2.04.111	Gene Expression Profiling, Protein Biomarkers, and Multimodal Artificial Intelligence for Prostate Cancer Management		
Original Policy Date:	March 1, 2016	**Effective Date:**	October 1, 2025
Section:	2.0 Medicine	**Page:**	Page 1 of 73

Policy Statement

I. Use of gene expression analysis, protein biomarkers, and multimodal artificial intelligence (MMAI) to guide management of prostate cancer is considered **investigational** in all situations.

Note: For individuals enrolled in health plans subject to the Biomarker Testing Law (Health & Safety Code Section 1367.667 and the Insurance Code Section 10123.209), Centers for Medicare & Medicaid Services (CMS) Local Coverage Determination (LCD) may also apply. Please refer to the Medicare National and Local Coverage section of this policy and to MolDX: Prostate Cancer Genomic Classifier Assay for Men with Localized Disease for reference.

NOTE: Refer to Appendix A to see the policy statement changes (if any) from the previous version.

Policy Guidelines

Plans may need to alter local coverage medical policy to conform to state law regarding coverage of biomarker testing.

Coding

See the Codes table for details.

Description

Description

Gene expression profile analysis and protein biomarkers have been proposed as a means to risk-stratify individuals with prostate cancer to guide treatment decisions. These tests are intended to be used either on prostate needle biopsy tissue to guide management decisions for active surveillance or therapeutic intervention, to guide radiotherapy use after radical prostatectomy (RP), or to guide medication selection after progression in metastatic castration-resistant prostate cancer.

Summary of Evidence

Initial Management Decision: Active Surveillance versus Therapeutic Intervention

For individuals who have clinically localized untreated prostate cancer who receive Prolaris, the evidence includes retrospective cohort studies of clinical validity using archived samples in patients of mixed risk categories. Relevant outcomes include overall survival (OS), disease-specific survival, quality of life (QOL), and treatment-related morbidity. For the low-risk group, the Prostate Testing for Cancer and Treatment trial showed 99% 10-year disease-specific survival in mostly low-risk patients receiving active surveillance. The low mortality rate estimated with tight precision makes it unlikely that a test intended to identify a subgroup of low-risk men with a net benefit from immediate treatment instead of active surveillance would find such a group. For the intermediate-risk group, the evidence of improved clinical validity or prognostic accuracy for prostate cancer death using Prolaris Cell Cycle Progression score in patients managed conservatively after a needle biopsy has shown

some improvement in areas under the receiver operating characteristic curve over clinicopathologic risk stratification tools. There is limited indirect evidence for potential clinical utility. The evidence is insufficient to determine that the technology results in an improvement in the net health outcome.

For individuals who have clinically localized untreated prostate cancer who receive Oncotype DX Prostate, the evidence includes case-cohort and retrospective cohort studies of clinical validity using archived samples in patients of mixed risk categories, and a decision-curve analysis examining indirect evidence of clinical utility. Relevant outcomes include OS, disease-specific survival, QOL, and treatment-related morbidity. Evidence for clinical validity and potential clinical utility of Oncotype DX Prostate in patients with clinically localized prostate cancer derives from a study predicting adverse pathology after RP. The validity of using tumor pathology as a surrogate for the risk of progression and cancer-specific death is unclear. It is also unclear whether results from an RP population can be generalized to an active surveillance population. The evidence is insufficient to determine that the technology results in an improvement in the net health outcome.

For individuals who have clinically localized untreated prostate cancer who receive Decipher Biopsy, the evidence includes retrospective cohort studies of clinical validity using archived samples in intermediate- and high-risk patients and no studies of clinical utility. Relevant outcomes include OS, disease-specific survival, QOL, and treatment-related morbidity. A test designed to identify intermediate-risk men who can receive active surveillance instead of RP or radiotherapy (RT) or high-risk men who can forego androgen deprivation therapy would need to show very high negative predictive value for disease-specific mortality at 10 years and improvement in prediction compared with existing tools used to select such men. The evidence is insufficient to determine that the technology results in an improvement in the net health outcome.

For individuals who have clinically localized untreated prostate cancer who receive the ProMark protein biomarker test, the evidence includes a retrospective cohort study of clinical validity using archived samples and no studies of clinical utility. Relevant outcomes include OS, disease-specific survival, QOL, and treatment-related morbidity. Current evidence does not support improved outcomes with ProMark given that only a single clinical validity study is available. The evidence is insufficient to determine that the technology results in an improvement in the net health outcome.

For individuals who have clinically localized untreated prostate cancer who receive ArteraAI Prostate Test, the evidence includes 1 meta-analysis and 5 retrospective analyses on archived samples from randomized clinical trials on prostate cancer patients of mixed risk categories to assess clinical validity and utility. Relevant outcomes include OS, disease-specific survival, QOL, and treatment-related morbidity. Evidence for clinical validity and potential clinical utility of ArteraAI Prostate Test in patients with clinically localized prostate cancer derives from a handful of studies comparing relevant outcomes against comparators like National Comprehensive Cancer Network (NCCN) and standard clinicopathologic risk-stratification tools. Multimodal artificial intelligence (MMAI) algorithms, that form the foundation of ArteraAI, have shown they can outperform comparators at prognosticating 10-year outcomes of interest (OS, distant metastasis [DM], biochemical failure [BF], and prostate cancer-specific survival [PCSS]). Additionally, MMAI was able to demonstrate it is predictive for short-term androgen deprivation therapy (ST-ADT) and can determine if prostate cancer patients would have a better net health outcome on RT alone or RT plus ST-ADT. Limitations of these studies are synonymous with retrospective analysis, including but not limited to, clinical heterogeneity of study populations, variability in data recording, and different conditions under which measurements occurred. No study reported management changes made in response to ArteraAI Prostate Test results, but current NCCN management algorithms recommend MMAI testing with ArteraAI for prostate cancer patients with NCCN intermediate-risk scores to indicate patients that should undergo ST-ADT regardless of RT dose or type. Moreover, NCCN notes that MMAI testing with ArteraAI may provide more accurate risk stratification to enable better management of cancer patients; however, it remains unclear on how this could be used in clinical practice as specific MMAI

cutoff values have not been published. The evidence is insufficient to determine that the technology results in an improvement in the net health outcome.

Management Decision After Radical Prostatectomy

For individuals who have localized prostate cancer treated with RP who receive Prolaris, the evidence includes retrospective cohort studies of clinical validity using archived samples. Relevant outcomes include OS, disease-specific survival, QOL, and treatment-related morbidity. No direct evidence is available to support the clinical utility of Prolaris for improving net outcomes of patients with localized prostate cancer following RP. The chain of evidence is also incomplete. Decision-curve analysis did not provide convincing evidence of meaningful improvement in net benefit by incorporating the cell cycle progression (CCP) score. Evidence of improved clinical validity or prognostic accuracy for prostate cancer death using the Prolaris Cell Cycle Progression score in patients after prostatectomy has shown some improvement in areas under the receiver operating characteristic curve over clinicopathologic risk stratification tools. Although Prolaris CCP score may have an association with biochemical recurrence (BCR), disease-specific survival outcomes were reported in only 1 analysis. A larger number of disease-specific survival events and precision estimates for discrimination measures are needed. The evidence is insufficient to determine that the technology results in an improvement in the net health outcome.

For individuals who have localized prostate cancer who are treated with RP and who receive the Decipher RP prostate cancer classifier, the evidence includes a study of analytic validity, prospective and retrospective studies of clinical validity using overlapping archived samples, decision-curve analyses examining indirect evidence of clinical utility, and prospective decision-impact studies without pathology or clinical outcomes. Relevant outcomes include OS, disease-specific survival, QOL, and treatment-related morbidity. The clinical validity of the Decipher RP genomic classifier has been evaluated in samples of patients with high-risk prostate cancer undergoing different interventions following RP. Studies reported some incremental improvement in discrimination. However, it is unclear whether there is consistently improved reclassification-particularly to higher risk categories-or whether the test could be used to predict which men will benefit from RT. The evidence is insufficient to determine that the technology results in an improvement in the net health outcome.

For individuals who have localized prostate cancer treated with RP who receive ArteraAI Prostate Test, the evidence includes 2 retrospective cohort studies of clinical validity using archived samples. Relevant outcomes include OS, disease-specific survival, QOL, and treatment-related morbidity.

ArteraAI proved to be prognostic for RP-specific endpoints of BCR and adverse pathology given the statistically significant association. Disease-specific survival outcomes were reported in both studies and the evidence of clinical validity and prognostic accuracy for MMAI scores via ArteraAI testing in patients after RP demonstrated statistically improved PCSM and OS when compared to standard clinicopathologic risk stratification tools. Limitations of these studies are synonymous with retrospective analysis, including but not limited to, clinical heterogeneity of study populations, variability in data recording, and different conditions under which measurements occurred. No study reported management changes made in response to ArteraAI Prostate Test results. Overall, ArteraAI Prostate Test is validated for disease-specific outcomes for prostate cancer patients who underwent RP and can provide additional prognostic information that may guide postoperative management, but further studies are needed to determine if MMAI can be used to decide specific treatment regimens that improve health outcomes. The evidence is insufficient to determine that the technology results in an improvement in the net health outcome.

Management Decision in Castration-Resistant Prostate Cancer

For individuals who have metastatic castration-resistant prostate cancer who receive the Oncotype DX AR-V7 Nuclear Detect, the evidence includes 1 prospective cohort study, 1 retrospective cohort study of clinical validity using archived samples, and no studies of clinical utility. Relevant outcomes

include OS, disease-specific survival, QOL, and treatment-related morbidity. Current evidence does not support improved outcomes with Oncotype DX AR-V7 Nuclear Detect, given that only 2 clinical validity studies meeting inclusion criteria were available. The evidence is insufficient to determine that the technology results in an improvement in the net health outcome.

Management Decision in Castration-Sensitive Prostate Cancer

For individuals who have metastatic castration-sensitive prostate cancer (mCSPC) who receive ArteraAI Prostate Test, the evidence includes 2 retrospective cohort studies of clinical validity using archived samples. Relevant outcomes include OS, disease-specific survival, QOL, and treatment-related morbidity. MMAI was able to estimate treatment effects and determine that MMAI high-risk mCRPC patients would derive benefit from metastasis-directed therapy (MDT) when compared to observation. Limitations of these studies are synonymous with retrospective analysis, including but not limited to, clinical heterogeneity of study populations, variability in data recording, and different conditions under which measurements occurred. No study reported management changes made in response to ArteraAI Prostate Test results. Overall, ArteraAI Prostate Test is prognostic for mCSPC patients and has the potential to guide treatment management, but further studies are needed to determine if MMAI can be used to decide specific treatment regimens that improve net health outcomes. The evidence is insufficient to determine that the technology results in an improvement in the net health outcome.

Additional Information

Not applicable.

Related Policies

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

Benefit Application

Benefit determinations should be based in all cases on the applicable member health services contract language. To the extent there are conflicts between this Medical Policy and the member health services contract language, the contract language will control. Please refer to the member's contract benefits in effect at the time of service to determine coverage or non-coverage of these services as it applies to an individual member.

Some state or federal law may prohibit health plans from denying FDA-approved Healthcare Services as investigational or experimental. In these instances, Blue Shield of California may be obligated to determine if these FDA-approved Healthcare Services are Medically Necessary.

Regulatory Status

SB 496

SB 496 requires health plans licensed under the Knox-Keene Act ("Plans"), Medi-Cal managed care plans ("MCPS"), and health insurers ("Insurers") to cover biomarker testing for the diagnosis, treatment, appropriate management, or ongoing monitoring of an enrollee's disease or condition to guide treatment decisions, as prescribed. The bill does not require coverage of biomarker testing for screening purposes. Restricted or denied use of biomarker testing for these purposes is subject to state and federal grievance and appeal processes. Where biomarker testing is deemed medically necessary, Plans and Insurers must ensure that the testing is provided in a way that limits disruptions in care.

Laboratory-Developed Tests and Regulatory Oversight

Clinical laboratories may develop and validate tests in-house and market them as a laboratory service; laboratory-developed tests must meet the general regulatory standards of the Clinical Laboratory Improvement Amendments (CLIA). Prolaris® (Myriad Genetics), Oncotype DX® Prostate and Oncotype DX AR-V7 Nuclear Detect (Genomic Health), Decipher gene expression profiling test (Decipher Corp), the ProMark™ protein biomarker test (Metamark Genetics), and Artera® Prostate Test are available under the auspices of the CLIA. Laboratories that offer laboratory-developed tests must be licensed by the CLIA for high-complexity testing. To date, the U.S. Food and Drug Administration (FDA) has chosen not to require any regulatory review of these tests.

In November 2015, the FDA's Office of Public Health Strategy and Analysis published a report suggesting FDA oversight of laboratory-developed tests.[17,] The FDA argued that many tests need more FDA oversight than the regulatory requirements of the CLIA. The CLIA standards relate to laboratory operations but do not address inaccuracies or unreliability of specific tests. Prolaris is among the 20 case studies in the document cited as needing FDA oversight. The report asserted that patients are potentially receiving inappropriate prostate cancer care because there is no evidence that results from the test meaningfully improve clinical outcomes.

Rationale

Background

Prostate Cancer

Prostate cancer is the most common cancer diagnosed among men in the U. S, and the second most common cancer overall.[1,] Autopsy studies in the era before the availability of prostate-specific antigen (PSA) screening have identified incidental cancerous foci in 30% of men 50 years of age, with incidence reaching 75% at age 80 years.[2,]

Localized prostate cancers may appear very similar clinically at diagnosis.[3,] However, they often exhibit diverse risk of progression that may not be captured by clinical risk categories (e.g., D'Amico criteria) or prognostic tools based on clinical findings, including PSA titers, Gleason grade, or tumor stage.[4,5,6,7,8,] In studies of conservative management, the risk of localized disease progression based on prostate cancer-specific survival rates at 10 years may range from 15%[9,10,] to 20%[11,] to perhaps 27% at 20-year follow-up.[12,] Among older men (aged >70 years) with low-risk disease, comorbidities typically supervene as a cause of death; these men will die with prostate cancer present, rather than from cancer itself. Other very similar appearing low-risk tumors may progress unexpectedly rapidly, quickly disseminating and becoming incurable.

Risk Stratification in Newly Diagnosed Disease

In the U.S., most prostate cancers are clinically localized at diagnosis due in part to the widespread use of PSA testing. Clinicopathologic characteristics are used to stratify patients by risk based on the extent of the primary tumor (T category), nearby lymph node involvement (N category), metastasis (M category), PSA level and Gleason score. The National Comprehensive Cancer Network (NCCN) and American Urological Association risk categories for clinically localized prostate cancer are similar, derived from the D'Amico criteria and broadly include low-, intermediate-, or high-risk as follows as well as subcategories within these groups:[13,14,]

- Low: T1-T2a and Gleason score ≤6/Gleason grade group 1 and PSA level ≤10 ng/mL;
- Intermediate: T2b-T2c or Gleason score 3+4=7/Gleason grade group 2 or Gleason score 4+3=7/Gleason grade group 3 or PSA level 10-20 ng/mL;
- High: T3a or Gleason score 8/Gleason grade group 4 or Gleason score 9-10/Gleason grade group 5 or PSA level >20 ng/mL.

Risk stratification is combined with patient age, life expectancy, and treatment preferences to make initial therapy decisions.

Principles of Risk Stratification and Biomarkers for Prostate Cancer

Predictive biomarkers and risk stratification methods are the primary tools within clinical practice that may aid in the treatment of individuals with localized and advanced prostate cancer. The NCCN uses multiple categories and subgroupings to capture prognostic risk and provide a method for risk-stratification to allow standardized treatment recommendations for individuals with localized and advanced prostate cancer. [13,] These tools are separated by type and category:
Type:

- "Standard Tools: These include clinical and/or pathologic variables routinely collected to assign a patient to an NCCN category and/or subgroup. Examples include TNM stage, Grade Group, PSA, and metastatic volume of disease."
- "Clinical and Pathologic Tools: These include clinical and/or pathologic tools that are generally derived from standard tools. Examples include multivariable models or nomograms, histologic variants, and PSA kinetics."
- "Advanced Tools: These involve an additional test above what is collected to assign an NCCN category or subgroup. These may include, but are not limited to, germline or somatic tests, gene expression tests, digital histopathology-based tests, imaging, and circulating markers."

Category:

- "Prognostic: Discriminates the risk of developing an oncologic endpoint (e.g., distant metastasis). The relative benefit of a treatment (i.e., the treatment effect or hazard ratio) is generally similar across a prognostic spectrum, although the absolute benefit of an intervention may vary by risk (i.e., number needed to treat [NNT])."
    - "Prognostic biomarkers independently discriminate and are associated with a clinically meaningful endpoint above and beyond standard tools relevant to that disease setting that ultimately helps guide a therapeutic decision."
- "Predictive: Discriminates a difference in the relative benefit of a specific treatment for an oncologic endpoint."
    - "Predictive biomarkers have been demonstrated to measure a biomarker-treatment interaction that ultimately helps guide a therapeutic decision in the context of a randomized trial, specifically randomizing the treatment of interest."

Monitoring After Prostatectomy

All normal prostate tissue and tumor tissue are theoretically removed during radical prostatectomy (RP), so the serum level of PSA should be undetectable following RP. Detectable PSA post-RP indicates residual prostate tissue and presumably persistent or recurrent disease. Prostate-specific antigen is serially measured following RP to detect early disease recurrence. The National Comprehensive Cancer Network recommends monitoring serum PSA every 6 to 12 months for the first 5 years and annually thereafter.[13,] Many recurrences following RP can be successfully treated.

The American Urological Association recommends that biochemical recurrence be defined as a serum PSA of 0.2 ng/mL or higher, which is confirmed by the second determination with a PSA level of 0.2 ng/mL or higher.[15,]

Castration-Resistant Prostate Cancer

Androgen deprivation therapy (ADT) is generally the initial treatment for patients with advanced prostate cancer. Androgen deprivation therapy can produce tumor response and improve quality of life but most patients will eventually progress on ADT. Disease that progresses while the patient is on ADT is referred to as castration-resistant prostate cancer. After progression, continued ADT is generally used in conjunction with other treatments. Androgen pathways are important in the progression of castration-resistant prostate cancer. Several drugs have been developed that either inhibit enzymes involved in androgen production or inhibit the androgen receptor, such as abiraterone and enzalutamide. Taxane chemotherapy with docetaxel or cabazitaxel may also be used after progression. Immunotherapy (sipuleucel-T) or radium 223 are options for select men.

Decision Framework for Evaluating Postate Cancer Biomarkers

Simon et al Framework

Many studies have investigated individual biomarkers or combinations of biomarkers associated with prostate cancer outcomes. Determining which studies constitute sufficient evidence that the test or biomarker is likely to be clinically useful depends on attributes of the test such as its performance and the quality of the study generating the results. Simon et al (2009) have described a framework to evaluate prognostic biomarker evidence.[16,] Study designs, such as prospective clinical trials or previously conducted clinical trials with archived tumor samples, constitute stronger evidence than studies with less planned and systematic patient recruitment and data collection. Randomized trials allow the determination of treatment-biomarker interactions that may be clinically important. In some clinical scenarios, demonstration of a treatment-biomarker interaction is not critical, because the decision to withhold chemotherapy in a low-risk group (to avoid chemotherapy-related morbidity) does not require the presence of a biomarker-treatment interaction. The study must generate an absolute estimate of outcomes in the patient group of interest that would result in a change in management (e.g., withholding of chemotherapy), and the study must have sufficient precision (narrow confidence intervals). Results of the same test across studies should show the consistency of results and more than 1 study demonstrating the desired result should be available.

Simon et al (2009) have proposed that at least 2 Simon et al (2009) category B studies showing results consistent with clinical utility are necessary to demonstrate adequate evidence of a biomarker.[16,] Simon et al (2009) also proposed that while "further confirmation in a separate trial of the results gained from a category A prospective trial is always welcome, compelling results from such a trial would be considered definitive and no other validating trial would be required."[16,]

Literature Review

Evidence reviews assess whether a medical test is clinically useful. A useful test provides information to make a clinical management decision that improves the net health outcome. That is, the balance of benefits and harms is better when the test is used to manage the condition than when another test or no test is used to manage the condition.

The first step in assessing a medical test is to formulate the clinical context and purpose of the test.

The test must be technically reliable, clinically valid, and clinically useful for that purpose. Evidence reviews assess the evidence on whether a test is clinically valid and clinically useful. Technical reliability is outside the scope of these reviews, and credible information on technical reliability is available from other sources.

Initial Management Decision: Active Surveillance versus Therapeutic Intervention

The divergent behavior of localized prostate cancers creates uncertainty whether to treat immediately or follow with active surveillance.[18,19,] With active surveillance, the patient will forgo immediate therapy and continue regular monitoring until signs or symptoms of disease progression are evident, at which point curative treatment is instituted.[20,] A patient may alternatively choose potentially curative treatment upfront.[21,] Surgery (i.e., radical prostatectomy [RP]) or external-beam radiotherapy (EBRT) is most commonly used to treat patients with localized prostate cancer.

