studies and systematic reviews of these studies have reported on the clinical validity of phi. In general, the comparator was a component of PSA (tPSA, fPSA) but have not included other risk factors from a standard clinical exam. Most of the primary studies included men with positive, negative, and inconclusive DRE and men with PSA levels outside of the 4- to 10-ng/mL range. African Americans have a high burden of morbidity and mortality but were not well represented in the study populations. There is no standardization of cutoffs used in a clinical setting for diagnosis and data on the diagnostic accuracy of phi for distinguishing clinically significant from insignificant cancer are lacking.

Clinical Utility
No studies directly measuring the effect of phi on clinical outcomes were found. An indirect chain of evidence might be used to demonstrate clinical utility if each link in the chain is intact. The phi test is associated with diagnosis of prostate cancer, although data on association with diagnosis of aggressive prostate cancer are lacking. The phi test provided better diagnostic information than other measures of PSA alone but comparison decisions made including other risk factors from clinical examination were not provided in most studies. Optimal cutoffs for classifying men into risk groups have not been standardized. No studies were found describing differences in management based on phi risk assessment. The indirect chain is incomplete.

Section Summary: pro-PSA and Prostate Health Index
The analytic validity of phi has been established. At least 4 systematic reviews including 1 HTA (NICE) have been reported and included many primary studies. In general, selected studies included some men outside of the intended use
population (PSA levels outside of the 4 to 10 ng/mL range and abnormal DRE). Comparisons to diagnosis with clinical examination were lacking. The cutoffs for categorizing men into risk groups in clinical practice have not been standardized and therefore there is heterogeneity in reporting of performance characteristics and decision curve analyses.

**Metabolic Biomarkers**

**Prostarix**

Prostarix (Metabolon/Bostwick Laboratories) is a post-DRE urine test based on a panel of biomarkers and is used in the early detection of prostate cancer. The results are intended to aid in the clinical decision to proceed to biopsy or repeat biopsy the prostate, particularly in patients who have a suspicious DRE and modestly elevated PSA (2.5-10 ng/mL). The test addresses metabolic abnormalities associated with prostate cancer. Prostarix measures the concentration of several metabolites: sarcosine, alanine, glycine, and glutamate, and these quantitative measurements are combined in a logistic regression algorithm to generate a Prostarix Risk Score. If PSA level and TRUS-determined prostate volume are available, they can be used with the metabolite measurements to generate the Prostarix-PLUS Risk Score. The test claims to have increased sensitivity and specificity over standard assessment tools to predict the likelihood of a positive prostate biopsy.

Two studies (described next) correlated the level of sarcosine in urine of prostate biopsy-positive and -negative patients, and found increased levels of sarcosine in the urine of patients with prostate cancer; however, is not clear in which patient population a test measuring urine sarcosine would be used, or what level of sarcosine would warrant a prostate biopsy. In addition, other studies done by different authors have shown conflicting results from those performed by Metabolon researchers.

In their initial study of the potential role of metabolomic profiles to delineate the role of sarcosine in prostate cancer progression, Sreekumar et al profiled 1126 metabolites across 262 prostate-derived clinical samples (42 tissue samples and 110 matched specimens of plasma and post-DRE urine from biopsy-positive cancer patients [n=59] and biopsy-negative control patients [n=51]). The authors reported that levels of sarcosine increased progressively in benign, localized prostate cancer, and metastatic disease.

Subsequently, the investigators used benign prostate tissue and localized prostate cancer obtained from a radical prostatectomy series from 1 university hospital. Urine specimens were collected from patients being screened for prostate cancer with PSA levels considered clinically significant (8.59±6.30 pg/nL). Urine was collected post-DRE but before prostate biopsy. Urine collected from patients undergoing prostatectomy was collected before surgery and used as a positive control. In total, 211 biopsy-positive and 134 biopsy-negative urine sediments were used. Using a logistic regression model, sarcosine levels were elevated in prostate cancer urine sediments compared with controls, with AUC of 0.71.
Genomic Biomarkers

PCA3 and Progensa
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 prostate cells released into the urine (PCA3 score), the test has significantly improved specificity compared with serum PSA and may better discriminate patients with benign findings on (first or second) biopsy from those with malignant biopsy results.

The Progensa PCA3 assay (Hologic Gen-Probe) has been approved by FDA to aid in the decision for repeat biopsy in men 50 years or older who have had 1 or more negative prostate biopsies and for whom a repeat biopsy would be recommended based on current standard of care. The Progensa PCA3 assay should not be used for men with atypical small acinar proliferation on their most recent biopsy. The manufacturer’s website states that the test is intended to identify men who have negative first biopsy results to determine who needs a follow-up biopsy and that a PCA3 score less than 25 is associated with a decreased likelihood of a positive biopsy.60

Analytic Validity
The analytic validity of the Progensa PCA3 has been reviewed by FDA61 and in an HTA for NICE.48 Limit of blank was reported as 0.50 pg/mL, limit of detection was 0.69 pg/mL, and limit of quantitation was 3.23 pg/mL. No assay interference was recorded in the SSED report. The SSED report included carryover studies with a 0% false-positive rate for negative samples interspersed with high-titer samples. Accuracy was calculated by percent recovery of PCA3 compared with ultraviolet-determined copies per milliliters of 8-member panel of female urine spiked with in vitro transcript; the minimum was 90% and the maximum was 118%. Precision as measured by percent coefficient of variation for within- and between-laboratory variation ranged from 12.3 to 25 for PCA3 score in 3 control samples. Linearity was assessed with 11 samples with in vitro transcripts in processed female urine, and the deviation from linearity for PCA3 was less than 9%.

Clinical Validity

Systematic Reviews
Several systematic reviews and meta-analyses have described the clinical validity of PCA3 and Progensa. The characteristics of the reviews are in Table 5. The reviews cover studies reported up to 2014. All primary studies included in the reviews were observational, although 1 study used the placebo arm from an RCT and a validation trial not included in the reviews is described below. The reviewers selected studies of men with positive, negative, or inconclusive DRE without restrictions on PSA levels. The 2 most recent reviews (Cui et al,62 Nicholson et al48) are detailed below.
In 2016 Cui et al reported results of systematic review including a search of PubMed and EMBASE for case-control or cohort studies. Quality was assessed using the QUADRAS tool. Pooled estimates were calculated using random-effects models and summarized ROCs when evidence of threshold effect was detected. Nicholson et al was described in the section on phi. In brief, an HTA commissioned by NICE included studies with men for whom initial prostate biopsy results were negative or equivocal.