Complications most commonly reported with RP or EBRT and with the greatest variability are incontinence (0%-73%) and other genitourinary toxicities (irritative and obstructive symptoms); hematuria (typically ≤5%); gastrointestinal and bowel toxicity, including nausea and loose stools (25%-50%); proctopathy, including rectal pain and bleeding (10%-39%); and erectile dysfunction, including impotence (50%-90%).[22,] In a population-based retrospective cohort study using administrative hospital data, physician billing codes, and cancer registry data, Nam et al (2014) estimated the 5-year cumulative incidence of admission to hospital for a treatment-related complication following RP or EBRT to be 22% (95% confidence interval [CI], 21.7% to 22.7%).[23,]

In the Prostate Testing for Cancer and Treatment (ProtecT) trial (2016), active surveillance, immediate RP, and immediate EBRT for the treatment of clinically localized prostate cancer were compared in 1643 men identified through prostate-specific antigen (PSA) testing.[24,] About 90% of the participants had a PSA level less than 10 ng/mL; two-thirds were Gleason score 6 and 20% were Gleason score 7; all were clinical stage T1c or T2. The mean age was 62 years. At a median of 10-year follow-up, prostate cancer-specific survival was high and similar across the 3 treatment groups: 98.8% (95% CI, 97.4% to 99.5%) in active surveillance, 99.0% (95% CI, 97.2% to 99.6%) in the surgery group, and 99.6% (95% CI, 98.4% to 99.9%) in the radiotherapy (RT) group. Surgery and RT were associated with lower incidences of disease progression and metastases compared with active surveillance.

Approximately 55% of men in the active surveillance group had received a radical treatment by the end of follow-up. Similarly, very high prostate cancer-specific survival and metastasis-free survival outcomes were reported by large, prospective cohorts of active surveillance patients in the U. S. and Canada.[25,26,]

The Prostate Cancer Intervention versus Observation Trial (PIVOT) randomized 731 men in the U. S. with localized prostate newly diagnosed cancer to RP or observation. The patients were 40% low-risk, 34% intermediate-risk, and 21% high-risk. Results from PIVOT also concluded that RP did not prolong survival compared with observation through 12 years and 19.5 years of follow-up in the primary analyses including all risk groups.[27,28,] However, among men with intermediate-risk tumors, surgery was associated with a 31% relative reduction in all-cause mortality compared with observation (hazard ratio [HR], 0.69; 95% CI, 0.49 to 0.98; absolute risk reduction, 12.6%).

An observational study by van den Bergh et al (2012), comparing sexual function of men with low-risk prostate cancer who chose active surveillance with men who received RT or RP, found that those who chose active surveillance were more often sexually active than similar men who received RP.[29,] In a 2011 report of quality of life (QOL) for men in the Scandinavian Prostate Cancer Group Study Number 4, after a median follow-up of more than 12 years, distress caused by treatment-related side effects was reported significantly more often by men assigned to RP than by men assigned to watchful waiting.[30,]

The American Urological Association (AUA), in joint guidelines (2017), has suggested that physicians recommend active surveillance for most men with low-risk localized prostate cancer but offer RP or RT to select low-risk, localized patients who have a high probability of progression on active surveillance.[22,] The guidelines also suggested that physicians recommend RP or RT plus androgen deprivation therapy (ADT) to patients with intermediate-risk prostate cancer and that RT alone or active surveillance may also be offered to select patients with favorable intermediate-risk localized cancer.

Clinical Context and Test Purpose

In men with newly diagnosed low- or favorable intermediate risk clinically localized prostate cancer, the purpose of gene expression profiling (GEP), protein biomarker testing, and multimodal artificial intelligence (MMAI) is to inform a decision whether to undergo immediate therapy or to forgo immediate therapy and begin active surveillance. In individuals with newly diagnosed unfavorable intermediate- or high-risk clinically localized prostate cancer, the purpose of GEP, protein biomarker testing, and MMAI is to inform a decision between local therapy alone (RP or RT) and treatment intensification (RT plus ADT).

Treatment decisions differ by patient risk:For newly diagnosed patients at low-risk, GEP, protein biomarker testing, and MMAI should identify a group of patients who should receive immediate RP or RT instead of active surveillance.

For newly diagnosed patients at favorable intermediate-risk, GEP, protein biomarker testing, and MMAI should identify a group of patients who can safely forgo immediate RP or RT and be followed with active surveillance.

For newly diagnosed patients at unfavorable intermediate- or high-risk, GEP, protein biomarker testing, and MMAI should identify a group of patients who can safely forgo ADT.

The following PICO was used to select literature to inform this review.

Populations

The relevant population of interest is individuals with newly diagnosed localized prostate cancer, who have not undergone treatment for prostate cancer, and who are deciding between therapeutic intervention and active surveillance, or between single and multimodal therapy.

Interventions

Gene expression profiling refers to the analysis of messenger RNA expression levels of many genes simultaneously in a tumor specimen and protein biomarkers.[31,32,33,34,35,36,] Three GEP tests and 1 protein biomarker test are intended to stratify biologically prostate cancers diagnosed on prostate needle biopsy: Prolaris, Oncotype DX Prostate Cancer Assay, and Decipher Biopsy are GEP tests that use archived tumor specimens as the messenger RNA source, reverse-transcriptase polymerase chain reaction amplification, and the TaqMan low-density array platform. A protein biomarker test, ProMark is an automated quantitative imaging method to measure protein biomarkers by immunofluorescent staining in defined areas in intact formalin-fixed paraffin-embedded biopsy tissue to provide independent prognostic information to aid in the stratification of patients with prostate cancer to active surveillance or therapy. MMAI refers to machine learning models that have been trained on histopathology slides and clinical data to predict responses to short-term androgen deprivation therapy (ST-ADT) and provide independent prognostic information to aid in the stratification of patients with prostate cancer. One MMAI test, ArteraAI Prostate Test, has been developed for these purposes in individuals with localized prostate cancer.

Comparators

Clinicopathologic risk stratification along with age/life expectancy and patient preference are currently being used to make decisions about prostate cancer management. Clinical characteristics (e.g., stage, biopsy Gleason grade, serum PSA level) and demographic characteristics (e.g., age, life expectancy) are combined to classify men according to risk. National Comprehensive Cancer Network (NCCN) and AUA have provided treatment recommendations based on risk stratification and life expectancy.[13,37,] The Kattan et al (2003) nomogram was developed to predict the risk of indolent cancer in a low-risk population considering active surveillance.[38,] The Cancer of the Prostate Risk Assessment (CAPRA) is a pretreatment nomogram that provides risk prediction of outcomes following RP developed from a cohort of RP patients.[39,]

Outcomes

Beneficial outcomes resulting from a true test result are prolonged survival, improved QOL, and reduction in unnecessary treatment-related adverse events. Harmful outcomes resulting from a false test result are recurrence, metastases or death, and unnecessary treatments. The outcomes of interest are listed in Table 1. The primary survival outcome of interest is disease-specific survival because overall survival (OS) is very high in this group.

Table 1. Outcomes of Interest for Individuals With Newly Diagnosed, Localized Prostate Cancer

Outcomes	Details
Overall survival	10-year survival
Disease-specific survival	10-year prostate cancer-free survival; 10-year prostate cancer death rate; 10-year recurrence rate; 10-year BCR; 10-year PCSM; 10-year DM) 10-year DDM
Quality of life	See Chen et al (2014)[40,] for NCI-recommended health-related quality of life measures for localized prostate cancer

Outcomes	Details
Treatment-related morbidity	Adverse events of radiotherapy, radical prostatectomy, or androgen-deprivation therapy

BCR: biochemical recurrence; DDM: death with distant metastasis; DM, distant metastasis; NCI: National Cancer Institute; PCSM: prostate cancer-specific mortality;.

Ten-year outcomes are of interest due to the prolonged natural history of localized prostate cancer.

Study Selection Criteria

For the evaluation of clinical validity of the Prolaris, Oncotype DX Prostate, ProMark protein biomarker, ArteraAI Prostate Test, and Decipher Biopsy tests, studies that meet the following eligibility criteria were considered:

- Reported on the accuracy of the marketed version of the technology (including any algorithms used to calculate scores);
- Included a validation cohort independent of the development cohort;
- Included a suitable reference standard (10 year prostate cancer-specific survival or death rate);
- Patient/sample clinical characteristics were described;
- Patient/sample selection criteria were described.

Prolaris

Prolaris is used to quantify expression levels of 31 cell cycle progression (CCP) genes and 15 housekeeper genes to generate a CCP score. This section reviews Prolaris for initial management decisions in newly diagnosed, localized cancer.

Clinically Valid

A test must detect the presence or absence of a condition, the risk of developing a condition in the future, or treatment response (beneficial or adverse).

Three studies reporting clinical validity related to newly diagnosed men with clinically localized prostate cancer are summarized in Table 2.

Table 2. Clinical Validity Studies Assessing Prolaris for Informing Initial Management Decisions

Study	Design	Dates	Sites	N	Population
Cuzick et al (2012)[41,]	Retrospective cohort from prospective registry	1990-1996	6 U.K. registries; not screen-detected	349	Clinically localized; 66% Gleason score 6-7; 46% PSA level ≤25 ng/mL
Cuzick et al (2015)[42,]	Retrospective cohort from prospective registry	1990-2003	3 U.K. registriesa; not screen-detected	761	Clinically localized; 74% Gleason score ≤7, mean PSA level 21 ng/mL
Lin et al (2018)[43,]	Validation cohort: Subset of Cuzick et al (2015) Clinical testing cohort: Consecutive men with biopsies submitted for testing to	1990-2003 2013-2016	3 U.K. registries[a]; not screen-detected NA; manufacturer database	585 19,215	• See Cuzick et al (2015) • Median PSA level, 5.6 ng/mL (IQR, 44-7.6 ng/mL) NCCN risk: • Low, 57% • Favorable intermediate, 20%

Study	Design	Dates	Sites	N	Population
	manufacturer				• Intermediate, 17% • High, 7%

IQR: interquartile range; NA: not available; NCCN: National Comprhensive Cancer Network; PSA: prostate-specific antigen.

[a] No overlap in population with Cuzick et al (2012).Cuzick et al (2012) examined the Prolaris prognostic value for prostate cancer death in a conservatively managed needle biopsy cohort.[41,] Cell cycle expression data were read blind to all other data. Patients were identified from 6 cancer registries in Great Britain and were included if they had clinically localized prostate cancer diagnosed by needle biopsy between 1990 and 1996; were younger than 76 years at diagnosis; had a baseline PSA measurement; and were conservatively managed. Potentially eligible patients who underwent RP, died, showed evidence of metastatic disease within 6 months of diagnosis, or received hormone therapy before diagnostic biopsy were excluded. The original biopsy specimens were retrieved and centrally reviewed by a panel of expert urologic pathologists to confirm the diagnosis and, where necessary, to reassign Gleason scores.[44,] Of 776 patients diagnosed by needle biopsy and for which a sample was available to review histology, needle biopsies were retrieved for 527 (68%), 442 (84%) of which had adequate material to assay. From the 442 samples, 349 (79%) produced a CCP score and had a complete baseline and follow-up information, representing 45% of 776 patients initially identified. The median follow-up time was 11.8 years. Ninety deaths from prostate cancer occurred within 2799 person-years.The primary, unadjusted analysis found a 1-unit increase in CCP score associated with a 2-fold increase (HR=2.02) in the risk of dying from prostate cancer (see Table 3). In a multivariate model including CCP, Gleason score, and PSA level, the adjusted HR for a 1-unit increase in CCP score was 1.65. However, changes in HRs may not reflect meaningful changes in absolute risk. As is shown in Table 4, Kaplan-Meier analyses of the 10 year risk of prostate cancer death are stratified by CCP score groupings. It appears that there might be a large change in risk for scores below 2 compared with above 2, but no CIs are reported so it is impossible to draw conclusions. Measures that would suggest improved discriminatory ability (e.g., area under the curve [AUC] or reclassification) compared with an existing nomogram were not reported in Cuzick et al (2012). The authors did not provide evidence that the test could correctly reclassify men initially at high-risk to lower risk to avoid overtreatment, or conversely, correctly reclassify those initially at low-risk to high-risk to avoid undertreatment.

Cuzick et al (2012) examined the Prolaris prognostic value for prostate cancer death in a conservatively managed needle biopsy cohort.[41,] Cell cycle expression data were read blind to all other data. Patients were identified from 6 cancer registries in Great Britain and were included if they had clinically localized prostate cancer diagnosed by needle biopsy between 1990 and 1996; were younger than 76 years at diagnosis; had a baseline PSA measurement; and were conservatively managed. Potentially eligible patients who underwent RP, died, showed evidence of metastatic disease within 6 months of diagnosis, or received hormone therapy before diagnostic biopsy were excluded. The original biopsy specimens were retrieved and centrally reviewed by a panel of expert urologic pathologists to confirm the diagnosis and, where necessary, to reassign Gleason scores.[44,] Of 776 patients diagnosed by needle biopsy and for which a sample was available to review histology, needle biopsies were retrieved for 527 (68%), 442 (84%) of which had adequate material to assay.

From the 442 samples, 349 (79%) produced a CCP score and had a complete baseline and follow-up information, representing 45% of 776 patients initially identified. The median follow-up time was 11.8 years. Ninety deaths from prostate cancer occurred within 2799 person-years.

The primary, unadjusted analysis found a 1-unit increase in CCP score associated with a 2-fold increase (hazard ratio [HR]=2.02) in the risk of dying from prostate cancer (see Table 3). In a multivariate model including CCP, Gleason score, and PSA level, the adjusted HR for a 1-unit increase in CCP score was 1.65. However, changes in HRs may not reflect meaningful changes in absolute risk. As is shown in Table 4, Kaplan-Meier analyses of the 10-year risk of prostate cancer death are stratified by CCP score groupings. It appears that there might be a large change in risk for scores below 2 compared with above 2, but no CIs are reported so it is impossible to draw conclusions.

Measures that would suggest improved discriminatory ability (e.g., area under the curve [AUC] or reclassification) compared with an existing nomogram were not reported in Cuzick et al (2012). The authors did not provide evidence that the test could correctly reclassify men initially at high-risk to

lower risk to avoid overtreatment, or conversely, correctly reclassify those initially at low-risk to high-risk toavoid undertreatment.

Cuzick et al (2015) examined 3 U.K. cancer registries from 1990 to 2003 to identify men with prostate cancer who were conservatively managed following needle biopsy, with follow-up through December 2012.[42,] The authors stated that the samples did not overlap with Cuzick et al (2012). Men were excluded if they had undergone RP or RT within 6 months of diagnosis. A combination of the CCP and CAPRA scores (called the combined clinical cell cycle risk [CCR] score) was used to predict prostate cancer death. There were 989 men who fit eligibility criteria; CCP scores were calculable for 761 (77%), and combined CCP and clinical variables were available for 585 (59%). Median age at diagnosis was 70.8 years, and the median follow-up was 9.5 years. The prostate cancer mortality rate was 17% (n=100), with 29% (n=168) dying from competing causes. Higher CCP scores were associated with increased 10-year risk of prostate cancer mortality (see Table 5): 7% (CCP score <0), 15% (CCP score 0-1), 36% (CCP score 1-2), and 59% (CCP score >2). For the CCR score, the HR for 10-year prostate cancer mortality increased to 2.17 (95% CI, 1.83 to 2.57). The C statistic for the CAPRA score was 0.74; adding the CCP score increased the C statistic to 0.78 (no CIs for the C statistic were reported).

Estimates with CIs for 10 year death rates for the CCR score are provided in a figure and given in Table 5 based on digitizing the figure. Note that the predictions appear to cross 100% for CCR of about 6. Treatment changes after 6 months were documented in only part of 1 of the 3 cohorts; at 24 months, 45% of the men in this cohort had undergone RT or prostatectomy.

Lin et al (2018)[43,] validated a CCR cutoff of 0.8 using a subset of 585 conservatively managed men from the Cuzick (2015) cohort. Of the 585 men, 60 had CCR scores of 0.8 or less. Among the 284 men who were at low- or intermediate-risk by NCCN criteria, 59 had CCR scores of 0.8 or less. The text reports that the estimated 10-year prostate cancer mortality risk was 2.7% for men with CCR scores below the threshold and 3.3% (95% CI, 1.9% to 5.7%) at the threshold in the full cohort, and 2.3% below the threshold and 2.9% (95% CI, 1.3% to 6.7%) at the threshold in the cohort that excluded high-risk men. However, the Kaplan-Meier curves show an estimated prostate cancer mortality at 10 years of 0% for men with CCR of 0.8 or less in both cohorts. The Kaplan-Meier curve estimated prostate cancer mortality at 10 years for men with CCR greater than 0.8 was 20% in the full cohort and 9% in the cohort excluding high-risk men (see Table 5; precision estimates not provided).

Tward et al (2021) reported the association of the CCR score with 10-year risk of metastasis and progression in men with unfavorable intermediate- or high-risk prostate cancer. However, this study did not meet inclusion criteria for this review because it did not provide survival outcomes.[45,]

Table 3. Univariate and Multivariate Associations Between CCP and Death From Prostate Cancer

Study	N	Unadjusted	Multivariate
		HR[c] (95% CI)	HR[c] (95% CI)
Cuzick et al (2012)[41,]	349	2.02 (1.62 to 2.53)	1.65 (1.31 to 2.09)[a]
Cuzick et al (2015)[42,]	585	2.08 (1.76 to 2.46)	1.76 (1.47 to 2.14)[b]

CCP: Cell Cycle Progression; CI: confidence interval; HR: hazard ratio.
[a] Adjusted for Gleason score and prostate-specific cancer level.
[b] Adjusted for Cancer of the Prostate Risk Assessment.
[c] For a 1-unit increase in CCP.

Table 4. Kaplan-Meier Estimates of Prostate Cancer Death at 10 Years by CCP Score Groupings in the Cuzick Validation Studies[c]

	Cuzick et al (2012)[41,]		Cuzick et al (2015)[42,]	
CCP Score	*N*	*10-Year Death Rate, %[a]*	*N*	*10-Year Death Rate, %[a]*
≤0	36	19.3	194	7
0 to ≤1	133	19.8	251	15
1 to ≤2	114	21.1	110	36
2 to ≤3	50	48.2	30[b]	59
>3	16	74.9		

CCP: Cell Cycle Progression.
[a] Confidence intervals were not reported.
[b] Grouped CCP score >2.
[c] No overlap in populations with Cuzick et al (2012) and Cuzick et al (2015).

Table 5. Predicted Risk of Prostate Cancer Death at 10 Years by CCR Score Groupings

Cuzick et al (2015)[42]			Lin et al (2018)[43] Using Data From Cuzick et al (2015)[42]		
Clinical Cell Cycle Risk Score	*N*	*10-Year Death Rate (95% CI), %*[a]	*CCR Score*	*N*	*10-Year Death Rate (95% CI), %*[d]
-1	NR	1.0 (0.2 to 1.8)			
0		2.2 (0.7 to 3.4)	≤0.8	Full[b]: 60 Modified[c]: 59	Full: 0 (CI NR) Modified: 0 (CI NR)
1		4.5 (2.3 to 7.0)	>0.8	Full[b]: 525 Modified[c]: 225	Full: 19.9. (CI NR) Modified: 8.7 (CI NR)
2		9.9 (6.4 to 13.0)			
3		20.2 (16.2 to 24.1)			
4		43.1 (34.1 to 51.2)			
5		73.5 (59.4 to 92.8)			
6		109.7 (82.0 to 120.8)			

CCR: combined clinical cell cycle risk; CI: confidence interval; NR: not reported.
[a] Estimated from digitizing a figure.
[b] Including all men from the validation cohort (»52% high-risk).
[c] Excluding high-risk men in the validation cohort.
[d] Based on the Kaplan-Meier plots.

Lin et al (2018) also reported reclassification of men using the CCR score threshold (0.8) in a group of 19,215 consecutive patients whose biopsies were sent for Prolaris testing between 2013 and 2016 (see Table 6).[43] According to the table of clinicopathologic features of patients, 14,685 of the 19,215 men had a low or favorable intermediate-risk by NCCN risk classification. However, in the reclassification table and the text describing the table (see Table 6), the authors said that only 8177 of the 19,215 men met NCCN criteria for active surveillance based on low/favorable intermediate-risk clinicopathologic features. It is not clear why fewer men were categorized as meeting NCCN low/favorable intermediate criteria for the purposes of demonstrating reclassification and, therefore, it is not clear how many of the 14685 men at low- or intermediate-risk by NCCN criteria would have been reclassified using the CCR threshold.

Table 6. Reclassification of NCCN Risk Stratification Criteria for Active Surveillance With the CCR Score[a]

NCCN Risk Group	CCR Score ≤0.8	CCR Score >0.8	Total
Met NCCN criteria for active surveillance[b]	7463	714	8177[b]
Did not meet NCCN criteria for active surveillance[b]	5758	52809	11038[b]
Total	13221	5994	19215

CCR: combined clinical cell cycle risk; NCCN: National Comprehensive Cancer Network.
[a] Adapted from Lin et al (2018).[43]
[b] Sample sizes here do not match the number of men reported to be low and favorable intermediate vs. intermediate and high-risk.

The purpose of the limitations tables (see Tables 7 and 8) is to display notable limitations identified in each study.