**Table 5. Characteristics of Systematic Reviews of the Clinical Validity for Progenesa for Diagnosing Prostate Cancer**

<table>
<thead>
<tr>
<th>Study</th>
<th>Dates</th>
<th>Key Inclusion Criteria</th>
<th>Design</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cui (2016)</td>
<td>Up to 2014</td>
<td>Biopsy as reference standard</td>
<td>Prospective, retrospective (case-control or cohort) OBS</td>
</tr>
<tr>
<td>Nicholson (2015) (NICE)</td>
<td>2000-2014</td>
<td>Initial prostate biopsy negative or equivocal, 6+ cores in initial biopsy, includes ±DRE</td>
<td>Prospective and mixed (prospective/retrospective) OBS (1 included cohort from RCT)</td>
</tr>
<tr>
<td>Bradley (2013) (AHRQ)</td>
<td>1990-2012</td>
<td>Comparative data</td>
<td>Cohort studies</td>
</tr>
</tbody>
</table>

DRE: digital rectal exam; OBS: observational; RCT: randomized controlled trial.

Results from the systematic reviews are shown in Table 6. Cui et al included 46 studies with over 12,000 men. The quality of the selected studies was rated as moderate to high. The most common PCA3 cutoff for categorizing low and high risk was 35; 25 studies had a PCA3 cutoff of 35. Most were performed in the United States and Europe; 5 were conducted in Asia. The estimates of AUC were lower for studies including men having repeated (0.68; 95% CI, 0.67 to 0.70) versus initial (0.80; 95% CI, 0.78 to 0.82) biopsies. AUC values were 0.74 (95% CI, 0.73 to 0.76) for studies with a cutoff value not equal to 35 and 0.77 (95% CI, 0.75 to 0.79) for studies with a cutoff value equal to 35, although the group with varying cutoff (≠35) had a greater range and more variable performance estimates.

Nicholson et al included 13 studies describing 11 cohorts including 1 from the placebo arm of a randomized controlled trial (RCT). The reviewers found that criteria for referral for repeat biopsy were often unclear and varied across studies. The criteria also differed on whether normal or abnormal DREs were included and the mean or median PSA, when reported, ranged from 4.9 to 11.0 ng/mL. The prevalence of cancer on repeat biopsy varied from 11.4% to 68.3%. Meta-analyses were not performed due to heterogeneity. Some studies used PCA3 scores as a continuous variable and others created risk categories. The addition of PCA3 to clinical assessment, as a continuous or categorical variable, generally led to improvement in AUC. The comparisons with respect to diagnostic odds ratio (OR) were mixed, although most studies found increased diagnostic accuracy for PCA3 plus clinical assessment compared to clinical assessment alone. Studies that fixed sensitivity and derived specificity and those that reported decision curve...
analysis had mixed results. The reviewers concluded that the clinical benefit of PCA3 in combination with clinical assessment was not confirmed.

Table 6. Results of Systematic Reviews of the Clinical Validity for PCA3 or Progensa for Diagnosing Prostate Cancer

<table>
<thead>
<tr>
<th>Study</th>
<th>Studies</th>
<th>N (Range)</th>
<th>Outcome</th>
<th>Results (95% CI)</th>
</tr>
</thead>
</table>
| Cui (2016)                 | 46      | 12,295 (NR)        | Diagnostic performance, any prostate cancer                            | • Sensitivity range, 47%-95%  
  Significant heterogeneity and threshold effect  
  Pooled estimate, 0.65 (0.63 to 0.66)  
  • Specificity range, 22%-100%  
  Significant heterogeneity and threshold effect  
  Pooled estimate: 0.73 (0.72 to 0.74)  
  • Negative LR:  
    Pooled estimate, 0.48 (0.44 to 0.52)  
  • AUC=0.75 (0.74 to 0.77) |
| Nichollson (2015) (NICE)   | 11      | 3336 (41-1072)     | Diagnostic performance, any prostate cancer                            | • AUC range:  
  Clinical assessment alone, 0.55-0.75  
  Clinical assessment plus PCA3, 0.61-0.76  
  Derived sensitivity (at given specificity) range:  
  Clinical assessment alone, 44%-48% (at 80%)  
  Clinical assessment plus PCA3, 39%-46% (at 80%)  
  • Derived sensitivity (at given specificity)  
    tPSA=91% (at 20%)  
    PCA3=96% (at 20%)  
  • Unable to compare performance for aggressive prostate cancer |
| Bradley (2013) (AHRQ)      | 43      | 9719 biopsies or prostatectomies (32-1246)  | Diagnostic performance, any prostate cancer and aggressive prostate cancer | • Sensitivity range, 47%-82%  
  • Pooled sensitivity, 85% (84% to 87%)  
  • Specificity range, 56%-89%.  
  • Pooled specificity, 96% (96% to 97%)  
  • Pooled negative LR=0.15 (0.13 to 0.18) |
| Ruiz-Aragon (2010)         | 14      | 3467 (30-563)      | Diagnostic performance, any prostate cancer                            | • Sensitivity range, 47%-82%  
  • Pooled sensitivity, 85% (84% to 87%)  
  • Specificity range, 56%-89%.  
  • Pooled specificity, 96% (96% to 97%)  
  • Pooled negative LR=0.15 (0.13 to 0.18) |

AUC: area under the curve; CI: confidence interval; LR: likelihood ratio; NR: not reported; tPSA: total prostate-specific antigen.

**Randomized Controlled Trials**

A 2014, the National Cancer Institute conducted a prospective trial to validate the diagnostic use of PCA3 to complement PSA-based detection of prostate cancer. 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 end point was the diagnosis of prostate cancer on biopsy and the secondary study end point 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 (PPV) 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, performance of Prostate Cancer Prevention Trial’s (PCPT) risk calculator was improved by adding PCA3 to the 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≤0.003).