Table 7. Study Relevance Limitations

Study	Population[a]	Intervention[b]	Comparator[c]	Outcomes[d]	Duration of Follow-Up[e]
Cuzick et al (2012)[41]	4. Not screen selected; higher risk than intended use	1. Thresholds not described		4. Reclassification not provided	

Study	Population[a]	Intervention[b]	Comparator[c]	Outcomes[d]	Duration of Follow-Up[e]
Cuzick et al (2015)[42],	4. Not screen selected; higher risk than intended use	1. Thresholds not described			
Lin et al (2018)[43],	Note. Validation cohort is from Cuzick (2015)				

The study limitations stated in this table are those notable in the current review; this is not a comprehensive gaps assessment.
[a] Population key: 1. Intended use population unclear; 2. Clinical context is unclear; 3. Study population is unclear; 4. Study population not representative of intended use.
[b] Intervention key: 1. Classification thresholds not defined; 2. Version used unclear; 3. Not intervention of interest.
[c] Comparator key: 1. Classification thresholds not defined; 2. Not compared to credible reference standard; 3. Not compared to other tests in use for same purpose.
[d] Outcomes key: 1. Study does not directly assess a key health outcome; 2. Evidence chain or decision model not explicated; 3. Key clinical validity outcomes not reported (sensitivity, specificity, and predictive values); 4. Reclassification of diagnostic or risk categories not reported; 5. Adverse events of the test not described (excluding minor discomforts and inconvenience of venipuncture or noninvasive tests).
[e] Follow-Up key: 1. Follow-up duration not sufficient with respect to natural history of disease (true-positives, true-negatives, false-positives, false-negatives cannot be determined).

Table 8. Study Design and Conduct Limitations

Study	Selection[a]	Blinding[b]	Delivery of Test[c]	Selective Reporting[d]	Data Completeness[e]	Statistical[f]
Cuzick et al (2012)[41],	1. Unclear if all men meeting criteria were included				2,3. 349 of 776 had sufficient data for inclusion	1. CIs not reported for KM estimates at 10 y for CCP
Cuzick et al (2015)[42],	1. Unclear if all men meeting criteria were included				2,3. 585 of 989 had sufficient data for inclusion	1. CIs not reported for KM estimates at 10 y for CCP
Lin et al (2018)[43],	Note. Used data from Cuzick (2015) for validation cohort					1. CIs not reported for KM estimates at 10 y for CCR

The study limitations stated in this table are those notable in the current review; this is not a comprehensive gaps assessment.
CCP: Cell Cycle Progression; CCR: combined clinical cell cycle risk; CI: confidence interval; KM: Kaplan-Meier.
[a] Selection key: 1. Selection not described; 2. Selection not random or consecutive (i.e., convenience).
[b] Blinding key: 1. Not blinded to results of reference or other comparator tests.
[c] Test Delivery key: 1. Timing of delivery of index or reference test not described; 2. Timing of index and comparator tests not same; 3. Procedure for interpreting tests not described; 4. Expertise of evaluators not described.
[d] Selective Reporting key: 1. Not registered; 2. Evidence of selective reporting; 3. Evidence of selective publication.
[e] Data Completeness key: 1. Inadequate description of indeterminate and missing samples; 2. High number of samples excluded; 3. High loss to follow-up or missing data.
[f] Statistical key: 1. Confidence intervals and/or p values not reported; 2. Comparison with other tests not reported.

In summary, Table 3 displays the association between CCP score adjusted for CAPRA; Table 4 shows the risk of death by groups of CCP score; and Table 5 shows predicted risk of death by CCR score, which is the combined CCP and CAPRA score. The CCR score is most relevant because it appears in the sample report provided by the manufacturer. Table 3 demonstrates an association between CCP and the risk of death on the relative scale but does not necessarily indicate that there is a difference in absolute risk that would be meaningful for clinical decision making. Table 4 displays the estimated absolute risk of death for the CCP score but notably absent are CIs that would help in interpretation.

However, given the data provided, several concerns arise. Even the lowest risk group shown in Cuzick et al (2012) has a 10-year death rate of 20%, which may be explained by the population characteristics (i.e., not PSA screen-selected, a third with Gleason >7 score and half with PSA level >25 ng/mL); however, a death rate of 20% is unlikely to be low enough to forgo immediate treatment.[41,]

Table 4 does not include the death rates by CCR score; however, the predicted 10-year prostate cancer death rates by CCR score were provided in a figure in Cuzick et al (2015). The predicted 10-year risk for CAPRA alone compared with CCR was provided in a dot plot in Cuzick et al (2015). The authors stated that CCR identified 11 men with a CAPRA score of 2 (indicating an estimated 10-year mortality rate of 4%) who "had a higher risk" based on CCR score. From the dot plot, it appears that for these 11 men, the 10-year mortality rate estimated by CCR score ranged from just greater than 4% to about 8%. The authors also indicated that for 31 men with CAPRA score of 3 (corresponding to the 10-year risk of death rate of 5.7%), the risk as estimated by CCR was less than 4.0% from the plot the CCR estimated risk appears to range from about 2.5% to 4% for those 31 men. It is not clear that either of these reclassifications would change the estimated mortality enough to alter treatment decisions.

Using data from Cuzick et al (2015) and a CCR cutoff of 0.8, Lin et al (2018) estimated that the 10-year death rate for men with low to favorable intermediate-risk was 0% in men with CCR score of 0.8 or less and 9% for men with CCR score greater than 0.8, but precision estimates were not provided.

Clinically Useful

A test is clinically useful if the use of the results informs management decisions that improve the net health outcome of care. The net health outcome can be improved if patients receive correct therapy, or more effective therapy, or avoid unnecessary therapy, or avoid unnecessary testing.

Direct Evidence

Direct evidence of clinical utility is provided by studies that have compared health outcomes for patients managed with and without the test. Because these are intervention studies, the preferred evidence would be from randomized controlled trials (RCTs).

BSC identified no studies that directly supported the clinical utility of Prolaris.

Chain of Evidence

Indirect evidence on clinical utility rests on clinical validity. If the evidence is insufficient to demonstrate test performance, no inferences can be made about clinical utility.

Three decision-impact studies have assessed the potential impact of Prolaris on physicians' treatment decisions in patients.[46,47,48,] The authors of these studies–Crawford et al (2014),[46,] Shore et al (2014),[47,] and Shore et al (2016)[48,] –have suggested that their findings supported the "clinical utility" of the test, based on whether the results would lead to a change in treatment. Pathology results were not reported for these studies. Given the lack of established clinical validity and no reported outcomes, it is uncertain whether any treatment changes were clinically appropriate.

In trying to construct a chain of evidence from clinical validity to clinical utility, there are several obstacles to drawing conclusions. First, as noted in the clinical validity section, it is not clear if the test provides incremental value over the CAPRA score for decision making. In the example of reclassification given by Cuzick et al (2015), 11 men with a CAPRA estimated 10-year mortality risk rate of 4% were reclassified as having higher 10-year mortality estimated by CCR score with risk ranging from just greater than 4% to about 8%, and 31 men with CAPRA 10-year mortality risk rate of 5.7% were reclassified as having lower estimated risk by CCR of about 2.5% to 4%.[42,] It is not clear that these reclassifications would change treatment decisions.

Given that the PIVOT trial supported RP for the intermediate-risk group, showing a 30% relative and 12% absolute benefit for OS, in order to be suitable for clinical decision making, the test would have to

identify a lower risk group of intermediate-risk men with very high negative predictive value (NPV) for survival with tight CIs. Because it is not clear how the Cuzick et al (2012) or Cuzick et al (2015) results would apply specifically to intermediate-risk men, it is not clear whether the test could be used to identify intermediate-risk men who can delay RP or RT.

Health Quality Ontario (2017) reported on a health technology assessment including a systematic review of the literature assessing the clinical utility of the Prolaris CCP.[49,] The literature search identified Crawford et al (2014)[46,] and Shore et al (2016).[48,] Reviewers concluded that the GRADE rating of the quality of evidence was very low and that there was no evidence on clinical outcomes of patients whose treatment was informed by CCP results.

Section Summary: Prolaris

In a cohort of men conservatively managed following needle biopsy, Cuzick et al (2012) suggested that the CCP score alone was more prognostic than either PSA level or Gleason score for tumor-specific mortality at 10-year follow-up based on HRs.[41,] Comparison with CAPRA score was not provided in Cuzick et al (2012). Cuzick et al (2015) found that discrimination improved somewhat by adding the CCP score to the CAPRA score, as reflected in the C statistic. [42,] Ten-year mortality rates based on CCP were inconsistent within Prolaris risk categories across Cuzick et al (2012) and Cuzick et al (2015). Numerical summaries of mortality rates for the CCR were provided in a figure in Cuzick (2015). The men included in the U.K. registries were not screen-selected, and a large proportion of the men in the validation studies were not low- or intermediate-risk.

No direct evidence is available to support the clinical utility of Prolaris for improving the net outcomes of patients with localized prostate cancer. The chain of evidence is also incomplete. The ProtecT trial showed 99% 10-year disease-specific survival in all 3 treatment groups: active surveillance, RT, and RP including predominately low-risk but also some intermediate-risk men. The American Urolocial Association has recommended active surveillance in low-risk men. The low mortality rate estimated with tight precision makes it unlikely that a test intended to identify a subgroup of low-risk men with a net benefit from immediate treatment instead of active surveillance would find such a group.

The PIVOT trial preplanned subgroup analysis showed a reduction in mortality for RP compared with observation for men with intermediate-risk; AUA has recommended RT or RP for such men. For intermediate-risk men, a test designed to identify men who can receive active surveillance instead of RP or RT would need to show very high NPV for disease-specific mortality at 10 years and improvement in prediction compared with existing tools used to select such men. To forgo evidence-based beneficial treatment, there would have to be a very high standard of evidence for the clinical validity of the test.

Oncotype DX Prostate

The Oncotype DX Prostate assay includes 5 reference genes and 12 cancer genes that represent 4 molecular pathways of prostate cancer oncogenesis: androgen receptor, cellular organization, stromal response, and proliferation. The assay results are combined to produce a Genomic Prostate Score (GPS), which ranges from 0 to 100. Higher GPS scores indicate more risk.

Clinically Valid

A test must detect the presence or absence of a condition, the risk of developing a condition in the future, or treatment response (beneficial or adverse).

Five studies reporting clinical validity are summarized in Table 9. One publication by Klein et al (2014) compiled results for 3 cohorts: 2 in test development including a contemporary (1997-2011) group of patients in a prostatectomy study (n=441; Cleveland Clinic database, 1987-2004) and a biopsy study (n=167; Cleveland Clinic database, 1998-2007); the third was an independent clinical validation study cohort (n=395; University of California, San Francisco [UCSF] Database, 1998-2011).[50,] A second study,

Cullen et al (2015), evaluated men with NCCN clinically very low- to intermediate-risk undergoing prostatectomy.[51,] The third study, Whalen et al (2016), evaluated men in a clinical practice setting.[52,] The study by van Den Eeden et al (2018) included men from a cancer registry[53,] and the study by Salmasi et al (2018) included men from an institutional database.[54,]

Table 9. Clinical Validity Studies Assessing Oncotype DX Prostate

Study	Design	Dates	Sites	N	Population
Klein et al (2014)[50,]	Case-cohort from prospective registry[a]	1998-2011	UCSF	395	Clinically localized; clinical stage T1/T2; PSA level ≤20 ng/mL, Gleason score ≤7; 3% African American
Cullen et al (2015)[51,]	Retrospective cohort from prospective longitudinal study	1990-2011	U.S. military centers	382	Clinically localized; clinical stage T1/T2; PSA level ≤20 ng/mL, Gleason score ≤7; 20% African American
Whalen et al (2016)[52,]	Prospective observational cohort (median follow-up, 5.2 y)	2013-2014	Mount Sinai Hospital	50	Clinically localized; clinical stage T1/T2; PSA level ≤20 ng/mL, Gleason score ≤7
Van Den Eeden et al (2018)[53,]	Retrospective cohort from registry (median follow-up, 9.8 y)	1995-2010	Kaiser Permanente Northern California	2259	Prostate cancer who underwent RP within 12 mo of diagnosis, NCCN risk: very low, 3%; low, 21%; intermediate, 67%; high, 9%; 11% African American
Salmasi et al (2018)[54,]	Retrospective cohort from institutional database	2010-2016	UCLA	1134	NCCN very low, low- or intermediate-risk prostate cancer treated with RP; 11% African American

NCCN: National Comprehensive Cancer Network; PSA: prostate-specific antigen; RP: radical prostatectomy; UCSF: University of California, San Francisco.
[a] Only the validation sample cohort is listed.[55,]

Results from the clinical validation study and prostatectomy study by Klein et al (2014) provided information on the potential clinical validity of this test.[50,] The cohorts included men with a mix of low- to low-intermediate clinical risk characteristics using NCCN or AUA criteria. The Klein (2014) clinical validation study (see Table 9) was prospectively designed, used masked review of prostatectomy pathology results, and as such met the Reporting Recommendations for Tumor Marker Prognostic Studies guidelines for biomarker validation.[56,] The prostatectomy study used a

case-cohort design to select a 1:3 ratio of recurrent to nonrecurrent patients. Favorable pathology was defined as freedom from high-grade or non-organ-confined disease. In the prostatectomy study, the ability of the GPS to stratify patients further within AUA groupings was related to the clinical recurrence-free interval in regression-to-the-mean estimated survival curves. Results of the Klein et al (2014) validation study showed that the GPS could refine the stratification of patients within specific NCCN criteria groupings, as summarized in Table 10. Proportions were estimated from a plot of GPS versus the percent likelihood of favorable pathology.[50,]

Table 10. Reclassification of Prostate Cancer Risk Categories With Oncotype DX Prostate

NCCN Risk Level	Estimated Mean Likelihood of Favorable Tumor Pathology	
	NCCN Criteria, %	*GPS + NCCN Criteria, Range, %*
Very low	84	63-91
Low	76	55-86
Intermediate	56	29-75

Adapted from the Klein et al (2014) validation study.[50,]
GPS: Genomic Prostate Score; NCCN: National Comprehensive Cancer Network.

The actual number of patients correctly or incorrectly reclassified across all 3 categories cannot be ascertained from the data provided. The results would suggest that the combination of GPS plus clinical criteria can reclassify patients on an individual basis within established clinical risk categories.

Extrapolation of this evidence to a true active surveillance population, for which the majority in the study would be otherwise eligible, is difficult because all patients had elective RP within 6 months of diagnostic biopsy.

The Klein et al (2014) prostatectomy study, although used to identify genes to include in the GPS, provided estimates of clinical recurrence rates stratified by AUA criteria[57,] compared with rates after further stratification according to the GPS from the validation study. The survival curves for clinical recurrence reached nearly 18 years based on the dates individuals in the cohort were entered into the database (1987-2004). The reclassifications are summarized in Table 11. The GPS groups are grouped by tertiles defined in the overall study. Absolute rates and precision estimates of clinical recurrence by GPS low-, intermediate-, and high-risk groups were not reported. These data would suggest the GPS can reclassify patient risk of recurrence based on a specimen obtained at biopsy. However, the findings do not necessarily reflect a clinical scenario of predicting disease progression in untreated patients under active surveillance.

Table 11. Reclassification of Prostate Cancer 10 Year Clinical Recurrence Risk With Oncotype DX Prostate

Overall 10-Year Risk (AUA Risk Level)	10-Year Risk (GPS Low-Risk Group), %	10-Year Risk (GPS Intermediate-Risk Group), %	10-Year Risk (GPS High-Risk Group), %
3.4% (low)	2.0	3.4	7.0
9.6% (intermediate)	2.8	5.1	14.3
18.2% (high)	6.2	9.2	28.6

Adapted from the Klein et al (2014) prostatectomy study.[50,]
AUA: American Urological Association; GPS: Genomic Prostate Score.

A retrospective cohort study by Cullen et al (2015) included men with NCCN-defined very low through intermediate-risk prostate cancer undergoing RP within 6 months of diagnosis.[51,] The sample was obtained from men enrolled in the Center for Prostate Disease Research longitudinal study at 2 U.S.

military medical centers. A Gleason score of 4 or 5 with the non-organ-confined disease was considered adverse pathology. Biopsies were available for 500 (57.9%) of 864 eligible patients; 382 (44.2% of eligible) with both adequate tissues for gene expression analysis and available RP pathology were included in the analysis. Selected patients were older (61.0 years vs. 59.7 years, p=.013) and had both higher Gleason scores (p<.001) and NCCN risk classification (29.8% vs 32.9% intermediate, p=.035). Median follow-up was 5.2 years and biochemical recurrence (BCR) occurred in

62 (15.4%). Estimates of 5-year BCR by GPS score are shown in Table 12. Adverse pathology was noted in 163 (34%) men. In an analysis adjusted for baseline characteristics, the GPS was associated with BCR-free survival and adverse pathology following RP (see Table 13). The GPS improved the C statistic for adverse pathology over NCCN risk alone from 0.63 to 0.72 (CIs not reported).

Comparisons with other predictors such as CAPRA or Gleason score alone were not reported. Study implications were limited by the low proportion of eligible men in the analysis and differences between excluded and included men.

Whalen et al (2016) prospectively evaluated the correlation between GPS and final pathology at RP in a clinical practice setting.[52,] Eligible men were 50 years of age and older with more than 10 years of life expectancy, PSA levels of 20 ng/mL or less, stage cT1c-cT2c newly diagnosed, untreated prostate cancer, who met NCCN classifications as very low-risk, low-risk, or low-intermediate risk. Men were enrolled from May 2013 to August 2014 at an academic medical center. Genomic Health reclassified patients' cancers as "less favorable," "consistent with," or "more favorable" than what would have been predicted by their NCCN risk group. Adverse pathology at RP was defined as any pT3 stage and primary Gleason grade of 4 or any-pattern 5. Fifty patients had RP pathology, and the reclassification results for these participants are discussed here; 21 (42%) met the definition of adverse pathology. The NCCN risk classification categorized 2 (4%) patients as very low-risk, 34 (68%) as low-risk, and 14 (28%) as a low-intermediate risk. Twenty-three (46%) patients were reclassified using GPS and the percentage with adverse pathology for the reclassification is shown in Table 14, as derived from data provided in the text. Confidence intervals were not provided.

Van Den Eeden et al (2018) reported on a retrospective study using a stratified cohort sampling design including 279 of 6184 men who were diagnosed with prostate cancer within a registry between 1995 and 2010 and underwent RP within 12 months of diagnosis, with a median follow-up of 9.8 years.[53,] Characteristics are shown in Table 9. In an analysis adjusted for NCCN risk classifications, the GPS was associated with BCR-free survival, distant metastasis (DM), and prostate cancer death following RP (see Table 13). Ten-year prostate cancer death by GPS score was displayed in a figure stratified by NCCN risk classification, which provides some information on potential reclassification.

Ten-year prostate cancer death appears to be close to zero for men who are NCCN low-risk regardless of GPS score, indicating little useful reclassification of NCCN low-risk men based on GPS. For NCCN intermediate-risk, the risk of prostate cancer death ranges from approximately 0 for a GPS of less than 40 to close to 40% for a GPS of 100. It is unclear how many men with GPS less than 40 were NCCN favorable intermediate-risk.

Salmasi et al (2018) reported on a retrospective cohort from a UCLA institutional database of men with NCCN very low-, low-, or intermediate-risk prostate cancer treated with RP between 2010 and 2016 who had undergone simultaneous 3 Tesla multiparametric magnetic resonance imaging fusion targeted and systematic biopsies within the 6-month period prior to RP (see Table 9). The association between GPS and adverse pathology is shown in Table 13. The authors also reported an AUC for a model including Gleason score, GPS, and highest Prostate Imaging Reporting and Data System score determined by magnetic resonance imaging was 0.79 (95% CI, 0.71 to 0.87). The AUC of other models had overlapping CIs; the AUC of a model with Gleason score and highest Prostate Imaging Reporting and Data System score was 0.69 (95% CI, 0.59 to 0.78); and another model including Gleason score and PSA level was 0.68 (95% CI, 0.58 to 0.78).

Table 12. Estimates of 5 Year Biochemical Recurrence With Oncotype DX Prostate

Genomic Prostate Score	N	5-Year Biochemical Recurrence (95% Confidence Interval), %[a]
10	Not reported	5.1 (2.7 to 9.1)
20		8.5 (5.8 to 13.4)
30		14.2 (10.2 to 19.0)
40		22.9 (18.0 to 28.8)
50		35.2 (27.1 to 45.4)

Genomic Prostate Score	N	5-Year Biochemical Recurrence (95% Confidence Interval), %[a]
60		53.8 (38.6 to 65.6)
70		71.8 (50.6 to 89.3)
80		87.3 (64.2 to 98.0)

Adapted from Cullen et al (2015).[51,]
[a] Estimated from digitizing a figure.