Other Relevant Clinical Studies
The pivotal study describing in the FDA SSED for Progensa included 495 men from 15 clinical sites who had at least 1 negative prostate biopsy and were recommended for repeat biopsy by their urologist. Prostate biopsy was performed per each site’s local standard procedure. A total of 433 (87.5%) were white, 45 (9.1%) were African American or black, and 11 (2.2%) were Asian. A valid PCA3 score and biopsy result was available for 466 men. Using a cutoff of PCA3 score of 25, the performance characteristics for positive biopsy were as follows: sensitivity, 77.5% (95% CI, 68.4% to 84.5%); specificity, 57.1% (95% CI, 52.0% to 62.1%); PPV, 33.6% (95% CI, 30.0% to 37.2%); and NPV, 90.0% (95% CI, 86.5% to 93.1%). Two hundred eight men in the study might have been spared an unnecessary repeat biopsy if a cutoff of 25 was used to recommend repeat biopsy. On the other hand, 23 of the men who had a biopsy positive for prostate cancer may have had their diagnosis delayed due to negative PCA3 result.

Clinical studies compared PCA3 to clinical examination and risk calculators and those focuses on distinguishing aggressive versus indolent cancer are particularly relevant. Ankerst et al (2008) reported that incorporating PCA3 score into the PCPT risk calculator improved the diagnostic accuracy of the calculator (from AUC of 0.653 to 0.696). Chun et al (2009), using a multivariate nomogram, demonstrated a 5% gain in predictive accuracy when PCA3 was incorporated with other predictive variables such as age, DRE results, PSA levels, prostate volume, and biopsy history. In a 2011 study of 218 patients with PSA values of 10 ng/mL or less, Perdona et al performed a head-to-head comparison of these 2 risk assessment tools and suggested both might have value in clinical decision making.

Several studies have evaluated the PCA3 score as a tool for distinguishing between patients with indolent cancers who may need only active surveillance and those with aggressive cancers who warrant aggressive therapy. Three studies from 2008—Haese et al, Nakanishi et al, and Whitman et al—demonstrated an association between PCA3 scores and evidence of tumor aggressiveness. However, these findings were not confirmed in a 2006 study by Bostwick et al or a 2008 study by vans Gils et al. Auprich et al (2011) reported that PCA3 scores
appeared to enhance identification of indolent disease but not pathologically advanced or aggressive cancer.\textsuperscript{75}

Tosoian et al (2010) reported on a short-term prospective cohort study evaluating \textit{PCA3} in relation to outcomes in an active surveillance program involving 294 patients.\textsuperscript{76} \textit{PCA3} did not distinguish patients who had stable disease from those with more aggressive features.

\textbf{Clinical Utility}

Clinical utility studies using assay results for decision making for initial biopsy, repeat biopsy, or treatment have not been reported, nor have studies of the effects of using assay results on clinical outcomes. Several studies using decision analysis to estimate the cost-benefit tradeoff between reduction in unnecessary biopsies and missed prostate cancers have been published. One group reported potential reductions in unnecessary biopsies of 48\% to 52\%, with attendant increases in missed prostate cancers of 6\% to 15\% using either a \textit{PCA3}-based nomogram\textsuperscript{77} or \textit{PCA3} level corrected for prostate volume (\textit{PCA3} density).\textsuperscript{78} Although both studies were prospective, neither assessed utility of the test for clinical decision making because all patients underwent biopsy. Merdan et al used decision analysis to simulate long-term outcomes associated with use of the \textit{PCA3} score to trigger repeat biopsy compared to the PCPT risk calculator in men with at least 1 previous negative biopsy and elevated PSA levels.\textsuperscript{79} They estimated that incorporating the \textit{PCA3} score (with a biopsy threshold of 25) into the decision to recommend repeat biopsy could avoid 55.4\% of repeat biopsies, with a 0.93\% reduction in the 10-year survival rate.

Given the lack of direct evidence of utility, an indirect chain of evidence would be needed to demonstrate clinical utility. The \textit{PCA3} test is associated with diagnosis of prostate cancer, although data on the association with diagnosis of aggressive prostate cancer are lacking. The \textit{PCA3} test provided better diagnostic information than other measures of PSA but comparison with decisions made using risk factors from clinical examination were not provided in most studies. No prospective studies were found describing differences in management based on \textit{PCA3} risk assessment. The indirect chain is incomplete.

\textbf{Section Summary: \textit{PCA3} and Progensa}

The analytic validity of Progensa has been established. At least 4 systematic reviews including 1 HTA have been reported and included many primary studies. Studies of \textit{PCA3} as a diagnostic test for prostate cancer have 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 \textit{PCA3} results. Cutoffs for recommending repeat biopsy with the Progensa test have suggested by the manufacturer and were used in a validation study for FDA approval. The clinical utility of \textit{PCA3} tests is uncertain, because there is no evidence that its use can change management in ways that improve outcomes.
Gene Hypermethylation and ConfirmMDx

Epigenetic changes—chromatin protein modifications that do not involve changes to the underlying DNA sequence but can change gene expression—have been identified in specific genes. An extensive literature has reported significant associations between epigenetic DNA modifications and prostate cancer. Several investigators have evaluated detection of hypermethylation products in biological fluids for early detection of prostate cancer. Promoters of 3 genes (APC, GSTP1, RARβ2) were identified early as potentially involved in prostate carcinogenesis. GSTP1 is the most widely studied methylation marker for prostate cancer, usually as a diagnostic application. Studies in the late 2000s of GSTP1 hypermethylation using tissue samples reported conflicting results. Sunami et al assayed blood from 40 healthy individuals and 83 men with prostate cancer using a 3-gene cohort (GSTP1, RASSF1, RARβ2) and demonstrated sensitivity of 28% for cancer patients. Trock et al (2012) conducted a small (86-patient) diagnostic exploratory cohort study and showed that hypermethylation of adenomatosus polyposis coli (APC) was associated with high sensitivity and high specificity for cancer on repeat biopsy.

In a 2012 meta-analysis by Van Neste et al, 30 peer reviewed studies of hypermethylation of GSTP1 and other genes in prostate tissue were evaluated. The pooled estimates of sensitivity for GSTP1 to distinguish prostate cancer from normal in biopsies (328 cases, 263 controls) was 82%, with 95% specificity, 95% NPV, and 85% PPV. The combination of GSTP1, APC, and RARβ had a sensitivity of 95%, specificity of 95%, NPV of 99%, and PPV of 95%. Van Neste et al suggested that a valuable first step in diagnostic use might be to test for methylated genes to select patients undergoing prostate biopsy who might not require repeat biopsy.

Following the 2012 meta-analysis, several studies reported associations between DNA hypermethylation at various gene loci (RASSF1A, APC, GSTP1, PTGS2, RAR-beta, TIG1, AOX1, C1orf114, GAS6, HAPLN3, KLF8, MOB3B) and prostate cancer. In contrast, Kachakova et al (2013) found that HIST1H4K hypermethylation was more likely due to aging than to prostate carcinogenesis.

ConfirmMDx (MDxHealth) is a commercially available test for gene methylation intended to distinguish true- from false-negative prostate biopsies to avoid the need for repeat biopsy in cases of a true negative and to identify men who may need a repeat biopsy. The test measures methylation of the genes GSTP1, APC, and RASSF1.

Analytic Validity

Goessl et al confirmed in 26 patients that neoplastic transformation could be identified in washings of prostate biopsies by GSTP1 promoter hypermethylation using methylation-specific polymerase chain reaction (PCR). Chu et al described a protocol for real-time, quantitative, methylation-sensitive PCR for detecting the methylation change in the 5’ regulatory sequence flanking the GSTP1 gene that was more sensitive than conventional nested PCR (test limitations were 0.048 and 0.64 ng DNA, respectively). Mehrotra et al confirmed that a field effect was detectable for APC, RARβ2, and RASSF1A up to 3 mm from the malignant core.
Van Neste et al described a study evaluating multiplex assay consisting of 3 genes: GSTP1, APC, and RASSF1. Thirty cancer-positive tissue samples and 12 cancer-free controls were analyzed with 4 singleplex versus 1 multiplex assay. A control gene (ACTB) was used to estimate DNA quantity and quality in 2 replicates. The ratio of ACTB copies ranged from 0.73 to 1.17 (outlier removed) for the multiplex assay, with median ratio of 1.0. ACTB copy numbers were higher for the multiplex assay than for a singleplex assay (median, 1.5-fold copy increase). A linear regression model yielded amplification factors of 1.57, 1.19, 4.13, and 1.25 for the ACTB, GSTP1, APC, and RASSF1 assays respectively with consistently high $R^2$ values (>0.90). Biopsies consisting of 10, 20, and 40 μm from formalin-fixed, paraffin-embedded (FFPE) tissue blocks from the minimization cohort were tested and compared (outliers were removed). The effect of the original sample volume on the relative DNA yield was minor, indicating that samples as small as 20 μm can be used to detect methylation. Older samples showed lower relative DNA yields (p<0.001), indicating that the age of FFPE samples does have a negative impact on DNA quality and quantity. Other measures of analytic validity were not found in the literature or the ConfirmMDx website. The laboratory that performs the analyses for ConfirmMDx is certified under CLIA.

**Clinical Validity**

Two blinded multicenter validation studies of the ConfirmMDx test have been performed. Partin et al reported results of the DOCUMENT study that evaluated archived, cancer-negative prostate biopsy core tissue samples from 350 men from 5 U.S. urology centers. All patients underwent repeat biopsy within 24 months. Men with 2 consecutive negative biopsies were classified as controls and men with a negative biopsy followed by a positive biopsy were classified as cases. Thirty (9%) men were excluded from analysis because of noneligibility (n=2), insufficient DNA (n=1), insufficient biopsy cores (n=23), or detection of adenocarcinoma in the first biopsy based on central pathology review (n=4); 320 men were included in analysis (92 cases, 228 controls). Median age was 62 years (range, not given). Median PSA level was 5.3 ng/mL; 23% of men had PSA less than 4 ng/mL and 10% had a PSA of 10 ng/mL or higher. Sixty percent of men had a normal DRE. Forty-two (13%) of the men were black, 232 (73%) were white, and 13 (4%) were Asian. The ConfirmMDx test, performed on the first biopsy, resulted in a NPV of 88% (95% CI, 85% to 91%), sensitivity of 62% (95% CI, 51% to 72%), and specificity of 64% (95% CI, 57% to 70%). The study was not powered to accurately determine the performance characteristics in a subgroup of black patients, but the estimated sensitivity was 77% (95% CI, 46% to 95%), specificity was 66% (95% CI, 46% to 82%), and NPV was 93% (85% CI 82% to 97%). Multivariate analysis of potential predictors of cancer on repeat biopsy, corrected for age, PSA, DRE, first biopsy histopathology characteristics, and race, showed that the ConfirmMDx test was the most significant independent predictor of patient outcome (OR=2.69; 95% CI, 1.60 to 4.51).

The MATLOC study reported by Stewart et al tested archived cancer-negative prostate biopsy needle core tissue samples from 498 men from the U.K. and Belgium. Patients underwent repeat biopsy within 30 months; cases had a
positive second biopsy while controls had a negative second biopsy. A total of 483 men were included in analysis (87 cases, 396 controls). The median PSA level was 5.9 ng/mL; 21% of men had PSA less than 4 ng/mL and 18% had PSA of 10 ng/mL or higher. Seventy-three percent of men had benign DRE. The ConfirmMDx test, performed on the first biopsy, resulted in a NPV of 90% (95% CI, 87% to 93%), sensitivity of 68% (95% CI, 57% to 77%), and specificity of 64% (95% CI, 59% to 69%). Multivariate analysis of potential predictors of cancer on repeat biopsy, corrected for patient age, PSA, DRE, and first biopsy histopathology characteristics, showed that the ConfirmMDx test was the most significant independent predictor of patient outcome (OR=3.17; 95% CI, 1.81 to 5.53).