Table 13. Univariate and Multivariate Association Between GPS and Outcomes

Study	Outcome	N	Unadjusted	Multivariate
			Ratio (95% CI)	Ratio (95% CI)
Klein et al (2014)[50,] validation study	Adverse pathology	395	OR=2.1 (1.4 to 3.2)	1.9 (1.3 to 2.8)[a]
Cullen et al (2015)[51,]	BCR	392	HR=2.9 (2.0 to 4.2)	2.7 (1.8 to 3.8)[b]
	Adverse pathology	392	HR=3.2 (2.1 to 5.0)	HR=2.7 (1.8 to 4.4)[c]
Whalen et al (2016)[52,]	Adverse pathology	50	NR	OR=1.4 (NR)[d]
Van Den Eeden et al (2018)[53,]	Distant metastasis	259	HR=2.8 (1.6 to 4.6)	HR= 2.3 (1.4 to 3.9)[a]
	Prostate-cancer death	259	HR=3.2 (1.8 to 5.7)	HR=2.7 (1.5 to 4.8)[a]
	BCR	259	HR=2.5 (1.6 to 3.9)	HR=2.1 (1.4 to 3.1)[a]
Salmasi et al (2018)[54,]	Adverse pathology	134	OR=3.8 (2.1 to 7.4)	OR=2.9 (1.5 to 5.9)[e]

BCR: biochemical recurrence; CI: confidence interval; GPS: Genomic Prostate Score; HR: hazard ratio; NCCN: National Comprehensive Cancer Network; NR: not reported; OR: odds ratio.
[a] Per 20-point increase in GPS; adjusted for NCCN risk group.
[b] Per 20-point increase in GPS; adjusted for NCCN risk group and medical center.
[c] Per 20-point increase in GPS; adjusted for NCCN risk group and age.
[d] As a continuous variable, adjusted for age, prostate-specific antigen level, clinical Gleason score, and NCCN risk category.
[e] Per 20-point increase in GPS; adjusted for Gleason score, magnetic resonance imaging score, and prostate-specific antigen level.

Table 14. Risk of Adverse Pathology With Oncotype DX Prostate

Overall AP Risk, %(NCCN Risk Level)	n	AP Risk, n (%)(GPS Less Favorable Group; n=5)	AP Risk, n (%)(GPS Consistent With Group; n=29)	AP Risk, n (%)(GPS More Favorable Group; n=18)
0% (very low)	2	-	0	-
32% (low)	34	5 (100)	6 (21)	0
71% (low-intermediate)	14	-	10 (34)	0

Adapted from Whalen et al (2016).[52,]
AP: adverse pathology; GPS: Genomic Prostate Score; NCCN: National Comprehensive Cancer Network.

Systematic Reviews

Brand et al (2016) combined the Klein et al (2014) and Cullen et al (2015) studies using a patient-specific meta-analysis.[58,] The GPS was compared with the CAPRA score, NCCN risk group, and AUA risk group. Reviewers tested whether the GPS added predictive value for the likelihood of favorable pathology above the clinical risk assessment tools. The model including the GPS and CAPRA score provided the best risk discrimination; the AUC improved from 0.68 to 0.73 by adding the GPS to the CAPRA score. The AUC improved from 0.64 to 0.70 by adding the GPS to the NCCN risk group. The improvements were reported to be significant, but the CIs for AUC were not provided.

Tables 15 and 16 display notable limitations identified in each study. This information is synthesized as a summary of the body of evidence and provides the conclusions on the sufficiency of the evidence supporting the position statement.

Table 15. Study Relevance Limitations

Study	Population[a]	Intervention[b]	Comparator[c]	Outcomes[d]	Duration of Follow-Up[e]
Klein et al (2014)[50], validation study	4. All patients had RP			1. Survival outcomes not included	
Cullen et al (2015)[51],	4. All patients had RP		3. No comparison to other risk predictors	1. Survival outcomes not included	1. 10-y outcomes not provided
Whalen et al (2016)[52],	4. All patients had RP			1. Survival outcomes not included	1. 10-y outcomes not provided
Van Den Eeden et al (2018)[53],	4. All patients had RP				
Salmasi et al (2018)[54],	4. All patients had RP			1. Survival outcomes not included	1. Follow-up duration unclear

The study limitations stated in this table are those notable in the current review; this is not a comprehensive gaps assessment.
RP: radical prostatectomy.
[a] Population key: 1. Intended use population unclear; 2. Clinical context is unclear; 3. Study population is unclear; 4. Study population not representative of intended use.
[b] Intervention key: 1. Classification thresholds not defined; 2. Version used unclear; 3. Not intervention of interest.
[c] Comparator key: 1. Classification thresholds not defined; 2. Not compared to credible reference standard; 3. Not compared to other tests in use for same purpose.
[d] Outcomes key: 1. Study does not directly assess a key health outcome; 2. Evidence chain or decision model not explicated; 3. Key clinical validity outcomes not reported (sensitivity, specificity, and predictive values); 4. Reclassification of diagnostic or risk categories not reported; 5. Adverse events of the test not described (excluding minor discomforts and inconvenience of venipuncture or noninvasive tests).
[e] Follow-Up key: 1. Follow-up duration not sufficient with respect to natural history of disease (true-positives, true-negatives, false-positives, false-negatives cannot be determined).

Table 16. Study Design and Conduct Limitations

Study	Selection[a]	Blinding[b]	Delivery of Test[c]	Selective Reporting[d]	Data Completeness[e]	Statistical[f]
Klein et al (2014)[50], validation study						1. CIs for reclassification not provided
Cullen et al (2015)[51],						1. CIs for AUC and reclassification not provided
Whalen et al (2016)[52],						1. CIs for reclassification not provided
Van Den Eeden et al (2018)[53],						
Salmasi et al (2018)[54],						

The study limitations stated in this table are those notable in the current review; this is not a comprehensive gaps assessment.
AUC: area under the curve; CI: confidence interval.
[a] Selection key: 1. Selection not described; 2. Selection not random or consecutive (i.e., convenience).
[b] Blinding key: 1. Not blinded to results of reference or other comparator tests.
[c] Test Delivery key: 1. Timing of delivery of index or reference test not described; 2. Timing of index and comparator tests not same; 3. Procedure for interpreting tests not described; 4. Expertise of evaluators not described.
[d] Selective Reporting key: 1. Not registered; 2. Evidence of selective reporting; 3. Evidence of selective publication.
[e] Data Completeness key: 1. Inadequate description of indeterminate and missing samples; 2. High number of samples excluded; 3. High loss to follow-up or missing data.
[f] Statistical key: 1. Confidence intervals and/or p values not reported; 2. Comparison with other tests not reported.

Clinically Useful

A test is clinically useful if the use of the results informs management decisions that improve the net health outcome of care. The net health outcome can be improved if patients receive correct therapy, or more effective therapy, or avoid unnecessary therapy, or avoid unnecessary testing.

Direct Evidence

Direct evidence of clinical utility is provided by studies that have compared health outcomes for patients managed with and without the test. Because these are intervention studies, the preferred evidence would be from RCTs.

BSC did not identify any studies that directly supported the clinical utility of Oncotype DX Prostate.

Chain of Evidence

Indirect evidence on clinical utility rests on clinical validity. If the evidence is insufficient to demonstrate test performance, no inferences can be made about clinical utility.

Decision-impact studies have assessed the potential impact of Oncotype DX Prostate on physicians' and patients' treatment decisions.[59,60,61,] As with the previously evaluated test, given the lack of established clinical validity and no reported outcomes, it is uncertain whether any treatment changes were clinically appropriate. With the exception of Carbunaru et al (2023), other decision-impact studies have indicated that men classified as low-risk by guidelines criteria, and thus meeting guidelines criteria for active surveillance, are more likely to receive active surveillance if they are *tested* with the Oncotype DX Prostate test.[60,62,63,64,] These arguments would suggest that the test may be a useful behavioral modifier. However, a comparison with educational or quality improvement initiatives designed to improve the uptake of active surveillance in low-risk men has not been provided. This is important to consider, since Carbunaru et al. (2023) found that higher GPS scores seemed to shift urologists' preferences from active surveillance to active treatment, but lower scores did not frequently shift preferences from active treatment back to active surveillance.

Furthermore, authors' noted that there were times when the urologists' treatment preferences did not align with NCCN recommendations for that patient's risk group (e.g., active surveillance in low-risk men).

Klein et al (2014)[50,] reported a decision-curve analysis[65,] that they proposed reflects the clinical utility of Oncotype DX Prostate. In this analysis, they compared the predictive impact of the GPS plus the CAPRA validated tool[66,] with the CAPRA score alone on the net benefit for the outcomes of patients with high-grade disease (Gleason score >4+3), high-stage disease, and combined high-grade and high-stage disease. They reported that, over a range of threshold probabilities for implementing treatment, "...incorporation of the GPS would be expected to lead to fewer treatments of patients who have favorable pathology at prostatectomy without increasing the number of patients with adverse pathology left untreated." For example, at a threshold risk of 40% (e.g., a man weighing the harms of prostatectomy vs. the benefit of active surveillance at 4:6), the test could identify 2 per 100 men with a high-grade or high-stage disease at a fixed false-positive rate, compared with using the CAPRA score alone. Thus, an individual patient could use the findings to assess his balance of benefits and harms (net benefit) when weighing the choice to proceed immediately to curative RP with its attendant adverse sequelae, or deciding to enter an active surveillance program. The latter would have an immediate benefit realized by forgoing RP but might be associated with greater downstream risks of disease progression and subsequent therapies. However, no CIs were presented for the decision-curve analysis.

Section Summary: Oncotype DX Prostate

The evidence from 5 studies on clinical validity for Oncotype DX Prostate has suggested the GPS can reclassify a patient's risk of recurrence or risk of adverse pathology at RP based on a biopsy

specimen.[50,51,52,] One study provided a figure with data on the reclassification of disease-specific survival using NCCN and GPS.[53,] Ten-year prostate cancer death appears to be close to zero for men who are NCCN low-risk regardless of GPS score, indicating little useful reclassification of NCCN low-risk men based on GPS. For NCCN intermediate-risk, the risk of prostate cancer death ranges from approximately 0 for a GPS of less than 40 to close to 40% for a GPS of 100. It is unclear how many of the men with a GPS less than 40 were NCCN favorable intermediate-risk. Moreover, generalizing this evidence to a true active surveillance population, for which most in the study would be otherwise eligible, is difficult because all patients had elective RP. Thus, the findings do not reflect a clinical scenario of predicting the risk of 10 year disease-specific survival in untreated patients under active surveillance. Some publications also lacked precision estimates for important variables such as risk estimates for recurrence or AUC estimates.

No direct evidence of clinical utility was found. The chain of evidence is also incomplete. Klein et al (2014) decision-curve analyses have suggested the potential for the combined GPS and CAPRA score data to help patients make decisions based on relative risks associated with immediate treatment or deferred treatment (i.e., active surveillance). This would reflect the clinical utility of the test. However, it is difficult to ascribe possible clinical utility of Oncotype DX Prostate in active surveillance because all patients regardless of clinical criteria elected RP within 6 months of diagnostic biopsy. Moreover, the validity of using tumor pathology as a surrogate for cancer-specific death is unclear. Reports from validation studies lack precision estimates for important variables such as risk estimates for recurrence.

The ProtecT trial showed 99% 10-year disease-specific survival in all 3 treatment groups: active surveillance, RT, and RP, including predominately low-risk but also some intermediate-risk men. AUA has recommended active surveillance in low-risk men. The low mortality rate estimated with tight precision makes it unlikely that a test intended to identify a subgroup of low-risk men with a net benefit from treatment instead of active surveillance would find such a group.

The PIVOT trial preplanned subgroup analysis showed a reduction in mortality for RP compared with observation for men at intermediate-risk; AUA has recommended RT or RP for such men. For intermediate-risk men, a test designed to identify men who can receive active surveillance instead of RP or RT would need to show very high NPV for disease-specific mortality at 10 years and improvement in prediction compared with existing tools used to select such men. For these men to forgo evidence-based beneficial treatment, there would have to be a very high standard of evidence for the clinical validity of the test.

Decipher Biopsy

This section reviews Decipher for initial management decisions in men with newly diagnosed, localized prostate cancer.

Clinically Valid

A test must detect the presence or absence of a condition, the risk of developing a condition in the future, or treatment response (beneficial or adverse).

Randomized Clinical Trial

Ross et al (2024) enrolled 227 individuals with low- or intermediate-risk localized prostate cancer and randomly assigned them to receive 1 year of enzalutamide therapy or active surveillance with a follow-up of 2 years.[67,] Transcriptional analyses were performed on tissue biopsy samples collected from patients either at screening, year 1, or year 2 with Decipher genomic classifier, androgen receptor activity (AR-A), or Prediction Analysis of Microarray 50 (PAM50). The primary endpoint of the study was to assess time to pathologic or therapeutic disease progression (pathologic disease progression defined as an increase in primary or secondary Gleason pattern of ≥1 or an increase of ≥15% in cancer-positive cores; therapeutic disease progression defined as the earliest occurrence of primary therapy for prostate cancer) in 2 distinct cohorts: 1) analytic cohort, comprised of samples

collected at screening and 2) expanded cohort, incorporated samples collected at any time during surveillance. Decipher analysis demonstrated a significant association with increased rates of pathologic or therapeutic disease progression for samples in the expanded cohort (n=114) (HR [95% CI] per 0.1; 1.17 [1.01 to 1.35]; p=.04). Additionally, Decipher scores displayed a significant association with therapeutic disease progression in the analytical cohort (n=95) (HR [95% CI] per 0.1; 1.51 [1.07 to 2.12]; p=.02) and patients with higher Decipher scores had greater benefit from enzalutamide (p=.03). Overall, this randomized prospective clinical trial suggests that Decipher gene expression profile analysis may be prognostic for low- to intermediate-risk prostate cancer. However, notable limitations include small sample size, homogeneity of samples, and lack of transcriptional analyses for all tissue samples which make it difficult to determine the clinical utility of Decipher testing.

Retrospective Studies

Three retrospective cohort studies reporting the clinical validity of Decipher Biopsy in men with newly diagnosed, localized prostate cancer are summarized in Tables 17 and 18.

Table 17. Characteristics of Clinical Validity Studies Assessing Decipher for Initial Management

Study	Study Population	Design	Comparator	Outcome	Sites	Dates
Berlin et al (2018)[68,]	Intermediate-risk PCa treated with curative-intent dose-escalated image-guided RT without neoadjuvant, concomitant or adjuvant ADT	Retrospective cohort from registry	NCCN risk groups	BCR, metastasis (5 y)	Tertiary care center, probably in Ontario	2005-2011
Nguyen et al (2017)[69,]	Treated with first-line RP or first-line RT plus ADT, had adverse pathology at surgery (defined as either preoperative PSA >20 ng/mL, stage pT3 or margin-positive, or RP grade group ≥4), the vast majority of whom had presented with intermediate- or high-risk PCa	Retrospective cohort from manufacturer database	NCCN risk groups; clinical nomogram (CAPRA)	Metastases (5 y)	7 tertiary referral clinics including Cleveland Clinic, Johns Hopkins	1987-2014
Tosoian et al (2021)[70,]	High-risk prostate cancer, defined as clinical stage T3a, Grade Group 4-5, or PSA >20 ng/ml. Patients had undergone RP or RT with ADT.	Retrospective cohort	NCCN risk groups; CAPRA	Metastases (5y)	11 centers	1995-2005

ADT: Androgen deprivation therapy; BCR: biochemical recurrence; CAPR: Cancer of the Prostate Risk Assessment l; NCCN: National Comprehensive Cancer Network; PCa: prostate cancer; PSA: prostate specific antigen; RP: radical prostatectomy; RT: radiotherapy.

The cumulative incidence of metastases at 5 years by risk group is shown in Table 18.

Table 18. Reported Prognostic Accuracies for Metastasis or PCa Mortality of Decipher as a Continuous Score and Comparators

Study	Outcome	AHR/AOR (95% CI) for Association Between GC and Outcome	AUC (95% CI)		
			GC	*Comparator*	*GC + Comparator*
Berlin (2018)[68,]	Metastasis (5 y)	2.1 (1.2 to 4.2)	0.86 (NR)	0.54 (NR)[a]	0.89 (NR)
Nguyen (2017)[69,]	Metastasis (5 y)	1.4 (1.1 to 1.8)	0.74 (0.63 to 0.83)	0.66 (0.53 to 0.77)[a]	0.74 (0.66 to 0.82)[a]
Tosoian et al (2021)[70,]	Metastasis (5 y)	1.33 per 0.1 unit (1.19 to 1.48)	NR	NCCN risk group: 0.46 (NR) CAPRA: 0.59 (NR)	GC + NCCN: 0.67 (NR) GC + CAPRA: 0.71 (NR)

AHR: adjusted hazard ratio; AOR: adjusted odds ratio; AUC: area under the curve; CI: confidence interval; GC: genomic classifier; NCCN: National Comprehensive Cancer Network; NR: not reported; PCa: prostate cancer.
[a] National Comprehensive Cancer Network risk categories.

Clinically Useful

A test is clinically useful if the use of the results informs management decisions that improve the net health outcome of care. The net health outcome can be improved if patients receive correct therapy, or more effective therapy, or avoid unnecessary therapy, or avoid unnecessary testing.

Direct Evidence

Direct evidence of clinical utility is provided by studies that have compared health outcomes for patients managed with and without the test. Because these are intervention studies, the preferred evidence would be from RCTs.

No published studies on the clinical utility of the Decipher test were identified. Zhu et a (2024) assessed patient data from the Surveillance, Epidemiology, and End Results registry to evaluate the potential impact of results from the Decipher test on treatment decisions.[71,] However, there were no reported outcomes, so it is uncertain whether any treatment changes were clinically appropriate.

Authors reported that when stratified by NCCN clinical risk stratification, testing was associated with conservative management in patients with very low/low and favorable intermediate risk. However, how these results compare to a group who did not receive testing was not reported.

Chain of Evidence

Indirect evidence on clinical utility rests on clinical validity. If the evidence is insufficient to demonstrate test performance, no inferences can be made about clinical utility.

Section Summary: Decipher Biopsy

For individuals who have clinically localized untreated prostate cancer who receive Decipher Biopsy, the evidence includes retrospective cohort studies of clinical validity using archived samples in intermediate-risk and high-risk patients and no studies of clinical utility. Relevant outcomes include OS, disease-specific survival, QOL, and treatment-related morbidity. A test designed to identify intermediate-risk men who can receive active surveillance instead of RP or RT or high-risk men who can forego ADT would need to show very high NPV for disease-specific mortality at 10 years and improvement in prediction compared with existing tools used to select such men.

ProMark Protein Biomarker Test

The ProMark assay includes 8 biomarkers that predict prostate pathology aggressiveness and lethal outcomes: *DERL1*, *PDSS2*, *pS6*, *YBX1*, *HSPA9*, *FUS*, *SMAD4*, and *CUL2*. The assay results are combined using predefined coefficients for each marker from a logistic regression model to calculate a risk score. A risk score is a continuous number between 0 and 1, which estimates the probability of "non-GS 6" pathology.

Clinically Valid

A test must detect the presence or absence of a condition, the risk of developing a condition in the future, or treatment response (beneficial or adverse).

Blume-Jensen et al (2015) reported on a study of 381 biopsies matched to prostatectomy specimens used to develop an 8-biomarker proteomic assay to predict prostate final pathology on prostatectomy specimen using risk scores.[72,]

Biomarker risk scores were defined as favorable if less than or equal to 0.33 and non-favorable if greater than 0.80, with a possible range between 0 and 1 based on false-negative and false-positive rates of 10% and 5%, respectively. The risk score generated for each patient was compared with 2 current risk stratification systems–NCCN guideline categories and the D'Amico system. Results from the study showed that, at a risk score of less than or equal to 0.33, the predictive values of the assay for favorable pathology in very low- and low-risk NCCN and low-risk D'Amico groups were 95%,

81.5%, and 87.2%, respectively, while the NCCN and D'Amico risk classification groups alone had predictive values of 80.3%, 63.8%, and 70.6%, respectively. The positive predictive value for identifying favorable disease with a risk score of less than or equal to 0.33 was 83.6% (specificity, 90%). At a risk score greater than 0.80, 77% had nonfavorable disease. Overall, 39% of the patients in the study had risk scores less than or equal to 0.33 or greater than 0.8, 81% of which were correctly identified with the 8-biomarker assay. Of the patients with intermediate-risk scores (>0.33 to ≤0.8), 58.3% had favorable disease.

The performance of the assay was evaluated in a second blinded validation study of 276 cases (see Table 19), also reported in Blume-Jensen et al (2015), to validate the assay's ability to distinguish "favorable" pathology (defined as Gleason score on prostatectomy ≤3+4 and organ-confined disease) from "non-favorable" pathology (defined as Gleason score on prostatectomy ≥4+3 or non-organ-defined disease). The second validation study separated favorable from non-favorable pathology (AUC=0.68; 95% CI, 0.61 to 0.74).

Table 19. Clinical Validity of ProMark

Study	Design[a]	Outcome	Site	N
Blume-Jensen et al (2015)[72,]	Retrospective cohort[a]	Favorable pathology at RP	Montreal, QC	276[a]

RP: radical prostatectomy.

[a] Only the validation sample cohort N.

Clinically Useful

A test is clinically useful if the use of the results informs management decisions that improve the net health outcome of care. The net health outcome can be improved if patients receive correct therapy, or more effective therapy, or avoid unnecessary therapy, or avoid unnecessary testing.

Direct Evidence

Direct evidence of clinical utility is provided by studies that have compared health outcomes for patients managed with and without the test. Because these are intervention studies, the preferred evidence would be from RCTs.