In 2016, Van Neste et al reported results of combined data from the DOCUMENT and MATLOC studies to investigate whether DNA methylation intensities were associated with high-grade (Gleason score, ≥ 7) prostate cancer. DNA methylation was the most significant and important predictor of high-grade cancer, resulting in a NPV of 96% (precision not reported).

**Clinical Utility**

In 2014, Wojno et al reported a field observation study in which practicing urologists at 5 centers used the ConfirmMDx test to evaluate at least 40 men with previous cancer-negative biopsies who were considered at risk for prostate cancer. Centers reported whether patients who had a negative test assay result had undergone a repeat biopsy at the time of the analysis. Median patient follow-up time after the assay results were received was 9 months. A total of 138 patients were included in the analysis. The median PSA level was 4.7 ng/mL. Repeat biopsies had been performed in 6 (4.3%) of the 138 men with a negative ConfirmMDx test, in which no cancer was identified.

In 2013, Aubry et al analyzed the expected reduction in biopsies associated with ConfirmMDx use. Using the MATLOC estimates of performance characteristics for ConfirmMDx, the authors estimated that 1106 biopsies per 1 million people would be avoided. The study did not include decision analysis comparing the tradeoff in reduction in biopsies and missed cancers. The analysis is limited by the uncertainty in performance characteristics.

MDxHealth completed enrollment into the PASCUAL trial in April 2015. The PASCUAL trial is a randomized clinical utility study of ConfirmMDx to evaluate the impact of the test on physician decisions for repeat biopsy. Results have not yet been published.

No studies were found that directly show the effects of using ConfirmMDx results on clinical outcomes. Given the lack of direct evidence of utility, an indirect chain of evidence would be needed to demonstrate clinical utility. The ConfirmMDx test is associated with diagnosis of prostate cancer and aggressive prostate cancer. The validity studies of the ConfirmMDx test included men in the intended use population but did not include comparison of performance characteristics to clinical examination with %fPSA. One survey of urologists who had previously used the ConfirmMDx test found that most ConfirmMDx negative patients did not have a
biopsy. Prospective data on utility should be available after completion of PASCUAL. No data are available on the longer term clinical outcomes of the men who did not have biopsy based on ConfirmMDx results. The indirect chain is incomplete.

**Section Summary: Gene Hypermethylation and ConfirmMDx**

Two clinical validation studies have reported on the clinical validity of the ConfirmMDx score in the intended use population. The studies did not provide estimates of validity compared to a standard clinical examination with %fPSA. No data are available on the long-term clinical outcomes or clinical utility of the test. The indirect chain of evidence is incomplete due to the limitations in evidence on the comparative clinical validity and utility.

**TMPRSS Fusion Genes and Mi-Prostate**

*TMPRSS2* is an androgen-regulated transmembrane serine protease that is preferentially expressed in normal prostate tissue. In prostate cancer, it may be fused to an ETS (E26 transformation-specific) family transcription factor (*ERG*, *ETV1*, *ETV4*, or *ETV5*), which modulates transcription of target genes involved in cell growth, transformation, and apoptosis. The result of gene fusion with an *ETS* transcription factor is that the androgen-responsive promoter of *TMPRSS2* upregulates expression of the *ETS* gene, suggesting a mechanism for neoplastic transformation. Fusion genes may be detected in tissue, serum, or urine.

*TMPRSS2:ERG* gene rearrangements have been reported in 50% or more of primary prostate cancer samples. Although *ERG* appears to be the most common ETS family transcription factor involved in the development of fusion genes, not all are associated with *TMPRSS2*. About 6% of observed rearrangements are seen with *SLC45A3*, and about 5% appear to involve other types or rearrangement.

In 2014, Yao et al published a systematic review with meta-analysis of *TMPRSS2:ERG* for the detection of prostate cancer. Literature was searched through July 2013, and 32 articles were included. 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 ($I^2>85\%$). It was unclear whether studies in screening populations were pooled with enriched patient samples (eg, elevated PSA and/or negative biopsy). There also was variability in the type of tissue samples analyzed (urine, prostatic secretions, biopsy, surgical specimens); the type of *TMPRSS2:ERG* assays used (fluorescence in situ hybridization [FISH], immunohistochemistry, real-time reverse transcriptase polymerase chain reaction, transcription-mediated amplification); and in *TMPRSS2:ERG* threshold cutoff values.

The Mi-Prostate (MiPS) is a test using the *TMPRSS2:ERG* gene to produce a risk probability for detection of prostate cancer and aggressive prostate cancer by standard biopsy. The probability score is calculated with logistic regression models that incorporate serum PSA, or the PCPT version 1.0, and urine *T2:ERG* and PCA3.
scores. The test was developed by, and is only available from, the University of Michigan MLabs, and may be used to make a decision about monitoring PSA levels or pursuing a prostate biopsy.

**Analytic Validity**
The MiPS test uses results from the Progensa PCA3 test that has demonstrated analytic validity in FDA submission. The amounts of urine $\text{TMPRSS2:ERG}$ are determined using transcription-mediated amplification assays. No peer-reviewed, full-length publications describing the analytic validity of the $\text{TMPRSS2:ERG}$ assays were identified.

**Clinical Validity**
Tomlins et al (2011) developed a transcription-mediated amplification assay to measure $\text{TMPRSS2:ERG}$ fusion transcripts in parallel with $\text{PCA3}$.\(^{106}\) Combining results from these 2 tests and incorporating them into the multivariate Prostate Cancer Prevention Trial risk calculator appeared to improve identification of patients with clinically significant cancer using Epstein criteria and high-grade cancer on biopsy. Although the study was large (1312 men at multiple centers), it was confounded by assay modifications during the study and by use of cross-validation rather than independent validation, using independent training and testing sets. Further studies are warranted.

In 2013, this same group evaluated 45 men using a multivariable algorithm that included serum PSA plus urine $\text{TMPRSS2:ERG}$ and $\text{PCA3}$ from a post-DRE sample.\(^{107}\) Samples were collected before prostate biopsy at 2 centers. For cancer prediction, sensitivity and specificity were 80% and 90%, respectively. AUC was 0.88.