No published studies on the clinical utility of the ProMark test were identified.

Chain of Evidence

Indirect evidence on clinical utility rests on clinical validity. If the evidence is insufficient to demonstrate test performance, no inferences can be made about clinical utility.

Because the clinical utility of the ProMark test has not been established, a chain of evidence supporting the test's clinical utility cannot be constructed.

Section Summary: ProMark Protein Biomarker Test

Data are insufficient to establish the clinical validity or the clinical utility of the ProMark test.

ArteraAI Prostate Test

The ArteraAI Prostate test is an artificial intelligence biomarker test that uses digital histopathology images and clinical variables (including but not limited to, combined Gleason score, clinical T-stage, baseline PSA) to prognosticate health outcomes and predict patients who will respond to ST-ADT. This section reviews ArteraAI Prostate Test for initial management decisions in individuals with newly diagnosed, localized prostate cancer.

Clinically Valid

A test must detect the presence or absence of a condition, the risk of developing a condition in the future, or treatment response (beneficial or adverse).

One meta-analysis and 5 retrospective analyses on randomized clinical trials reporting clinical validity and utility related to newly diagnosed individuals with clinically localized prostate cancer are summarized in Table 20.

Table 20. Clinical Validity/Utility Studies Assessing ArterAI Prostate Test for Informing Initial Management Decisions

Study	Design	Randomized Clinical Trials (N)	Development/Training Cohort (n)	Validation Cohort (n)	Population of Validation Cohort	
Esteva et al (2022)[73,]	Retrospective analysis on RCTs	NRG/RTOG 9202, 9408, 9413, 0126 and 9910 (5654)	5654		931	Individuals with localized prostate cancer who received definitive RT, with or without use of ADT
Spratt et al (2023)[74,]	Retrospective analysis on RCTs	NRG/RTOG 9202, 9413, 9910, 9408 and 0126 (5727)	2024		1594	Primarily individuals with intermediate-risk prostate cancer (defined as a Gleason score of 7 or a Gleason score of ≤6 with a PSA of 10 to 20 ng/ml or clinical stage T2b and not high risk) assigned to receive RT ± 4 months of ADT
Gerrard et al (2024)[75,]	Retrospective analysis on RCTs	NRG/RTOG protocols 9202, 9408, 9413, 9910, 0126, 0415, 0521, and 9902 (7026)	Prognostic Performance Cohort: 5259 Predictive Performance Cohort: 3977	Prognostic Performance Cohort: 1767 Predictive Performance Cohort: 1509		Individuals with localized prostate cancer

Study	Design	Randomized Clinical Trials (N)	Development/Training Cohort (n)	Validation Cohort (n)	Population of Validation Cohort
Spratt et al (2024)[76,]	Meta-analysis of RCTs using a retrospective analysis	NRG/RTOG 9202, 9408, 9413, 0521 and 9910 (1088)	426	1088	Individuals with prostate cancer with ≥1 NCCN high/very high-risk factor (defined as cT3–4, Gleason 8–10, PSA >20 ng/ml, and primary Gleason pattern 5)
Ross et al (2024)[77,]	Retrospective analysis on RCTs	NRG/RTOG 9902 (397)	337	318	Individuals with localized high-risk (defined as PSA between 20 and 100 ng/ml and Gleason score ≥7 or had clinical stage ≥T2 and Gleason score ≥8) prostate cancer who received long-term AS with RT alone (AS + RT) or with adjuvant combination chemotherapy (AS + RT + CT)
Tward et al (2024)[78,]	Retrospective analysis on RCTs	NRG/RTOG 9202, 9408, 9413, 9902, 9910, 0126, 0415, and 0521 (9787)	7067	2486	Individuals with localized prostate cancer and were treated with first-line RT, with or without 4-28 months

Study	Design	Randomized Clinical Trials (N)	Development/Training Cohort (n)	Validation Cohort (n)	Population of Validation Cohort
					of ADT and with or without CT

ADT: androgen deprivation therapy; AS: androgen suppression; CT: chemotherapy; NCCN: National Comprehensive Cancer Network; PSA: prostate-specific antigen; RCT: randomized clinical trial; RT: radiotherapy.

Esteva et al (2022) developed and trained a multimodal artificial intelligence (MMAI) system with clinical data and digital histopathology using prostate biopsies from 5 phase 3 randomized clinical trials (RCT; N=5654).[73,] For each patient, the MMAI model was trained on clinical variables—including the NCCN variables (combined Gleason score, clinical T-stage, baseline PSA), as well as age, Gleason primary, and Gleason secondary—and digitized histopathology slides (median of 2 slides). In a head-to-head comparison, the MMAI system outperformed the NCCN risk-stratification tool and demonstrated superior discriminatory performances in distinct endpoints (5- and 10-year DM, 5- and 10-year biochemical failure, 10-year prostate cancer-specific survival, and 10-year overall survival (Table 21). The MMAI architecture from this seminal study became the foundation for the ArteraAI Prostate Test.

Spratt et al (2023) trained and validated an MMAI algorithm using digital pathology images from prostate tissue and clinical data from 5 phase 3 RCTs (N=5727), in which prostate cancer patients received RT with or without ADT, to predict survival outcomes via DM and prostate cancer-specific mortality (PCSM) endpoints.[74,] The MMAI model, for the overall validation cohort, demonstrated that prostate cancer patients who received RT plus ST-ADT were less likely to develop DM at 15 years when compared to patients who received RT alone (sub-distribution hazard ratio [sHR], 0.64; 95% CI, 0.45 to 0.90; p=.01). Moreover, the MMAI architecture was further developed into a binary predictive model (positive or negative) to determine whether or not patients with intermediate-risk prostate cancer would derive differential benefit from ST-ADT (Table 22). Patients who were classified as positive to receive benefit from ST-ADT therapy by the MMAI model and received RT plus ST-ADT were significantly less prone to develop DM (15-year DM estimates) than patients who were predicted positive and received RT alone (sHR, 0.34; 95% CI, 0.19 to 0.63; p<.001).

Gerrard et al (2024) set out to analytically validate the 2 most prominent MMAI algorithms developed by ArteraAI.[75,] The 2 algorithms included an algorithm with prognostic performance (Esteva et al 2022) and a second algorithm that is predictive for treatment benefit from ST-ADT (Spratt et al 2023).

The algorithms were assessed for analytical accuracy using intraclass correlation coefficient (ICC).

Both algorithms were considered to be analytically valid with reported analytical accuracy ICCs of 0.991 and 0.934 for the prognostic and predictive algorithms, respectively. Clinical validity/utility for both algorithms was reported (Tables 22 and 23) and it was concluded that the MMAI model is prognostic for DM and PCSM endpoints. Additionally, patients who were predicted as biomarker positive and received RT plus ST-ADT had significantly reduced risk of DM compared to patients that were biomarker positive and received RT alone.

Spratt et al (2024) performed a meta-analysis on the prognostic MMAI model (Esteva et al 2022) for NCCN high/very high-risk (H/HV) prostate cancer patients from 6 phase 3 RCTs.[76,] Univariate analysis was performed as a continuous score (per increment in standard deviation) and multivariate analyses were conducted to demonstrate the independent effect of the MMAI model while differentiating them from the number of NCCN H/VH risk factors (defined as cT3–4, Gleason 8–10, PSA >20 ng/ml, and primary Gleason pattern 5). Overall, the MMAI algorithm was prognostic for better health outcomes in 3 distinct endpoints (DM, PCSM, and death with distant metastasis [DDM]; Table 23).

Tward et al (2024) retrospectively analyzed a cohort of 9787 patients with localized prostate cancer from 8 different RCTs, who were treated with either RT, ADT, or chemotherapy, using MMAI models developed by ArteraAI.[78,] This study's primary goal was to compare the risk stratification of MMAI algorithms with the standard prognostic factors of NCCN risk groups through a patient-level meta-analysis, the NCCN 6-tiered risk groupings were collated into 3 groups to form the D'amico risk categories[79,], with the objective to compare the reclassification betwixt NCCN and MMAI in regard to the strength of the association with the DM endpoint. The discriminatory ability for MMAI versus NCCN was assessed using time-dependent area under the receiver operating characteristic curves (tdAUCs) over time and demonstrated that each system was able to prognosticate for DM. At 10-years, the tdAUCs were compared for the 2 risk-stratification tools, and MMAI outperformed NCCN for DM (p<.001), DDM (p<.001), and PCSM (p<.001). The MMAI models continued to establish prognostic ability for DM, DMM, and PCSM with increased scores displaying significant associations with these endpoints leading to worse health outcomes (see Table 23). The MMAI model was able to remain prognostic for DM within clinical and treatment subgroups. The overall 10-year risk for DM was similar for both NCCN and MMAI low-risk groups, however, 13% of patients were classified as low-risk for the MMAI biomarker. Furthermore, roughly 57% of NCCN intermediate- and 6% high-risk patients were reclassified as MMAI low-risk, 0% of NCCN low- and 2.5% intermediate-risk patients were reclassified as MMAI high-risk, and 46% of NCCN high-risk were reclassified as MMAI intermediate-risk, thus highlighting the inadequacy of the NCCN risk-stratification tool and the potential to minimize over- and undertreatment for individuals within MMAI risk groups.

Ross et al (2024) conducted an external validation study of the prognostic MMAI model developed by ArteraAI (Esteva et al 2022) using NCCN high-risk prostate cancer patients from a phase 3 RCT (NRG/RTOG 9902; n=318). [77,] The prognostic MMAI model outperformed clinical and pathological variables for determining DM and PCSM endpoints in a population of individuals at a high risk for disease progression (Table 23).

Table 21. Validation Results for the Subset of Patients from the Validation Set (n=931)[73,]

Clinical Outcome	NCCN AUC estimates (95% CI)	MMAI AUC (95% CI)	Differential AUC estimate (MMAI - NCCN)	Comparative test p-value
Distant Metastasis (5-year)	0.72 (0.67 to 0.78)	0.83 (0.78 to 0.88)	0.11	<.001
Distant Metastasis (10-year)	0.69 (0.64 to 0.74)	0.78 (0.73 to 0.84)	0.09	<.001
Biochemical Failure (5-year)	0.61 (0.57 to 0.64)	0.69 (0.65 to 0.73)	0.08	<.001
Biochemical Failure (10-year)	0.62 (0.58 to 0.66)	0.68 (0.63 to 0.72)	0.06	<.004
Prostate Cancer-Specific Survival (10-year)	0.67 (0.61 to 0.73)	0.77 (0.70 to 0.83)	0.10	<.001
Overall Survival (10-year)	0.57 (0.54 to 0.61)	0.65 (0.61 to 0.69)	0.08	<.001

AUC: Area under the curve; CI: confidence interval; MMAI: multimodal artificial intelligence; NCCN: National Comprehensive Cancer Network.

Table 22. Short-Term Androgen Deprivation Therapy Predictive MMAI Model for Distant Metastasis

Study	n	sHR (95% CI)	p-value
Spratt et al (2023)[74,]	543	0.34 (0.19 to 0.63)	<.001
Gerrard et al (2024)[75,]	276	0.33 (0.15 to 0.72)	.006

CI = confidence interval; sHR = sub-distribution hazard ratio

Table 23. Prognostic Performance of MMAI for Distant Metastasis and Prostate Cancer Specific Mortality

Study	Endpoints	Variable	sHR (95% CI)	p-value
Gerrard et al (2024) [75,]	DM	UVA	2.41 (2.05 to 2.82)	<.001
	PCSM	UVA	2.59 (2.17 to 3.10)	<.001
Spratt et al (2024) [76,]	DM	UVA	2.05 (1.74 to 2.43)	<.001
	PCSM	UVA	2.03 (1.73 to 2.38)	<.001
	DDM	UVA	2.04 (1.73 to 2.42)	<.001
	DM	MVA	1.90 (1.57 to 2.31)	<.001
	PCSM	MVA	2.06 (1.67 to 2.54)	<.001
	DDM	MVA	2.12 (1.72 to 2.62)	<.001
Tward et al (2024) [78,]	DM	UVA	2.66 (2.31 to 3.07)	<.001
	MMAI intermediate vs low	UVA	2.69 (1.72 to 4.20)	<.001
	MMAI high vs low	UVA	10.4 (6.88 to 15.7)	<.001
	PCSM	UVA	2.16 (1.87 to 2.50)	<.001
	MMAI intermediate vs low	UVA	1.69 (1.13 to 2.55)	.01
	MMAI high vs low	UVA	5.73 (3.93 to 8.37)	<.001
Ross et al (2024) [77,]	DM	UVA	2.33 (1.60 to 3.38)	<.001
	PCSM	UVA	3.54 (2.38 to 5.28)	<.001
	DM	MVA	NR	NR
	PCSM	MVA	NR	NR

DDM, death with distant metastases; DM: distant metastasis; MMAI: multimodal artificial intelligence; MVA = multivariate analysis; NR: not reported, PCSM: prostate cancer-specific mortality; sHR: sub-distribution hazard ratio; UVA: univariate analysis

Tables 24 and 25 display notable limitations identified in each study. This information is synthesized as a summary of the body of evidence and provides the conclusions on the sufficiency of the evidence supporting the position statement.

Table 24. Study Relevance Limitations

Study	Population[a]	Intervention[b]	Comparator[c]	Outcomes[d]	Duration of Follow-Up[e]
Esteva et al (2022)[73,]		1. Classification thresholds are not defined		4. Reclassification of risk categories not reported	
Spratt et al (2023) [74,]		1. Classification thresholds are not defined	3. No comparison to other risk predictors	4. Reclassification of risk categories not reported	
Gerrard et al (2024) [75,]		1. Classification thresholds are not defined	3. No comparison to other risk predictors	4. Reclassification of risk categories not reported	
Spratt et al (2024) [76,]		1. Classification thresholds are not defined		4. Reclassification of risk categories not reported	
Tward et al (2024) [78,]		1. Classification thresholds are not defined			
Ross et al (2024) [77,]		1. Classification thresholds are not defined		4. Reclassification of risk categories not reported	

The study limitations stated in this table are those notable in the current review; this is not a comprehensive gaps assessment.
[a] Population key: 1. Intended use population unclear; 2. Clinical context is unclear; 3. Study population is unclear; 4. Study population not representative of intended use.

[b] Intervention key: 1. Classification thresholds not defined; 2. Version used unclear; 3. Not intervention of interest.
[c] Comparator key: 1. Classification thresholds not defined; 2. Not compared to credible reference standard; 3. Not compared to other tests in use for same purpose.
[d] Outcomes key: 1. Study does not directly assess a key health outcome; 2. Evidence chain or decision model not explicated; 3. Key clinical validity outcomes not reported (sensitivity, specificity, and predictive values); 4. Reclassification of diagnostic or risk categories not reported; 5. Adverse events of the test not described (excluding minor discomforts and inconvenience of venipuncture or noninvasive tests).
[e] Follow-Up key: 1. Follow-up duration not sufficient with respect to natural history of disease (true-positives, true-negatives, false-positives, false-negatives cannot be determined).

Table 25. Study Design and Conduct Limitations

Study	Selection[a]	Blinding[b]	Delivery of Test[c]	Selective Reporting[d]	Data Completeness[e]	Statistical[f]
Esteva et al (2022)[73],	2. Retrospective analysis		2. Samples tested retrospectively			
Spratt et al (2023) [74],	2. Retrospective analysis		2. Samples tested retrospectively			
Gerrard et al (2024) [75],	2. Retrospective analysis		2. Samples tested retrospectively			
Spratt et al (2024) [76],	2. Retrospective analysis		2. Samples tested retrospectively			
Tward et al (2024) [78],	2. Retrospective analysis		2. Samples tested retrospectively			
Ross et al (2024) [77],	2. Retrospective analysis		2. Samples tested retrospectively			

The study limitations stated in this table are those notable in the current review; this is not a comprehensive gaps assessment.
AUC: area under the curve; CI: confidence interval.
[a] Selection key: 1. Selection not described; 2. Selection not random or consecutive (i.e., convenience).
[b] Blinding key: 1. Not blinded to results of reference or other comparator tests.
[c] Test Delivery key: 1. Timing of delivery of index or reference test not described; 2. Timing of index and comparator tests not same; 3. Procedure for interpreting tests not described; 4. Expertise of evaluators not described.
[d] Selective Reporting key: 1. Not registered; 2. Evidence of selective reporting; 3. Evidence of selective publication.
[e] Data Completeness key: 1. Inadequate description of indeterminate and missing samples; 2. High number of samples excluded; 3. High loss to follow-up or missing data.
[f] Statistical key: 1. Confidence intervals and/or p values not reported; 2. Comparison with other tests not reported.

Clinically Useful

A test is clinically useful if the use of the results informs management decisions that improve the net health outcome of care. The net health outcome can be improved if patients receive correct therapy, or more effective therapy, or avoid unnecessary therapy, or avoid unnecessary testing.

Direct Evidence

Direct evidence of clinical utility is provided by studies that have compared health outcomes for patients managed with and without the test. Because these are intervention studies, the preferred evidence would be from randomized controlled trials (RCTs).

BSC identified no studies that directly supported the clinical utility of ArteraAI.

Chain of Evidence

Indirect evidence on clinical utility rests on clinical validity. If the evidence is insufficient to demonstrate test performance, no inferences can be made about clinical utility.

In trying to construct a chain of evidence from clinical validity to clinical utility, there are several obstacles to drawing conclusions. First, in the example of reclassification given by Tward et al (2024), the overall 10-year risk for DM was similar for both NCCN and MMAI low-risk group with approximately 15% reclassifying as MMAI low. As previously stated, roughly 57% of NCCN intermediate- and 6% high-risk patients were reclassified as MMAI low-risk, 0% of NCCN low- and 2.5% intermediate-risk patients were reclassified as MMAI high-risk, and 46% of NCCN high-risk were reclassified as MMAI intermediate-risk. Despite these results, the utility of risk stratification is still unclear, and more studies are needed to determine how to reconcile discordant scores. Spratt et al (2023) and Gerrard et al (2024) were able to show that the predictive MMAI models for RT alone vs RT plus ST-ADT identified individuals who would and wouldn't benefit from ST-ADT using DM endpoints. However, both studies are limited by having mixed risk groups (albeit Spratt et al [2023] used mostly intermediate-risk patients) within the study population and lack of comparators make it difficult to extrapolate how the results would apply specifically to intermediate-risk men or improve on the current standard for risk stratification, currently it is unclear whether the test could be used to identify intermediate-risk men who can delay RP or RT. Furthermore, no studies have reported de-intensification strategies for high-risk populations in order to improve the net health outcome. Overall, it is still unclear that reclassifications would change treatment decisions.

Section Summary: ArteraAI Prostate Test

For individuals who have clinically localized untreated prostate cancer who receive ArteraAI Prostate Test, the evidence includes 1 meta-analysis and 5 retrospective analyses on archived samples from randomized clinical trials on prostate cancer patients of mixed risk categories to assess clinical validity and utility. Relevant outcomes include overall survival (OS), disease-specific survival, quality of life (QOL), and treatment-related morbidity. Evidence for clinical validity and potential clinical utility of ArteraAI Prostate Test in patients with clinically localized prostate cancer derives from a handful of studies comparing relevant outcomes against comparators like NCCN and standard clinicopathologic risk-stratification tools. Multimodal artificial intelligence (MMAI) algorithms, that form the foundation of ArteraAI, have shown they can outperform comparators at prognosticating 10-year outcomes of interest (OS, DM, biochemical failure [BF], and prostate cancer-specific survival [PCSS]). Additionally, MMAI was able to demonstrate it is predictive for ST-ADT and can determine if prostate cancer patients would have a better net health outcome on RT alone or RT plus ST-ADT.

Limitations of these studies are synonymous with retrospective analysis, including but not limited to, clinical heterogeneity of study populations, variability in data recording, and different conditions under which measurements occurred, etc. No study reported management changes made in response to ArteraAI Prostate Test results, but current NCCN management algorithms recommend MMAI testing with ArteraAI for prostate cancer patients with NCCN intermediate-risk scores to indicate patients that should undergo ST-ADT regardless of RT dose or type. Moreover, NCCN notes that MMAI testing with ArteraAI may provide more accurate risk stratification to enable better management of cancer patients, however, it still remains unclear on how this could be used in clinical practice as specific MMAI cutoff values have not been published. The evidence is insufficient to determine that the technology results in an improvement in the net health outcome.

Management Decision after Radical Prostatectomy

Clinical Context and Test Purpose

The purpose of gene expression profiling (GEP) and protein biomarker testing in patients who have prostate cancer and who have undergone RP is to inform management decisions.

For example, the optimal timing of RT after RP is debated. Adjuvant RT may maximize cancer control outcomes; early salvage RT (at first evidence of a rising serum PSA level) can minimize overtreatment

and still lead to acceptable oncologic outcomes.[72,] Adjuvant RT in men with pT3 or margin-positive cancer has been compared with observation in RCTs; such comparisons have shown that adjuvant RT improves the biochemical and local control rates among patients with adverse pathology at RP.[80,81,] Although the observation arms in these trials included men who received adjuvant therapy, the trials did not directly compare early salvage RT with immediate adjuvant RT because they included varying or unspecified thresholds for the initiation of salvage therapy RT.

Several observational analyses have shown conflicting conclusions whether adjuvant RT is favored over early salvage RT.[80,82,81,] RCTs comparing adjuvant with early salvage RT are underway. Guidelines have recommended that adjuvant RT be offered to patients with adverse pathologic findings at RP, and salvage RT is offered to patients with PSA or local recurrence after RP.[15,83,] However, many men treated with RT will never experience recurrence after surgery and therefore receive no benefit while experiencing harm from RT. Therefore, a test that could be used to identify men who meet criteria for adjuvant or early salvage RT but can safely receive observation instead would be useful.

Other post-RP clinical questions for which GEP or protein biomarker testing might be useful is in guiding systemic treatment (ADT and/or chemotherapy) in men receiving RT.
The second question addressed in this evidence review is: Does GEP or protein biomarker testing, compared with clinicopathologic risk stratification or when used with clinicopathologic risk stratification, improve outcomes in men following RP?
The following PICO was used to select literature to inform this review.

Populations
The relevant population of interest is individuals who have undergone RP for prostate cancer, and who are deciding on subsequent management such as adjuvant RT or no adjuvant RT. The Decipher results report says that "Decipher is intended for use in those patients who present with specific risk factors for the recurrence of prostate cancer after RP: (1) stage T2 disease with positive surgical margins, or (2) stage T3 disease, or (3) rising prostate-specific antigen (PSA) levels after initial PSA nadir."

Interventions
Prolaris, described in the previous section, is also intended to classify individuals who have undergone RP.

Decipher is a tissue-based tumor 22-biomarker GEP test intended to classify high-risk individuals who have undergone RP. The cutpoints 0.45 and 0.60 are used to categorize men using a low-, intermediate-, and high-risk genomic classifier (GC) on the Decipher test results report.

Comparators
Clinicopathologic risk stratification is currently being used to make decisions about prostate cancer management following RP. Clinical characteristics (e.g., stage, biopsy Gleason grade, serum PSA level, surgical margin, disease involvement) and demographic characteristics (e.g., age, life expectancy) are combined to classify men according to risk. As described previously, NCCN and AUA provide risk stratification guidelines.[13,15,] The Stephenson nomogram[84,85,] and Cancer of the Prostate Risk Assessment-Surgical (CAPRA-S) nomogram[86,] can be used to predict outcomes after RP.

Outcomes
Beneficial outcomes resulting from a true test result are prolonged survival, improved QOL, and reduction in unnecessary treatment-related adverse events. Harmful outcomes resulting from a false test result are recurrence, metastases or death, and unnecessary treatments. The outcomes of interest are listed in Table 26.

Table 26. Outcomes of Interest for Individuals After Radical Prostatectomy

Outcome	Details
Overall survival	10-year survival
Disease-specific survival	10-year prostate cancer-free survival; 10-year prostate cancer death rate; 10-year recurrence rate; 10-year BCR; 10-year PCSM; 10-year DM; adverse pathology
Quality of life	See Chen et al (2014)[40] for NCI-recommended health-related quality of life measures for localized prostate cancer
Treatment-related morbidity	Adverse events of radiotherapy or radical prostatectomy

BCR: biochemical recurrence; DM: distant metastasis; NCI: National Cancer Institute; PCSM: prostate cancer-specific mortality.

Ten-year outcomes are of interest due to the prolonged natural history of prostate cancer and the low number of events observed.

Prolaris

Prolaris used for initial management decisions was described in the previous section. This section reviews Prolaris for management after RP.

Clinically Valid

A test must detect the presence or absence of a condition, the risk of developing a condition in the future, or treatment response (beneficial or adverse).

Five studies reporting clinical validity in the post-RP management setting are summarized in Table 27. Four of these studies: Cuzick et al (2011),[87] Cooperberg et al (2013),[55] Bishoff et al (2014)[88] and Swanson et al (2021)[89] reported on post-RP patients. Koch et al (2016)[90] reported on post-RP patients with BCR. Freedland et al (2013)[91] reported on post-RT patients but is included in this section for completeness.

Table 27. Clinical Validity Studies Assessing Prolaris for Post-RP or Post-RT Management

Study	Design	Population	Dates	Sites	N
After prostatectomy					
Cuzick et al (2011) [87]	Retrospective cohort from prospective registry	Clinical stage T1/T2; no neoadjuvant therapy; 71% PSA level ≤10 ng/mL, 96% Gleason score ≤7	1985-1995	Scott and White Clinic	366
Cooperberg et al (2013)[55]	Retrospective cohort from prospective registry	98% PSA level ≤20 ng/mL, 95% Gleason score ≤7; no neoadjuvant or adjuvant therapy	2005-2006	Martini Clinic	283
Bishoff et al (2014)[88]	Retrospective cohort from medical records	Clinical stage T1/T2; median PSA level 5.5-7.2 ng/mL; between 91% and 94% Gleason score ≤7; between 3% and 19% with adjuvant therapy	1994-2005	Durham VAMC	176
			1997-2004	Intermountain Healthcare	123
			1994-2005	Durham VAMC	176
			1997-2004	Intermountain Healthcare	123
Koch et al (2016)[90]	Retrospective cohort from medical records	Median PSA level 6.5-11 ng/mL; 64% Gleason score ≤7; no adjuvant RT	1995-2010	Indiana University SOM	47
Swanson et al (2021)[89]	Retrospective cohort from prospective registry	46% considered to have a low risk of disease progression, 35% to have an intermediate risk, and 19% high risk according to CAPRA-S	1985-1997	Scott and White hospital (Temple, TX)	360
After external beam radiotherapy					

Study	Design	Population	Dates	Sites	N
Freedland et al (2013)[91,]	Retrospective cohort, source unclear	97% clinical stage T1/T2; Median PSA level 8 ng/mL; 88% Gleason score ≤7; 53% no concurrent hormone use; 57% African American	1991-2006	Durham VAMC	141

CAPRA-S: Cancer of the Prostate Risk Assessment Postsurgical; PSA: prostate-specific antigen; RP: radical prostatectomy; RT: radiotherapy; SOM: School of Medicine; UCSF: University of California, San Francisco; VAMC: Veterans Affairs Medical Center.

Cuzick et al (2011) examined the potential use of the Prolaris CCP test combined with a clinical score following RP, using a retrospective cohort of archived samples from a tumor registry.[87,] The study also included a cohort of men with localized prostate cancer detected from specimens obtained during transurethral resection of the prostate, which is not a population of interest here, and so is not described. Men conservatively managed after RP between 1985 and 1995 were identified from a tumor registry (n=366 with CCP scores). The primary endpoint was time to BCR, and the secondary endpoint was prostate cancer death. Myriad Genetics assessed CCP scores blindly. The median age of patients was 68 years (median follow-up, 9.4 years). Gleason scores were 7 or lower in 96%, but margins were positive in 68%. Cancers were clinically staged as T3 in 34%; following RP, 64% was judged pathologic stage T3. CCP score was associated with BCR (see Table 15). Analyses of prostate cancer deaths in the RP cohort were problematic, due to only 12 (3%) deaths. The clinical score included PSA level, stage, positive surgical margins, and Gleason score. The AUC for BCR within 5 years in the RP cohort was 0.825 for the clinical score and 0.842 for the CCR score. Although the CCP increased the AUC by 2%, whether that improvement is clinically useful is unclear because reclassification data and analysis of net benefits are lacking.

Swanson et al (2021) published a reanalysis of 360 patients from the cohort first reported in Cuzick et al (2011).[89,] After a median follow-up of 16 years, 163 (45%) of the cohort developed BCR, 41 (11%) developed metastatic disease, and 33 (9%) died from prostate cancer. The CCR score (a combination of CAPRA-S and the CCP molecular score) was prognostic of prostate cancer death, but the estimate was imprecise (HR per unit score, 3.40; 95%CI, 1.52 to 7.59). The study authors illustrated the added value of CCR for predicting disease-specific mortality by comparing predicted risk using CCR to risk predicted by a CAPRA-S-only model in a Kaplan-Meier curve; however, precision estimates were not presented.

Cooperberg et al (2013) evaluated the CCP score in an RP cohort and the incremental improvement over the CAPRA-S score for predicting BCR using a prospective-retrospective design (conforming to a PRoBE study design).[55,] A prognostic model was developed from the RP cohort described by Cuzick et al (2011).[87,] The validation cohort was obtained from patients identified from the UCSF Urologic Oncology Database. Tissue sufficient to obtain a CCP score was available for 413 men (69% of the 600 eligible samples). Both UCSF and Myriad Genetics performed statistical analyses. In the validation cohort, 95% had Gleason scores of 7 or lower, 16% of samples had positive margins, 4% had seminal vesicle invasion, and 23% had extracapsular extension. BCR occurred in 82 (19.9%) men.

The association with BCR is shown in Table 28. The AUC for BCR with CAPRA-S alone was 0.73, increasing to 0.77 for the combined CCR score.

Bishoff et al (2014) examined the prognostic ability of the CCP score in 3 cohorts: the Martini Clinic (n=283, simulated biopsies from formalin-fixed paraffin-embedded RP specimen), Durham Veterans Affairs Medical Center (n=176, diagnostic biopsies), and Intermountain Healthcare (n=123, diagnostic biopsies).[88,] The combined analysis included all 582 patients. Gleason scores were 7 or lower in 93% of men. In the combined cohorts, a unit increase in the CCP score increased the adjusted HR for BCR by 1.47 (see Table 28). Metastatic events (n=12) were too few to draw conclusions.

Koch et al (2016) evaluated whether the CCP score could discriminate between systemic disease and local recurrence in patients with BCR after RP.[90,] All 60 patients given RP as primary therapy at an academic medical center between 1995 and 2010 for whom samples were available and who had a BCR and either developed metastatic disease or received salvage EBRT with at least 2 years of follow-up were eligible for retrospective analysis. Data from 5 patients were excluded for failing to meet clinical eligibility requirements (no clarification provided) or because data were incomplete; sample blocks from 3 patients contained insufficient tumor for assay and data from 6 patients were excluded due to lack of "passing" CCP scores. Forty-seven patients were included in the analysis.

Outcomes were classified into 3 categories: (1) metastatic disease (n=22), (2) nonresponse to salvage EBRT (n=14), and (3) durable response to salvage EBRT (n=11). Analyses were performed with a binary outcome (categories 1 and 2 combined). For each 1-unit change in the CCP score, the univariate odds ratio for metastatic disease or nonresponse was 3.72 (see Table 22). Multivariate analysis was performed; however, due to the very small number of participants in the durable response group, CIs were very wide.

Table 28. Univariate and Multivariate Associations Between Prolaris CCP and Outcomes in Post-RP Clinical Validation Studies

Study	Outcomes	Median FU, y	N	Unadjusted	Multivariate
				Ratio (95% CI)	Ratio (95% CI)
Cuzick et al (2011)[87,]	BCR	9.4	366	HR=1.89 (1.54 to 2.31)	1.77 (1.40 to 2.22)[a]
	Prostate cancer death		337	HR=2.92 (2.38 to 3.57)	2.56 (1.85 to 3.53)[b]
Cooperberg et al (2013)[55,]	BCR	7	413	HR=2.1 (1.6 to 2.9)	1.7 (1.3 to 2.4)[c]
Bishoff et al (2014)[88,]	BCR	5/7[f]	582	HR=1.60 (1.35 to 1.90)	1.47 (1.23 to 1.76)[d]
Koch et al (2016)[90,]	Metastatic disease or nonresponse	9.4	47	OR=3.72 (1.29 to 10.7)	10.4 (2.05 to 90.1)[e]
Swanson et al (2021)[89,]	Prostate cancer death	16	360	HR=2.11 (1.68 to 2.65)	3.40 (1.52 to 7.59)[c]

BCR: biochemical recurrence; CCP: Cell Cycle Progression; CI: confidence interval; FU: follow-up; HR: hazard ratio; OR: odds ratio; PSA: prostate-specific antigen; RP: radical prostatectomy.
[a] Per 1-unit increase in CCP. Adjusted for PSA level, Gleason score, pathologic T stage and grade, positive surgical margins, extracapsular extension, bladder involvement, seminal vesicle involvement, positive lymph node, and age.
[b] Per 1-unit increase in CCP. Adjusted for Gleason score, PSA level, Ki67, and cancer extent.
[c] Per 1-unit increase in CCP. Adjusted for Cancer of the Prostate Risk Assessment-Surgical.
[d] Per 1-unit increase in CCP. Adjusted for PSA level, Gleason score, and adjuvant treatment.
[e] Per 1-unit increase in CCP. Adjusted for Gleason score, time from surgery to BCR, and PSA level.
[f] Not reported for 3 cohorts.

Although not a study of management post-RP, Freedland et al (2013) described the prognostic ability of the CCP score for predicting BCR in men who received primary EBRT.[92,] The retrospective data included 141 men diagnosed with prostate cancer who were treated with EBRT from 1991 to 2006, with biopsy samples and follow-up of at least 3 years. Nineteen (13%) men experienced BCR by 5 years. The univariate HR for BCR for each 1-unit increase in CCP was 2.55 (95% CI, 1.43 to 4.55). The multivariable HR for BCR associated with a 1-unit increase in CCP, including adjustment for pretreatment PSA level, Gleason, percent positive cores, and concurrent ADT, was 2.11 (95% CI, 1.05 to 4.25).

Systematic Reviews

As described in the previous Prolaris section, results of an industry-sponsored systematic review and meta-analysis were reported.[93,] Seven published studies were identified; all have been reviewed in the previous paragraphs (needle biopsy conservative management cohorts, postprostatectomy cohorts, and EBRT cohort). Including 4 validity studies[87,55,88,91,] that reported outcomes of BCR in post-RP cohorts, the pooled estimate of the HR, calculated with random-effects meta-analytic methods,

for BCR for a 1-unit increase in CCP score was 1.9 (95% CI, 1.6 to 2.3). Two studies reported outcomes for disease-specific mortality.[41,87,] Since only one of those was a post-RP study, the pooled HRs are not relevant here. There was evidence of heterogeneity in both models; reviewers did not report any variables associated with heterogeneity.

Clinically Useful

A test is clinically useful if the use of the results informs management decisions that improve the net health outcome of care. The net health outcome can be improved if patients receive correct therapy, or more effective therapy, or avoid unnecessary therapy, or avoid unnecessary testing.

Direct Evidence

Direct evidence of clinical utility is provided by studies that have compared health outcomes for patients managed with and without the test. Because these are intervention studies, the preferred evidence would be from RCTs.

BSC did not identify any studies directly supporting the clinical utility of Prolaris.

Chain of Evidence

Indirect evidence on clinical utility rests on clinical validity. If the evidence is insufficient to demonstrate test performance, no inferences can be made about clinical utility.

Decision Curves

In a decision-curve analysis, Cooperberg et al (2013) found the CAPRA-S score superior to CCP alone (as well as treat-none or treat-all strategies) in men after prostatectomy.[55,] A combined CCR predictor appeared only slightly better than CAPRA-S alone for thresholds of approximately 30% or more. For example, at a threshold of 30% (i.e., meaning a man would value the harm-to-benefit of treatment such as RT as 3:7), the CCR score would detect about 2 more men per 100 likely to experience BCR if the false-positive rate was fixed. However, the lack of CIs for the decision-curve analysis, together with the small difference, is consistent with an uncertain net benefit obtained by adding CCP to the CAPRA-S score. Also, it is not clear whether the group of patients identified as high-risk of experiencing BCR would have a net benefit from adjuvant instead of early salvage RT.

Section Summary: Prolaris

Five identified studies examined the clinical validity of Prolaris in men after RP using a BCR or systemic disease endpoint. Cuzick et al (2011) found that the CCP score offered little improvement in the AUC (2%) over clinicopathologic predictors and did not examine reclassification.[87,] Cooperberg et al (2013) found the AUC for BCR improved from 0.73 (CAPRA-S alone) to 0.77 by adding CCP score.[55,] Bishoff et al (2014)[88,] and Koch et al (2016)[90,] did not report any classification or discrimination measures. Koch et al (2016) was performed in patients who had a BCR following RP. Swanson et al (2021) published a reanalysis of 360 patients from the cohort first reported in Cuzick et al (2011).[89,] After a median follow-up of 16 years 163 (45%) of the cohort developed biochemical recurrence, 41 (11%) developed metastatic disease, and 33 (9%) died from prostate cancer. The CCR score was prognostic of prostate cancer death but the estimate was imprecise (HR per unit score, 3.40; 95% CI, 1.52 to 7.59).

No direct evidence is available to support the clinical utility of Prolaris for improving net outcomes of patients with localized prostate cancer following RP. The chain of evidence is also incomplete.

Decision-curve analysis did not provide convincing evidence of meaningful improvement in net benefit by incorporating the CCP score.

Prolaris CCP score may have an association with BCR, but disease-specific survival outcomes were reported in only one analysis. A larger number of disease-specific survival events and precision estimates for discrimination measures are needed.

Decipher Prostate Radical Prostatectomy

Decipher used for initial management decisions was described in the previous section. This section reviews Decipher for management after RP.

The Decipher test classifies as low-risk those patients who can delay or defer RT after prostatectomy, or as high-risk those who would potentially benefit from early radiation. The GC is a continuous risk score between 0 and 1, with higher risk scores indicating a greater probability of developing metastasis.

Clinically Valid

A test must detect the presence or absence of a condition, the risk of developing a condition in the future, or treatment response (beneficial or adverse).

Randomized Clinical Trial

Morgan et al (2025) enrolled 356 patients with prostate cancer who had undergone RP within 9 months, stage pT3-4N0 cancer and/or positive surgical margins, and post-RP prostate-specific antigen (PSA) of <0.1ng/ml to determine the clinical impact of Decipher genomic classifier (GC) on management decisions following RP. [94,] Patients were stratified into 2 cohorts: GC cohort (intervention) or usual care (UC) cohort with the primary endpoint being if a patient received adjuvant therapy defined as therapy preceding biochemical recurrence within 18 months following RP. There was no significant difference in the use of adjuvant therapy (8.7% vs 9.7%, p=.8) between the 2 study arms indicating that Decipher testing does not improve quality of life for post-RP. However, after modeling for adjuvant treatment by GC categories, a high GC score (>0.6) was significantly more likely to receive adjuvant therapy as opposed to lower GC (≤0.6) scores (OR, 6.9; 95% CI, 1.8 to 26; p=.005) and patients that were receiving usual care (OR, 3.5; 95% CI, 1.5 to 8.2; p=.005). However, having a low GC (<0.45) score did not indicate a lower likelihood of receiving adjuvant therapy.

Important limitations of this study include onset assumptions that were not met and resulted in less power to detect statistical differences and the potential to introduce provider bias related to GC testing. Additionally, the assumption that the use of adjuvant therapy is correct and/or meaningful needs to be further explored to ensure the association results in beneficial health outcomes for patients. Overall, these data do not provide sufficient evidence that GC testing improved the net health outcome for prostate cancer patients following RP.