In 2016, Tomlins et al published results of a validation study of the MiPS score in 1244 prospectively collected, post-DRE urine samples from 7 U.S. clinics.\(^{108}\) A total of 1225 of the specimens had sufficient materials for both $\text{TMPRSS2:ERG}$ and $\text{PCA3}$ analysis and were included. Eighty percent of patients were presenting for initial biopsy. Seventy-three percent were white; the number of African Americans was not given. Approximately 25% of the men were older than 70. Twenty-three percent had an abnormal DRE and the median PSA level was 4.7 ng/mL. The AUCs for predicting any cancer using PSA alone, PCPT risk calculator alone, and the MiPS score alone were 0.59, 0.64, and 0.76, respectively (CIs not given, $p<0.001$ for MiPS vs PCPT). The AUCs for predicting high-grade cancer were 0.65, 0.71, and 0.78, respectively ($p<0.001$ for MiPS vs PCPT). A MiPS score threshold for recommending biopsy has not been provided and so sensitivity and NPV were not calculated.

**Clinical Utility**
Tomlins et al (2016) also used decision curve analysis to estimate the number of biopsies that would have been performed and cancers that would have been missed using a MiPS risk cutoff for biopsy in their cohort.\(^{108}\) Compared with a biopsy-all strategy, using a MiPS cutoff for aggressive cancer of 15% would have avoided 36% of biopsies while missing 7.0% of any prostate cancer and 1.6% of high-grade prostate cancer diagnoses. Using the PCPT risk calculator cutoff of 15%
for aggressive cancer would have avoided 68% of biopsies while missing 25% of any cancer and 8% of high-grade cancer.

No studies were found that directly show the effects of using MiPS results on clinical outcomes. Given the lack of direct evidence of utility, an indirect chain of evidence would be needed to demonstrate clinical utility. The MiPS test is associated with diagnosis of prostate cancer and aggressive prostate cancer. The clinical validity study of the MiPS test included men with relevant PSA levels but also included men with positive DRE who would not likely forgo biopsy. The clinical validation study included comparison of performance characteristics to standard risk calculators; comparison to %fPSA was not provided. Confirmation of performance characteristics is needed. No prospective data are available on using the MiPS score for decision making. No data are available on the longer term clinical outcomes of the men who did not have biopsy based on MiPS results. The indirect chain is incomplete.

Section Summary: TMPRSS Fusion Genes and Mi-Prostate
Concomitant detection of *TMPRSS2:ERG* and *PCA3* may more accurately identify men with prostate cancer. However, current evidence is insufficient to support its use. Estimated accuracy varies across available studies. The Mi-Prostate (MiPS) test has preliminary data suggesting improved clinical validity compared to the PCPT risk calculator in a validation study but independent confirmation of clinical validity and comparison to %fPSA is needed. Data on analytic validity and clinical utility are lacking.

**Prostate Core Mitomics Test**
The Prostate Core Mitomics Test (PCMT; Mitomics; formerly Genesis Genomics) is a proprietary test that 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 a positive PCMT means that the patient is at high risk of undiagnosed prostate cancer. The website also states that physicians should consider using PCMT for patients who have a negative initial biopsy but continue to have elevated PSA, rising PSA, irregular DRE, atypical small acinar proliferation, high-grade prostatic intraepithelial neoplasia, or inconclusive biopsy.  

**Analytic Validity**
No peer-reviewed, full-length publications on the analytic validity of the commercially available PCMT test were identified.

**Clinical Validity**
A 2006 study retrospectively analyzed mitochondrial DNA mutations from 3 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 (66.7%) of 24 had mutations in all 3 tissue types, 22 (91.7%) of 24 had mutations in malignant samples, 19 (79.2%) of 24 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, 7 (58.3%) patients had between 1 and 5 alterations, mainly in noncoding regions. The authors concluded that the mutations found in the malignant group versus the control group differed significantly and that mitochondrial DNA mutations are an indicator of malignant transformation in prostate tissue.

In 2008, Maki et al reported the discovery and characterization of a 3.4-kb mitochondrial genome deletion and its association with prostate cancer. 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 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 that 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 had 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 on 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<0.000) and benign and proximal (p<0.003) samples. The PTM samples closely resembled the malignant sample, with no statistically significant resolution between the scores (p<0.833), to which the authors attributed to 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 AUC ROC was 0.83 for the deletion. A blinded, external validation study showed a sensitivity and specificity of 83% and 79% and the AUC ROC of 0.87.

In 2010, Robinson et al. assessed the clinical value of the 3.4-kb deletion described in the Maki study in predicting rebiopsy outcomes. Levels of the deletion were measured by quantitative PCR in prostate biopsies negative for cancer from 101 patients who underwent repeat biopsy within 1 year and had known outcomes. Of the 101 first biopsies, the diagnosis was normal in 8, atypical and/or had prostatic intraepithelial neoplasia 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. Final data were 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 AUC ROC of 0.75. NPV was 91%.

**Clinical Utility**
No peer-reviewed, full-length publications on the clinical utility of the commercially available PCMT test was identified.

**Section Summary: Prostate Core Mitomics Test**
The PCMT test has preliminary data on performance characteristics in a small validation study but independent confirmation of clinical validity is needed. The studies did not provide estimates of validity compared to a standard clinical examination. No data is available on the long-term clinical outcomes. Data on analytic validity and clinical utility are lacking.

**Candidate Gene Panels and Single-Nucleotide Polymorphism Testing**
Because no single gene marker that is both highly sensitive and highly specific for diagnosing prostate cancer has been found, particularly in men already known to have elevated PSA levels, some investigators are combining several markers into a single diagnostic panel. Although promising in concept, only single studies of various panels have been published, and none apparently is offered as a clinical service.

Single-nucleotide polymorphisms (SNPs) occur when a single nucleotide is replaced with another, and they are the most common type of genetic variation in humans. They occur normally throughout the genome and can act as biological markers for disease association. Genome-wide association studies have identified correlations between prostate cancer risk and specific SNPs. However, it is generally accepted that individually, SNP-associated disease risk is low and of no value in screening for disease, although multiple SNPs in combination may account for a higher proportion of prostate cancer. Investigators have begun to explore the use of algorithms incorporating information from multiple SNPs to increase the clinical value of testing.

Ma et al (2014) examined various algorithms for cancer diagnosis and prognosis using urine and plasma levels of multiple genes, including **PCA3**, **PSA**, **TMPRSS2**, and **ERG**. One algorithm distinguished prostate cancer from benign prostatic hypertrophy with an AUC of 0.78. Another algorithm distinguished men with a Gleason score 7 or higher from men with a Gleason score less than 7 (AUC=0.88). Combination of these 2 algorithms into a scoring system predicted the presence of a Gleason score 7 or higher in 75% of men. Qu et al (2013) reported preliminary results of a 3-gene panel (androgen receptor (AR), **PTEN**, **TMPRSS2:ERG** analyzed by FISH. Thirty-one percent of 110 archived primary tumor samples and 97 metastatic tumor samples from a separate cohort of patients were analyzable. Chromosomal abnormalities were detected in 53% of primary prostate cancers and in 87% of metastatic tumors (p<0.001).

In 2015, Leyten et al reported on development of a gene panel using specimens from 133 patients that included 3 urinary biomarkers (**HOXC6**, **TDRD1**, **DLX1**).
When the gene panel was used with PSA, the combined AUC for predicting high-grade prostate cancer was 0.81 (95% CI, 0.75 to 0.86), which was higher than the concurrently measured Progensa AUC of 0.68 (95% CI, 0.62 to 0.75). Xiao et al (2016) reported on the development of an 8-gene panel (PMP22, HPN, LMTK2, FN1, EZH2, GOLM1, PCA3, GSTP1) that was able to distinguish high grade prostate cancer from indolent prostate cancer with a sensitivity of 93% (95% CI, 88% to 97%), a specificity of 70% (95% CI, 36% to 104%), a PPV of 98% (95% CI, 95% to 100%), and an NPV of 61% (95% CI, 25% to 97%) in a specimen cohort of 158 men.116

A 2012 Agency for Healthcare Research and Quality report on multigene panels in prostate cancer risk assessment reviewed the literature on SNP panel tests for assessing risk of prostate cancer.117 All 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. Reviewers concluded that the evidence on the available SNP panels did not permit meaningful assessment of analytic validity, the limited evidence on clinical validity was insufficient to conclude that SNP panels would perform adequately as a screening test and that there was no evidence on the clinical utility of current panels.

Kader et al (2012) evaluated a panel of 33 prostate cancer–associated SNPs identified from genome-wide association studies in 1654 men.118 Genetic score was a significant (p<0.001) independent predictor of prostate cancer (OR=1.72; 95% CI, 1.44 to 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 but with 15% fewer biopsies.

In a 2010 review by Ioannidis et al, 27 gene variants across a large range of chromosomal locations were identified that increased risk for prostate cancer, although in all cases, the observed incremental risk was modest (OR≤1.36).119

Lindstrom et al (2011), in a study of 10,501 cases of prostate cancer and 10,831 controls, identified 36 SNPs showing association with prostate cancer risk, including 2 (rs2735893, rs266849) that showed differential association with Gleason score.120 Per allele odds ratios ranged from 1.07 to 1.44.

Ishak and Giri (2011) reviewed 11 replication studies involving 30 SNPs (19 in men of African descent, 10 in men with familial prostate cancer).121 Odds ratios were positively associated with prostate cancer, although the magnitude of association was generally small (range, 1.11-2.63).
Section Summary: Candidate Gene Panels and Single-Nucleotide Polymorphism Testing
Numerous studies have demonstrated the association of many gene panels and SNPs with prostate cancer. These studies, in early stages of development, have generally shown a modest degree of association with future risk for prostate cancer. The clinical utility of these tests is uncertain; there is no evidence that information obtained from gene panels or SNP testing can be used to change clinical management in ways that will improve outcomes.

Summary of Evidence
For individuals for whom an initial prostate biopsy or a rebiopsy is being considered who receive genetic and protein biomarker testing, the evidence includes systematic reviews and meta-analyses and primarily observational studies. Relevant outcomes are overall survival, disease-specific survival, test accuracy and validity, other test performance measures, resource utilization, hospitalizations, quality of life, and treatment-related mortality and morbidity. The evidence supporting clinical utility varies by test but has not been directly shown for any biomarker test. In general, performance of biomarker testing for predicting biopsy compared with clinical examination, including the ratio of free or unbound prostate-specific antigen (PSA) to total PSA, is lacking. However, procedures for referrals for biopsy based on clinical examination vary, making it difficult to quantify performance characteristics for this comparator. There is considerable variability in biopsy referral practices based on clinical examination alone and many biomarker tests do not have standardized cutoffs to recommend biopsy. Therefore, having prospective, comparative information on how test results are expected to be used or actually being used in practice and the associated effects on outcomes will be needed to determine if the tests improve net health outcomes. Many of the test validation populations have included men with a positive digital rectal exam, prostate-specific antigen (PSA) level outside of the gray zone, or older men for whom the information for test results are less likely to be informative. African Americans have a high burden of morbidity and mortality, but have not been well represented in these study populations. It is unclear how to monitor men with low biomarker risk scores who continue to have symptoms or high or rising PSA levels. Comparative studies of the many biomarkers are lacking and it is unclear how to use the tests in practice, particularly when test results are contradictory. The evidence is insufficient to determine the effects of the technology on health outcomes.

Practice Guidelines and Position Statements
American Urological Association
In 2013, the American Urological Association (AUA) published guidelines for the early detection of prostate cancer. Based on systematic review of the literature to 2013, AUA 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 [prostate-specific antigen] 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
In 2013, the Evaluation of Genomic Applications in Practice and Prevention Working Group published the following recommendations for PCA3 testing in prostate cancer, based on the Agency for Healthcare Quality and Research comparative effectiveness review summarized earlier:

- Evidence was insufficient to recommend “PCA3 testing to inform decisions for when to rebiopsy 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 DRE [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.”
- “…[T]he overall certainty of clinical validity to predict the diagnosis of prostate cancer using PCA3 is deemed ‘low.’ … [C]linical use for diagnosis is discouraged unless further evidence supports improved clinical validity.”
- “…[T]he overall certainty of net health benefit is deemed ‘low.’ … [C]linical use is discouraged] unless further evidence supports improved clinical outcomes.”