The clinical validity of the Decipher test (GC) has been reported in multiple studies to predict metastasis, mortality, or BCR after RP in men with postoperative high-risk features like pathologic stage T2 with positive margins, pathologic stage T3 disease, or a rising PSA level (see Tables 29 and 30).[95,96,92,97,98,99,100,101,102,103,104,105,]

Table 29. Characteristics of Clinical Validity Studies Assessing the Decipher Genomic Classifier

Study	Study Population	Design	Comparator	Outcome	Sites	Dates
Feng et al (2021)[95,]	Recurrent disease after RP with a PSA of 0.2-4.0 ng/mL, pathologic T3 disease (tumor spread beyond the prostate) or T2 disease (tumor contained within the prostate) with a positive surgical margin and no evidence of nodal or metastatic disease	Ancillary study of specimens from an RCT	Standard clinicopathologic variables	Distant metastasis (primary), prostate cancer death and OS (secondary)	Multiple sites in US and Canada	1998-2003
Spratt et al (2018)[106,]	Clinically localized PCa after RP; serious PSA levels post-RP documented; no neoadjuvant ADT; 31% with detectable PSA 8 wk post-RP	Retrospective cohort from registry	Clinicopathological risk factors (e.g., preop PSA,	Metastases (5 y)	MD Anderson, Durham VA,	1990-2015

Study	Study Population	Design	Comparator	Outcome	Sites	Dates
			SM, RP grade group)		Thomas Jefferson	
Karnes et al (2018)[107],	Clinically localized PCa after RP; pathologic GS ≥7, pT3, pN1, or margin-positive; no neoadjuvant treatment; ≥10 y follow-up for patient alive	Retrospective cohort from registry	Clinicopathologic risk factors (e.g., preop PSA, EPE, GS); clinical nomogram (CAPRA-S)	PCa mortality (10 y)	Mayo Clinic, Johns Hopkins, Cleveland Clinic, Durham VA	1987-2010
Freedland et al (2016)[103],	Clinically localized PCa after RP; received postoperative SRT; pathologic node-negative disease; undetectable post-RP PSA; no neoadjuvant or adjuvant treatment; 32% African American	Retrospective cohort from registry	Clinicopathologic risk factors (e.g., preop PSA, EPE, GS); Clinical nomogram (Briganti, CAPRA-S)	Metastases	Durham VA, Thomas Jefferson, Mayo Clinic	1991-2010
Glass et al (2016)[104],	Clinically localized PCa after RP; preop PSA >20 ng/mL, stage pT3, margin- positive, or pathologic GS ≥8; no neoadjuvant or adjuvant treatment; 2% African American	Retrospective cohort from registry	Clinical risk factors (age at diagnosis); Clinical nomogram (CAPRA-S)	Clinical recurrence (10 y)	Kaiser Permanente Northwest	1997-2009
Ross et al (2016)[108],	Clinically localized PCa after RP; CAPRA-S score ≥3, pathologic GS ≥7, post-RP PSA nadir <0.2 ng/mL, and sufficient tissue and clinical data; no nodal disease prior to surgery; no treatment before metastasis; 8% African American	Case cohort from registry	Clinicopathologic risk factors (e.g., preop PSA, EPE, GS); clinical nomogram (CAPRA-S, Eggener)	Metastases (10 y)	Johns Hopkins	1992-2010
Ross (2016)[108],	Clinically localized PCa after RP; stage pT3 or margin-positive; achieve PSA nadir after surgery; no node-positive; no neoadjuvant treatment; no hormone-only treatment prior to metastasis; no SRT for PSA >10 ng/mL	Retrospective cohort from registry	Clinical variables (e.g., ART, MRD-SRT, SRT, no-RT); clinical nomogram (CAPRA-S)	Metastasis (10 y)	Mayo Clinic, Johns Hopkins, Durham VA, Thomas Jefferson	1990-2010
Cooperberg et al (2015)[98],	Clinically localized PCa after RP; preop PSA >20 ng/mL, stage pT3b, or pathologic GS ≥8; no neoadjuvant treatment; achieve PSA nadir after surgery	Case cohort from registry	Clinicopathologic risk factors (e.g., preop PSA, EPE, GS); clinical nomogram (CAPRA-S)	PCa mortality	CapSURE Registry	2000 - 2006
Den et al (2015)[96],	Clinically localized PCa after RP; pT3 or margin-positive disease; received post-RP RT; no neoadjuvant treatment; no lymph node invasion	Retrospective cohort from registry	Clinicopathologic risk factors (e.g., preop PSA, EPE, GS); clinical nomogram (CAPRA-S)	Metastases	Thomas Jefferson, Mayo Clinic	1990-2009

Study	Study Population	Design	Comparator	Outcome	Sites	Dates
Klein et al (2015)[92]; Klein et al (2016)[105]	Clinically localized PCa after RP; preop PSA >20 ng/mL, stage pT3, margin-positive or pathologic GS ≥8; pathologic node-negative disease; undetectable post-RP PSA; no neoadjuvant or adjuvant treatment; ≥5 y follow-up for censored patients; 8% African American	Retrospective cohort from registry	Clinicopathologic risk factors (e.g., pre-op PSA, EPE, GS); clinical nomogram (Stephenson, CAPRA-S)	Metastases (5 y, 10 y)	Cleveland Clinic	1993-2001
Den et al (2014)[97]	Clinically localized PCa after RP; pT3 or margin-positive disease; received post-RP RT; no neoadjuvant treatment; 39% BCR; 13% African American	Retrospective cohort from registry	Clinicopathologic risk factors (e.g., preop PSA, EPE, GS); clinical nomogram (Stephenson, CAPRA-S)	BCR	Thomas Jefferson	1999-2009
Ross et al (2014)[99,a] (BCR only)	Clinically localized PCa with BCR after RP; preop PSA >20 ng/mL, pathologic GS ≥8, SVI or Mayo Clinic nomogram score ≥10; no neoadjuvant treatment	Case cohort from registry	Clinicopathologic risk factors (e.g., preop PSA, EPE, GS); clinical nomogram (Stephenson, CAPRA-S)	Metastases (5 y)	Mayo Clinic	2000 - 2006
Erho et al (2013)[101],(validation)	Clinically localized PCa after RP; 32% no evidence of disease post-RP within 7 y of follow-up; 34% BCR post-RP with no clinical metastasis within 5 y of BCR; 34% clinical metastasis within 5 y of BCR	Nested case-control from registry	Clinicopathologic risk factors (e.g., preop PSA, EPE, GS)	Metastases	Mayo Clinic	1987-2001
Karnes et al (2013)[100]	Clinically localized PCa after RP; preop PSA >20 ng/mL, pathologic GS ≥8, SVI or Mayo Clinic nomogram score ≥10; no neoadjuvant treatment	Case cohort from registry	Clinicopathologic risk factors (e.g., preop PSA, EPE, GS); clinical nomogram (Stephenson)	Metastases (5 y)	Mayo Clinic	2000 - 2006

ADT: androgen deprevation therapy; ART: adjuvant radiotherapy; BCR: biochemical recurrence; CARPA-S: Cancer of the Prostate Risk Assessment Postsurgical; EPE: extraprostatic extension; GS: Gleason Score; MRD: minimal disease residual; PCa: prostate cancer; preop: preoperative; PSA: prostate-specific antigen; RP: radical prostatectomy; RT: radiotherapy; SM: surgical margins; SRT: salvage radiotherapy; SVI: seminal vesicle invasion.
[a] Appears to be subgroup with BCR from Karnes et al (2013).

Table 30. Reported Prognostic Accuracies for Metastasis or PC Mortality of Decipher as a Continuous Score and Comparators

Study	Outcome	AHR/AOR (95% CI) for Association Between GC and Outcome	AUC (95% CI)		
			GC	*Comparator*	*GC + Comparator*
Feng et al (2021)[95]	• Metastasis • PCSM • OS	• 1.17 (1.05 to 1.832) p=.006	NR	NR	NR

Study	Outcome	AHR/AOR (95% CI) for Association Between GC and Outcome	AUC (95% CI)		
		• 1.39 (1.20 to 1.63); p<.001 • 1.17 (1.06 to 1.29), p=.002			
Spratt (2018)[106,]; 95% received RT	Metastasis	NR	0.86 (0.80 to 0.94)	0.69 (0.41 to 0.89)[b]	0.83 (0.70 to 1)
Karnes (2018)[107,]	PCa mortality	1.3 (1.2 to 1.5)	0.73 (0.67 to 0.78)	0.73 (0.68 to 0.78)	0.76 (0.71 to 0.82)
Freedland (2016)[103,]	Metastasis post-RT	1.6 (1.1 to 2.1)	0.85 (0.73 to 0.88)	0.65 (0.54 to 0.81)[g]	NR
Ross (2016)[108,]	Metastasis	1.3 (1.1 to 1.5)	0.76 (0.65 to 0.84)	0.77 (0.69 to 0.85)[b]	0.87 (0.77 to 0.94)
Glass (2016)[104,]	Metastasis	1.5 (p=.011)	0.80 (0.64 to 0.92)	0.73 (0.49 to 0.95)[c]	0.84 (0.70 to 0.96)
Cooperberg (2015)[98,]	PCa mortality	1.8 (1.5 to 2.3)	0.78 (0.68 to 0.87)	0.75 (0.55 to 0.84)[b]	
Klein (2015)[92,]; Klein (2016)[105,]	Metastasis 5 y Metastasis 10 y	1.5 (1.1 to 2.1); 1.7 (1.1 to 2.8)	0.77 (0.66 to 0.87); 0.80 (0.58 to 0.95)	0.75 (0.65 to 0.84)[c]; 0.75 (0.64 to 0.87)[h]	0.79 (0.65 to 0.85) 0.88 (0.76 to 0.96)
Den (2015)[96,]	Metastasis post-RT	1.9 (p<.001)	0.78 (0.64 to 0.91)	0.70 (0.49 to 0.90)[b]	0.85 (0.79 to 0.93)
Ross (2014)[99,]	Metastasis	1.4 (p=.003)	0.82 (0.76 to 0.86)	0.70 (0.66 to 0.75)[a]	0.75 (0.69 to 0.80)
Den (2014)[97,]	Metastasis	NR	0.70 (0.49 to 0.90)[d]	0.78 (0.64 to 0.91)	0.80 (0.68 to 0.93)
Erho (2013)[101,]	Metastasis	1.4 (p<0.001)	0.75 (0.70 to 0.81)[e]	0.69 (0.60 to 0.77)[a,e]	0.74 (0.65 to 0.82)[a,e]
Karnes (2013)[100,]	Metastasis	1.5 (p<0.001)	0.79 (0.68 to 0.87)	0.64 (0.55 to 0.72)[d,f]	

AHR: adjusted hazard ratio; AOR: adjusted odds ratio; AUC: area under the curve; CI: confidence interval; GC: genomic classifier; NR: not reported; OS: overall survival; PCa: prostate cancer; PCSM: prostate-cancer specific mortality; RT: radiotherapy.
[a] Clinical classifier includes Gleason score, extracapsular extension, positive surgical margins, seminal vesicle invasion, or lymph node involvement.
[b] Cancer of the Prostate Risk Assessment-Surgical.
[c] Stephenson nomogram.
[d] Only reported vs. single clinical predictors.
[e] AUC CI obtained by digitizing figure.
[f] Gleason score.
[g] Briganti score.
[h] National Comprehensive Cancer Network risk categories.
[I] With detectable PSA post-RP.

All studies were conducted retrospectively from registry data or clinical records. The development study had a nested case-control design.[101,] The 5- and 10-year results of 1 study were published separately.[92,105,] Four were case-cohort studies and 8 used retrospective cohorts. Nine studies were supported by GenomeDx (now Decipher Corp), which offers the Decipher test. The cutpoints used to classify men into low-, intermediate- and high-risk by GC score were updated in 2016. Only 1 study (Karnes et al [2018][107,]) has reported 10-year prostate cancer-specific survival after the update in the cutpoints.

Several studies,[98,99,100,101,108,106,108,] including the test (validation) sample from the development study, examined men observed following RP and undergoing adjuvant or salvage RT. Median follow-up periods ranged from 6.4 to 16.9 years. The distributions of Gleason scores in the studies varied from

17.8% to 49.3% for those with Gleason scores of 8 or higher and from 0.4% to 15.1% for those with scores of 6 or lower. Extracapsular extension of the tumor ranged from 42.7% and 72.3% of men across studies.

Association between GC continuous score and metastasis or prostate cancer-specific mortality is shown in Table 31. The GC AUCs for predicting metastases are shown in Table 30. Among the 69 men developing metastases in Karnes et al (2013), of the 29 with Gleason scores of 7 or lower, 10 were correctly reclassified to the highest GC risk (score >0.6), but of the 40 men with Gleason scores of 8 or higher, 10 were incorrectly reclassified to the lowest GC risk group (score <0.4).[100,]

The cumulative incidence of metastases by risk group is shown in Table 32. Three studies reported prostate cancer-specific mortality; only one of which included 10 year outcomes. Precision estimates were not provided. Values in the tables below may be estimated from figures when exact values were not provided in article text or tables.

Table 31. Metastasis by GC Risk Group

Study	FU Time, y	N	Patients in Risk Group, %			Metastasis Rate, %		
			Low	*Int*	*High*	*Low*	*Int*	*High*
Feng et al (2021)[95,]	13	352	42	38	20	6.2	8.7	15.3
Spratt et al (2018)[106,]	10	561	46	28	26	0	3	23
Ross et al (2016)[108,]	5	422	57	27	16	7	10	22
Freedland et al (2016)[103,]	10	170	51	31	18	3	8	33
Glass et al (2016)[104,]	10	224	NR	NR	NR	0	3	
Ross et al (2016)[102,]	10	260	73	17	10	8	20	32
Klein et al (2015)[92,]								
Den et al (2015)[96,]	5	188	41	39	20	0	9	29
Den et al (2014)[97,]	5	139	21	38	41	0	5	17
Ross et al (2014)[99,]	5	85	NR	NR	NR	9	54	
Karnes et al (2013)[100,]	5	219	51	22	27	2	6	22

FU: follow-up; GC: genomic classifier; Int: intermediate; NR: not reported.

For prostate cancer mortality, compared with CAPRA-S, Cooperberg et al (2015) found that the GC improved reclassification somewhat-of the 19 men with CAPRA-S scores of 5 or lower, 12 were correctly reclassified to the highest GC risk, and 1 was incorrectly reclassified with a CAPRA-S score greater than 6 to low-risk; all men had CAPRA-S scores of 3 or more.[98,] Feng et al (2021) reported prostate specific mortality and OS according to GC category but did not provide data on reclassification.[95,]

Karnes et al (2018) reported the preferred outcome for this review (10-year prostate cancer-specific survival).[107,] The authors found that adding the GC to CAPRA improved the AUC from 0.73 to 0.76 with highly overlapping CIs. The 10-year cumulative incidence of prostate cancer-specific mortality by CAPRA and GC risk categories are shown in Table 33. Samples sizes and precision estimates for the cross-tabulations were not provided.

Table 32. Prostate-Cancer-Specific Mortality by Genomic Classifier Risk Group

Study	FU, y	N	Patients in Risk Group, %			Prostate Cancer Mortality		
			Low	Int	High	Low	Int	High
Karnes et al (2018)[107,]	10	561	58	17	25	12	13	45
Cooperberg et al (2015)[98,]	5	185	54	22	24	6	3	30
Feng et al (2021)[95,]	13	352	42	38	20	0.7	2.4	9.8

FU: follow-up; Int: intermediate.

Table 33. Cross-Tabulation of 10 Year Cumulative Incidence of Prostate Cancer-Specific Mortality by GC and CAPRA

CAPRA-S Risk Category	Decipher GC Risk Category, %	
	Low/Intermediate (≤0.6)	High (>0.6)
Low-risk (<6)	2.8 (CI NR)	18 (CI NR)
High-risk (≥6)	5.5 (CI NR)	30 (CI NR)

Adapted from Karnes et al (2018).[107,]
CAPRA: Cancer of the Prostate Risk Assessment; CI: confidence interval; GC: genomic classifier; NR: not reported.

Tables 34 and 35 display notable limitations identified in each study. The limitations analysis focuses on 10-year prostate cancer-specific mortality outcomes (i.e., Karnes et al [2018][107,]).

Table 34. Study Relevance Limitations

Study	Population[a]	Intervention[b]	Comparator[c]	Outcomes[d]	Duration of Follow-Up[e]
Karnes et al (2018)[107,]					

The study limitations stated in this table are those notable in the current review; this is not a comprehensive gaps assessment.
[a] Population key: 1. Intended use population unclear; 2. Clinical context is unclear; 3. Study population is unclear; 4. Study population not representative of intended use.
[b] Intervention key: 1. Classification thresholds not defined; 2. Version used unclear; 3. Not intervention of interest.
[c] Comparator key: 1. Classification thresholds not defined; 2. Not compared to credible reference standard; 3. Not compared to other tests in use for same purpose.
[d] Outcomes key: 1. Study does not directly assess a key health outcome; 2. Evidence chain or decision model not explicated; 3. Key clinical validity outcomes not reported (sensitivity, specificity, and predictive values); 4. Reclassification of diagnostic or risk categories not reported; 5. Adverse events of the test not described (excluding minor discomforts and inconvenience of venipuncture or noninvasive tests).
[e] Follow-Up key: 1. Follow-up duration not sufficient with respect to natural history of disease (true-positives, true-negatives, false-positives, false-negatives cannot be determined).

Table 35. Study Design and Conduct Limitations

Study	Selection[a]	Blinding[b]	Delivery of Test[c]	Selective Reporting[d]	Data Completeness[e]	Statistical[f]
Karnes et al (2018)[107,]	2. Unclear if included men were consecutive or random samples of those meeting eligibility criteria					1. CIs for prostate cancer-specific mortality by GC low/high-risk and reclassification not provided

The study limitations stated in this table are those notable in the current review; this is not a comprehensive gaps assessment.
CI: confidence interval.
[a] Selection key: 1. Selection not described; 2. Selection not random or consecutive (i.e., convenience).
[b] Blinding key: 1. Not blinded to results of reference or other comparator tests.
[c] Test Delivery key: 1. Timing of delivery of index or reference test not described; 2. Timing of index and comparator tests not same; 3. Procedure for interpreting tests not described; 4. Expertise of evaluators not described.
[d] Selective Reporting key: 1. Not registered; 2. Evidence of selective reporting; 3. Evidence of selective publication.
[e] Data Completeness key: 1. Inadequate description of indeterminate and missing samples; 2. High number of samples excluded; 3. High loss to follow-up or missing data.
[f] Statistical key: 1. Confidence intervals and/or p values not reported; 2. Comparison with other tests not reported.

Clinically Useful

A test is clinically useful if the use of the results informs management decisions that improve the net health outcome of care. The net health outcome can be improved if patients receive correct therapy, or more effective therapy, or avoid unnecessary therapy, or avoid unnecessary testing.

Direct Evidence

Direct evidence of clinical utility is provided by studies that have compared health outcomes for patients managed with and without the test. Because these are intervention studies, the preferred evidence would be from RCTs.

No studies reporting direct evidence were identified.

Chain of Evidence

Indirect evidence on clinical utility rests on clinical validity. If the evidence is insufficient to demonstrate test performance, no inferences can be made about clinical utility.

Decision Curves

Studies have included decision curves comparing the net benefit of different strategies using metastases or survival as the outcome (see Table 36).[96,92,98,99,100,102,107,109,107,] In observational and RT samples from Karnes et al (2013)[100,]and Ross et al (2014),[99,] using a 15% to 25% range of thresholds for decision making (i.e., suspected probability of developing metastases) would be expected to identify correctly as few as no men or as many as 4 per 100 likely to experience metastases. This range of thresholds assumes several things: it assumes those making the decisions are relying on the GC result for adjuvant RT decisions, compared with treating based on the best comparator test, and it assumes no increase in false-positives. No CIs were provided for the net benefit estimates and uncertainty cannot be evaluated. In the 2 observation-only samples, although the GC improved the net benefit over a "treat none" strategy over 15% to 25% thresholds, it appeared to offer little over the comparator test (e.g., about 1 additional patient would be likely to experience metastases without an increase in false-positives).[92,102,]In Ross et al (2014), the net benefit for CAPRA-S score exceeded that of the GC, with the net benefit of the GC plus CAPRA-S score being slightly better than the CAPRA-S score alone.[102,] Finally, among men undergoing RT, decision curves suggested that the test would identify 3 or 4 men developing metastases per 100 tested at a fixed false-positive rate. Lobo et al (2015)[109,] reported an individualized decision analysis comparing the GC with "usual care" using data from the cohorts in Karnes et al (2013) and Den et al (2014). The usual care probabilities of receiving each treatment were derived from the published literature. A 6% threshold for the GC score was used for GC-based treatment. Using the cohort from Karnes et al (2013), the estimated 10-year probability of metastasis or death was 0.32 (95% CI, 0.32 to 0.33) for usual care compared with 0.31 (95% CI, 0.30 to 0.32) for GC-based treatment. In the cohort from Den et al (2014), the estimated 10-year probability of metastasis or death was 0.28 (95% CI, 0.27 to 0.29) for usual care compared with 0.26 (95% CI, 0.25 to 0.27) for GC-based treatment.

Table 36. Reported Net Benefit of the Decipher Classifier versus Comparators

Study	Outcome	Range of Net Benefit versus	
		Treat None	Best Comparator
Spratt et al (2018)[106,]	Metastasis	-0.003 to 0.002	NR
Karnes et al (2018)[107,]	PCa mortality	0.06 to 0.09	0.045 to 0.095
Ross et al (2016)[108,]	Metastasis	0.045 to 0.075	0.09 to 0.12
Freedland (2016)[103,]	Metastasis	0.01 to 0.045	0 to 0.02
Lobo et al (2015)[109,] with Karnes et al (2013)[100,] cohort	Metastasis or death	NR	0.017
Cooperberg et al (2015)[98,]	PCa mortality	0.003[a]	NR
Klein et al (2015)[92,]	Metastasis	0.008 to 0.025	0.000 to 0.012
Den et al (2015)[96,]	Metastasis post-RT	0.02 to 0.03	-0.01 to 0.001
Lobo et al(2015)[109,] with Den et al(2014)[97,] cohort	Metastasis or death	NR	0.015
Ross et al (2014)[99,]	Metastasis	0.09 to 0.13	0.036 to 0.040
Karnes et al (2013)[100,]	Metastasis	0.009 to 0.020	-0.004 to 0.003

NR: not reported; PCa: prostate cancer; RT: radiotherapy.
[a] For 25% threshold.

Changes in Management

Several studies have compared physician's treatment recommendations before and after receiving GC results.[62,110,111,112,113,114,] Because the studies did not include information on outcomes and clinical validity has not been established, it is not known whether these treatment decisions represent a clinical improvement in management.

The Association Between the Genomic Classifier and Treatment Effects

Ross et al (2016) reported on results of a retrospective, comparative study of RT after RP for 422 men with pT3 disease or positive margins.[108,] The men were from 4 cohorts previously described (Karnes et al [2013][100,]; Den et al [2014][97,]; Ross et al [2016][108,]; Freedland et al [2016][103,]). The 4 treatment groups were adjuvant RT (n=111), minimal residual disease salvage RT (n=70), salvage RT (n=83), and no RT (n=157). The primary endpoint was a metastasis. Thirty-seven men developed metastasis, and the median follow-up was 8 years. Both CAPRA-S (HR=1.39; 95% CI, 1.18 to 1.62) and Decipher (HR=1.28; 95% CI, 1.08 to 1.52) were independently associated with metastasis in multivariable analysis. There was no evidence that the treatment effect was dependent on genomic risk (interaction p=.16 for CAPRA-S, p=.39 for Decipher). Men with low CAPRA-S or low Decipher scores had a low-risk of metastatic events regardless of treatment selection, and men with high CAPRA-S or Decipher scores benefitted from adjuvant RT compared with the other treatments.