National Comprehensive Cancer Network
National Comprehensive Cancer Network guidelines (v.2.2016) recommend that any man with a PSA level greater than 3 ng/mL undergo workup for benign disease, repeat prostate-specific antigen (PSA) and DRE. The guidelines also recommend consideration of %fPSA, phi, and 4Kscore in patients with a PSA level greater than 3 ng/mL who have not yet had a biopsy and consideration of %fPSA, phi, 4Kscore, PCA3, and ConfirmMDx in men who had a negative biopsy but are thought to be at higher risk (category 2A evidence). The panel noted that these tests may be especially useful in men with PSA levels between 3 and 10 ng/mL. Guideline authors note:

“The panel believes that no biomarker test can be recommended over any other at this time The optimal order of biomarker tests and imaging is unknown; and it remains to unclear how to interpret results of multiple tests in individual patients – especially when results are contradictory.”

U.S. Preventive Services Task Force Recommendations
The U.S. Preventive Services Task Force published recommendations for prostate cancer screening on May 2012. Genetic and protein biomarkers addressed in this evidence review, including PCA3, were not mentioned.

Medicare National Coverage
There is no national coverage determination (NCD). In the absence of an NCD, coverage decisions are left to the discretion of local Medicare carriers. Palmetto GBA has issued a local coverage determination for positive coverage for ConfirmMDx Epigenetic Molecular Assay (date effective 11/03/14). Palmetto GBA issued a draft non-coverage policy determination in 2016 for the 4Kscore.
Ongoing and Unpublished Clinical Trials
Some currently unpublished trials that might influence this review are listed in Table 7.

Table 7. Summary of Key Trials

<table>
<thead>
<tr>
<th>NCT No.</th>
<th>Trial Name</th>
<th>Planned Enrollment</th>
<th>Completion Date</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Ongoing</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NCT01739062</td>
<td>Prostate Cancer Risk Assessment Using Genetic Markers in General Practice</td>
<td>1298</td>
<td>Jun 2017</td>
</tr>
<tr>
<td>NCT00773773</td>
<td>A Study to Assess if a Combination of Serum Measurements of Molecular Biomarkers and Serum Protein Profiling Can be Used to Predict Which Patients Undergoing Prostatic Biopsy Will be Diagnosed With Cancer</td>
<td>500</td>
<td>Oct 2017</td>
</tr>
<tr>
<td>NCT02241122</td>
<td>Improved Prostate Cancer Diagnosis - Combination of Magnetic Resonance Imaging Targeted Biopsies and Biomarkers (Multi-IMPROD)</td>
<td>400</td>
<td>Nov 2017</td>
</tr>
<tr>
<td>NCT01632930</td>
<td>Medical Economics of Urinary PCSA3 Test for Prostate Cancer Diagnosis</td>
<td>900</td>
<td>Dec 2021</td>
</tr>
<tr>
<td><strong>Unpublished</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>NCT02250313a</td>
<td>PASCUAL (Prostate Assay Specific Clinical Utility at Launch) Study</td>
<td>600</td>
<td>Mar 2015 (completed)</td>
</tr>
</tbody>
</table>

NCT: national clinical trial.
a Denotes industry-sponsored or cosponsored trial.

References


55. Boegemann M, Stephan C, Cammann H, et al. The percentage of prostate-specific antigen (PSA) isoform [-2]proPSA and the Prostate Health Index improve the diagnostic accuracy for clinically relevant prostate cancer at initial and repeat biopsy compared with total PSA and percentage free PSA in men aged <=65 years. BJU Int. Jan 2016;117(1):72-79. PMID 25818705


calculator and Chun’s nomogram, two risk estimators incorporating prostate cancer antigen 3.


**Billing Coding/Physician Documentation Information**

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<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>81479</td>
<td>Unlisted molecular pathology procedure</td>
</tr>
<tr>
<td>81539</td>
<td>Oncology (high-grade prostate cancer), biochemical assay of four proteins (Total PSA, Free PSA, Intact PSA, and human kallikrein-2 [hK2]), utilizing plasma or serum, prognostic algorithm reported as a probability score (New code as of 1/1/2017)</td>
</tr>
<tr>
<td>81599</td>
<td>Unlisted multianalyte assay with algorithmic analysis</td>
</tr>
<tr>
<td>0005U</td>
<td>Oncology (prostate) gene expression profile by real-time RT-PCR of 3</td>
</tr>
</tbody>
</table>
genes (ERG, PCA3, and SPDEF), urine, algorithm reported as risk score (new code added 2/1/2017)

**ICD-10 Codes**
- **C61** Malignant neoplasm of prostate
- **Z12.5** Encounter for screening for malignant neoplasm of prostrate

Code 0010M deleted 1/1/2017, S3721 deleted 1/1/2016

**Additional Policy Key Words**
N/A

**Policy Implementation/Update Information**
- **7/1/14** New policy; considered investigational (or a benefit exclusion)
- **7/1/15** Policy revised to focus on diagnostic testing (as well as SNP testing for cancer risk assessment). Policy statements revised to include an expanded list of diagnostic genetic and protein biomarker tests as investigational. Prognostic testing is being moved to a different policy. Title changed “Genetic and Protein Biomarkers for the Diagnosis and Cancer Risk Assessment of Prostate Cancer.”
- **7/1/16** Information about Medicare local coverage determination for ConfirmMDx Epigenetic Molecular Assay added to Medicare section. No policy statement changes.
- **12/1/16** The Prostate Health Index (phi) biomarker test was added to the investigational policy statement.
- **7/1/17** No policy statement changes.

State and Federal mandates and health plan contract language, including specific provisions/exclusions, take precedence over Medical Policy and must be considered first in determining eligibility for coverage. The medical policies contained herein are for informational purposes. The medical policies do not constitute medical advice or medical care. Treating healthcare providers are independent contractors and are neither employees nor agents Blue KC and are solely responsible for diagnosis, treatment and medical advice. No part of this publication may be reproduced, stored in a retrieval system or transmitted, in any form or by any means, electronic, photocopying, or otherwise, without permission from Blue KC.