Section Summary: Decipher Radical Prostatectomy Prostate Cancer Classifier

Clinical validity has been evaluated in overlapping validation samples (including the development test set). The validation studies consisted of observational data obtained from registries or medical records with archived samples. Although each study evaluated different outcomes (i.e., metastasis, prostate cancer-specific mortality, BCR) in samples with different populations, all studies reported some incremental improvement in discrimination. The CIs of AUC frequently overlapped between Decipher and comparators. Only 1 study (Karnes et al [2018][107,]) reported 10-year disease-specific survival. Estimates with CIs of outcomes, particularly disease-specific mortality at 10 years, by GC low-, intermediate-, and high-risk are needed as well as reclassification analyses of prostate cancer-specific survival compared with comparators. Results did not consistently demonstrate meaningful improvement in reclassification; possibly most importantly to lower risk categories. It is not clear whether the group of patients identified as low-risk using Decipher could be managed with an observation instead of adjuvant or early salvage RT.

ArteraAI Prostate Test

ArteraAI used for initial management decisions was described in the previous section. This section reviews ArteraAI for management after RP.

Clinically Valid

A test must detect the presence or absence of a condition, the risk of developing a condition in the future, or treatment response (beneficial or adverse).

Two studies reporting clinical validity in the post-RP management setting have evaluated the previously described ArterAI MMAI models for DM and prostate cancer specific mortality (PCSM) on archived biopsy samples.

Bjartell et al (2025) retrospectively analyzed prospectively collected data from 143 individuals with prostate cancer who underwent RP and enrolled in the Urology Prostate Cancer biomarker study in Sweden between 2004 through 2010.[115,] The primary and secondary objectives of the study were to validate the previously developed MMAI models using time to BCR endpoint, defined as 2 successive PSA measurements ≥0.2 ng/mL post-RP, and to evaluate the newly developed biopsy-based model on the surgical endpoint of adverse pathology at RP, defined as Gleason grade group 3 or higher, pT3b or higher, and/or node positive disease. Estimated 5-year BCR rates for individuals with MMAI intermediate- or high-risk scores were significantly worse than individuals with MMAI low-risk scores (p<.001). Furthermore, in both univariable and multivariable analysis the MMAI models were

considered to be prognostic for post-RP endpoints of BCR and adverse pathology (Table 37). The MMAI models were able to risk stratify individuals within NCCN and Cancer of the Prostate Risk Assessment (CAPRA) risk groups, with 29% of individuals with NCCN high-risk disease were reclassified as MMAI intermediate- (68%) or low-risk (11%) and 22% of individuals with CAPRA high-risk disease were reclassified as MMAI intermediate- (74%) or low-risk (7%). Limitations of these studies are synonymous with retrospective analysis, including but not limited to, clinical heterogeneity of study populations' disease, variability in data recording, and different conditions under which measurements occurred, etc. The study excluded patients that didn't meet eligibility criteria and lead to a small number of individuals who comprise of a less racially and ethnically diverse population potentially introducing bias.

Table 37. Univariable and Multivariable analyses of MMAI scores for BCR and Adverse Pathology at Radical Prostatectomy

Analysis	Endpoint	Level	Effect Size[a] (95% CI)	p-value
Univariable	BCR	Continuous[c]	2.45 (1.77 to 3.38)	<.001
	BCR	Intermediate-High vs Low[d]	5.39 (1.91 to 15.23)	=.002
	AP	Continuous[b]	4.85 (2.54 to 10.78)	<.001
	AP	Intermediate-High vs Low[d]	25.9 (6.64 to 173)	<.001
Multivariable	BCR	Continuous[c]	2.13 (1.44 to 3.14)	<.001 (adjusted for NCCN)
	BCR	Intermediate-High vs Low[d]	3.45 (1.20 to 9.86)	=.02
	BCR	Continuous[c]	1.99 (1.29 to 3.07)	=.002 (adjusted for CAPRA)
	BCR	Intermediate-High vs Low[d]	3.48 (1.24 to 9.75)	=.02
	AP	Continuous[b]	5.07 (2.26 to 13.68)	<.001 (adjusted for NCCN)
	AP	Intermediate-High vs Low[d]	41.7 (6.47 to 851)	<.001
	AP	Continuous[c]	4.59 (2.32 to 10.51)	<.001 (adjusted for CAPRA)
	AP	Intermediate-High vs Low[d]	23.7 (5.87 to 162)	<.001

CAPRA: Cancer of the Prostate Risk Assessment; CI: confidence interval; NCCN: National Comprehensive Cancer Network.
a. Effect size refers to subdistribution hazard ratio and odds ratio for BCR and adverse pathology endpoints, respectively.
b. Patients in parentheses represent events and number of patients belonging to the Intermediate-High group
c. MMAI score per 1 standard deviation increase.
d. Pre-established MMAI risk groups.

Li et al (2025) evaluated archived biopsy samples from patients (N=1032) that underwent RP from 1993 to 2001 in the Prostate, Lung, Colorectal, and Ovarian Cancer Screening RCT for the validation of PCSM and OS using ArteraAI MMAI models previously described by Esteva et al (2022).[115,] The results continued to demonstrate that the MMAI models developed for PCMS and DM were prognostic of PCSM (HR, 2.31; 95% CI, 1.6 to 3.35; p<.001 and HR, 1.96; 95% CI, 1.35 to 2.85; p<.001, respectively) and OS (HR, 1.22; 95% CI, 1.01 to 1.47; p=.04 and HR, 1.19; 95% CI, 1.02 to 1.4; p=.03, respectively) with a medium follow-up time of 17 years (interquartile range=14.3, 19.3 years). Notable limitations for this study include a small sample size due to lack of digitized RP slides for all patients from the PLCO RCT, MMAI models were developed using prostate biopsy slides, incomplete clinical data (Gleason scores), potential bias from multiple hypothesis testing, clinical heterogeneity of study populations, variability in data recording, and different conditions under which measurements occurred. Overall, the MMAI models provide further prognostic information and have the potential to identify patients who may benefit from secondary treatments post-RP.

Clinically Useful
A test is clinically useful if the use of the results informs management decisions that improve the net health outcome of care. The net health outcome can be improved if patients receive correct therapy, or more effective therapy, or avoid unnecessary therapy, or avoid unnecessary testing.

Direct Evidence
Direct evidence of clinical utility is provided by studies that have compared health outcomes for patients managed with and without the test. Because these are intervention studies, the preferred evidence would be from RCTs.

BSC did not identify any studies directly supporting the clinical utility of ArteraAI as all studies were from retrospective analyses on data compiled from RCTs and no study reported management changes made in response to ArteraAI Prostate Test results.

Chain of Evidence
Indirect evidence on clinical utility rests on clinical validity. Although disease-specific survival outcomes were reported in these studies, the clinical utility of risk stratification is still unclear, and more studies are needed to determine how to reconcile discordant scores. It is still unclear how reclassifications would change treatment decisions and whether the test could be used to identify individuals who would benefit from additional treatment post-RP.

Section Summary: ArteraAI Prostate Test
For individuals who have localized prostate cancer treated with RP who receive ArteraAI Prostate Test, the evidence includes 2 retrospective cohort studies of clinical validity using archived samples. Relevant outcomes include OS, disease-specific survival, QOL, and treatment-related morbidity.

ArteraAI proved to be prognostic for RP-specific endpoints of BCR and adverse pathology given the statistically significant association. Disease-specific survival outcomes were reported in both studies and the evidence of clinical validity and prognostic accuracy for MMAI scores via ArteraAI testing in patients after RP demonstrated statistically improved PCSM and OS when compared to standard clinicopathologic risk stratification tools. Limitations of these studies are synonymous with retrospective analysis, including but not limited to, clinical heterogeneity of study populations, variability in data recording, and different conditions under which measurements occurred. No study reported management changes made in response to ArteraAI Prostate Test results. Overall, ArteraAI Prostate Test is validated for disease-specific outcomes for prostate cancer patients who underwent RP and can provide additional prognostic information that may guide postoperative management, but further studies are needed to determine if MMAI can be used to decide specific treatment regimens that improve health outcomes. The evidence is insufficient to determine that the technology results in an improvement in the net health outcome.

Management Decision in Castration-Resistant Prostate Cancer
Clinical Context and Test Purpose
In men with metastatic castration-resistant prostate cancer (mCRPC), the purpose of protein biomarker assessment of circulating tumor cells (CTCs) is to inform a decision whether to administer androgen receptor signaling (ARS) inhibitors (e.g., abiraterone, enzalutamide), or a taxane (e.g., docetaxel).

Multiple approved therapeutic options exist for the treatment of men with mCRPC, which are given in conjunction with continued ADT. In particular, ARS inhibitors and taxane-based chemotherapy have both demonstrated effectiveness in prolonging survival but head-to-head comparisons of ARS inhibitors and taxanes in RCTs are lacking. Optimal sequencing of available treatments has also not been established. Guidelines have suggested that both ARS inhibitors and chemotherapy are appropriate for men with mCRPC who have sufficiently good performance status to tolerate chemotherapy as first-line treatment of mCRPC. In practice, sequencing depends on several factors

such as sites and extent of disease, rates of progression, ease and convenience of administration, side effects, comorbidities, and patient preferences. However, unless a man has rapidly progressive, symptomatic disease, ARS inhibitors are generally used as first-line treatment of mCRPC because they are orally administered and have lower toxicity. After disease progression on first-line ARS inhibitor, men could then receive another ARS inhibitor or another systemic therapy, usually a taxane. A test that could inform the choice of second-line therapy would fill an unmet management need. The androgen receptor isoform encoded by splice variant 7 lacks the ligand-binding domain that is the target of the ARS inhibitors enzalutamide and abiraterone. Therefore, detection of androgen receptor splice variant 7 messenger RNA (AR-V7) in CTCs from men with mCRPC might be associated with a lack of response to enzalutamide and abiraterone but not with lack of response to taxanes. The following PICO was used to select literature to inform this review.

Populations

The relevant population of interest is individuals with mCRPC who have progressed on an ARS inhibitor (e.g., enzalutamide, abiraterone), have a good performance status (i.e., are able to tolerate chemotherapy), and who are deciding between a second ARS inhibitor or a taxane.

Interventions

The test being considered is the Oncotype DX AR-V7 Nuclear Detect. Detection of AR-V7 in men with progressive mCRPC is associated with resistance to the ARS inhibitors abiraterone and enzalutamide.[116,] The Oncotype DX AR-V7 Nuclear Detect test is a liquid biopsy test that detects CTCs with nuclear expression of the AR-V7 truncated protein. The test reports a score of AR-V7-positive or -negative. Scher et al (2016) described the development of the test and results in the development cohort in which they observed longer OS for men taking taxanes compared with ARS inhibitors when AR-V7-positive CTCs were detected before therapy (HR=0.24; 95% CI, 0.10 to 0.57).[117,] Scher et al (2017) explored whether expanding the AR-V7 scoring criteria to include both nuclear and cytoplasmic AR-V7 localization improved prediction in the same development cohort and concluded that the expanded "nuclear-agnostic" AR-V7 scoring criterion was less prognostic for men on ARS inhibitor therapy.[118,]

Decisions about management of localized prostate cancer are typically made by patients, urologists, and oncologists in the secondary or tertiary care setting.

Comparators

Since there are no head-to-head comparisons of ARS inhibitors and taxanes in RCTs to determine optimal second- and subsequent-line therapies, in standard clinical care, physicians and men with mCRPC are making treatment decisions based on patient preference, disease characteristics, and comorbidities.

Outcomes

Beneficial outcomes resulting from a true test result are prolonged survival, improved QOL, and reduction in unnecessary treatment-related adverse events. Harmful outcomes resulting from a false test result are unnecessary treatments and shortened survival. The primary survival outcome of interest is OS.

In a systematic review of randomized phase 3 trials of systemic therapies for CRPC, which included 23 trials (total N=13,909 men), the median OS was 19 months.[110,] Outcomes with at least 1 year of follow-up of those surviving would be preferred.

Oncotype DX AR-V7 Nuclear Detect

Oncotype DX AR-V7 Nuclear Detectis used to detect nuclear-localized AR-V7 protein in CTCs of men with mCRPC who have failed first-line therapy and are considering additional ARS inhibitor therapy.

Clinically Valid

A test must detect the presence or absence of a condition, the risk of developing a condition in the future, or treatment response (beneficial or adverse).

Two studies were not included in this assessment of clinical validity because they reported results in the developmental cohort.[117,111,] Two published clinical validity studies met selection criteria.[118,] Characteristics of the studies are provided in Table 38. Scher et al (2018) reported results of a blinded validation study including 142 samples from patients with histologically confirmed, progressing mCRPC from 3 centers in the U. S. and the United Kingdom from 2012 to 2016. The samples were collected prior to the administration of second-line or greater ARS inhibitors or taxanes. Armstrong et al (2019) reported results of the PROPHECY trial, a prospective validation study of AR-V7 detection in men with high-risk mCRPC starting abiraterone or enzalutamide treatment.

Table 38. Characteristics of Clinical Validity Studies Assessing Oncotype DX AR-V7

Study	Study Population	Design	Outcome Measure	Threshold for Positive Index Test	Blinding of Assessors
Scher et al (2018)[118,]	Men with progressing mCRPC undergoing change in therapy	Retrospective; unclear whether samples were consecutive or randomly chosen from eligible	OS (68 men with 12-mo follow-up, 15 men with 24 m follow-up, 6 men with 36-mo follow-up)	At least 1 CTC with an intact nucleus and nuclear-localized AR-V7 signal-to-noise ratio above a prespecified background intensity	Yes
Armstrong et al (2019)[112,]	Men with progressive, high-risk mCRPC initiating standard-of-care treatment with enzalutamide or abiraterone. Prior exposure to enzalutamide or abiraterone was permitted for men who were planning to receive the alternative agent	Prospective, consecutive	PFS (primary); Response rates (PSA and radiographic); OS (secondary)	Johns Hopkins and Epic AR-V7 assays; results for both assays reported	Yes

CTC: circulating tumor cell; mCRPC: metastatic castration-resistant prostate cancer; OS: overall survival; PFS: progression-free survival; PSA: prostate-specific antigen.

Results of the validation studies are shown in Table 39. In Scher et al (2018), median follow-up time in surviving men was not provided. Sixty-eight men were still in the risk set at 12 months. Numerically, men treated with ARS inhibitors had the longest OS if they were AR-V7-negative and had the shortest OS if they were AR-V7-positive. The unadjusted HR for OS for ARS inhibitors versus taxanes was statistically significantly greater than one (favoring ARS inhibitors) in the AR-V7-negative men, while there was no statistically significant difference in OS (but with an unadjusted HR favoring taxanes) in AR-V7-positive men. A test of interaction for AR-V7 status by treatment was not provided.

The analysis was further stratified by a binary prognostic risk score (high vs. low) developed from the training cohort and including clinical biomarkers (see Table). However, the additional stratification resulted in the group that was AR-V7-positive and receiving ARS inhibitors including fewer than 10

men for both high- and low-risk. In Armstrong et al (2019), detection of AR-V7 in CTCs was associated with shorter PFS and OS.

Table 39. Results of Clinical Validity Studies Assessing Oncotype DX AR-V7

Study	Initial N	Final N	Excluded Samples	AR-V7+, %	Median OS (mo) by AR-V7 and Next-Line Therapy			
					AR-V7+, ARS Inhibitor	AR-V7+, Taxane	AR-V7-, ARS Inhibitor	AR-V7-, Taxane
Scher et al (2018)[118,]	248	142 (70 before ARS inhibitor tx, 72 before taxane tx)	144 (93 obtained before first-line tx, 24 duplicates, 23 second-line tx other than ARS inhibitor or taxane, 2 insufficient material, 2 missing clinical data)	24	7.3	14.3	19.8	12.8
HR (95% CI); p ARS vs. taxane					AR-V7+:0.6 (0.3 to 1.4);.25	AR-V7-:1.7 (1.0 to 2.8);.05		
Interaction p					Not reported			
Armstrong et al (2019)[112,]	118	107	2 unevaluable (1%)	10	ARS inhibitor: 8.4 Taxane: NR	ARS inhibitor: 25.5 Taxane: NR		
HR (95% CI); p ARS vs. taxane					NR	NR		
Interaction p					NR			

ARS: androgen receptor signaling; CI: confidence interval; HR: hazard ratio; NR: not reported; OS: overall survival; tx: treatment.

Table 40. Cross-Tabulation of AR-V7 Status and Clinical Risk Score

		Risk Score		
		High	Low	Total
AR-V7 status	Positive	24	10	34
	Negative	46	62	108
	Total	70	72	142

Adapted from Scher et al (2018).[118,]

Tables 41 and 42 display notable limitations identified in each study.

Table 41. Study Relevance Limitations of Clinical Validity Studies Assessing Oncotype DX AR-V7

Study	Population[a]	Intervention[b]	Comparator[c]	Outcomes[d]	Duration of Follow-Up[e]
Scher et al (2018)[118,]					1. Median follow-up in surviving men not clear but overall <50% of men had 12-mo follow-up
Armstrong et al (2019)[112,]					

The study limitations stated in this table are those notable in the current review; this is not a comprehensive gaps assessment.
[a] Population key: 1. Intended use population unclear; 2. Clinical context is unclear; 3. Study population is unclear; 4. Study population not representative of intended use.
[b] Intervention key: 1. Classification thresholds not defined; 2. Version used unclear; 3. Not intervention of interest.
[c] Comparator key: 1. Classification thresholds not defined; 2. Not compared to credible reference standard; 3. Not compared to other tests in use for same purpose.
[d] Outcomes key: 1. Study does not directly assess a key health outcome; 2. Evidence chain or decision model not explicated; 3. Key clinical validity outcomes not reported (sensitivity, specificity, and predictive values); 4. Reclassification of diagnostic or risk categories not reported; 5. Adverse events of the test not described (excluding minor discomforts and inconvenience of venipuncture or noninvasive tests).
[e] Follow-Up key: 1. Follow-up duration not sufficient with respect to natural history of disease (true-positives, true-negatives, false-positives, false-negatives cannot be determined).

Table 42. Study Design and Conduct Limitations of Clinical Validity Studies Assessing Oncotype DX AR-V7

Study	Selection[a]	Blinding[b]	Delivery of Test[c]	Selective Reporting[d]	Data Completeness[e]	Statistical[f]
Scher et al (2018)[118],	2. Unclear if original 248 samples included were consecutive or randomly chosen from eligible					1. Interaction p value not provided
Armstrong et al (2019)[112],	2. Unclear if consecutive or convenience sample					

The study limitations stated in this table are those notable in the current review; this is not a comprehensive gaps assessment.
[a] Selection key: 1. Selection not described; 2. Selection not random or consecutive (i.e., convenience).
[b] Blinding key: 1. Not blinded to results of reference or other comparator tests.
[c] Test Delivery key: 1. Timing of delivery of index or reference test not described; 2. Timing of index and comparator tests not same; 3. Procedure for interpreting tests not described; 4. Expertise of evaluators not described.
[d] Selective Reporting key: 1. Not registered; 2. Evidence of selective reporting; 3. Evidence of selective publication.
[e] Data Completeness key: 1. Inadequate description of indeterminate and missing samples; 2. High number of samples excluded; 3. High loss to follow-up or missing data.
[f] Statistical key: 1. Confidence intervals and/or p values not reported; 2. Comparison with other tests not reported.

Clinically Useful

A test is clinically useful if the use of the results informs management decisions that improve the net health outcome of care. The net health outcome can be improved if patients receive correct therapy, or more effective therapy, or avoid unnecessary therapy, or avoid unnecessary testing.

Direct Evidence

Direct evidence of clinical utility is provided by studies that have compared health outcomes for patients managed with and without the test. Because these are intervention studies, the preferred evidence would be from RCTs.

No studies reporting direct evidence were identified.

Chain of Evidence

Indirect evidence on clinical utility rests on clinical validity. If the evidence is insufficient to demonstrate test performance, no inferences can be made about clinical utility.

Because the clinical validity of the Oncotype DX AR-V7 test has not been established, a chain of evidence supporting the test's clinical utility cannot be constructed.

Section Summary: Oncotype DX AR-V7 Nuclear Detect

Multiple, high-quality studies of the marketed version of the test (including current algorithms and cutoffs), in populations independent of the developmental cohort, that include the intended use population and have consistent and precise results are needed to characterize the performance characteristics.

One retrospective analysis of 142 men from the U. S. and the United Kingdom including men with progressing mCRPC undergoing a change in therapy is available. The median follow-up in surviving men is unclear, but, overall, 68 men had 12 months of follow-up, 15 men had 24 months of follow-up, and 6 men had 36 months of follow-up. Men treated with ARS inhibitors had the longest OS if they were AR-V7-negative (median, 19.8 months) and had the shortest OS if they were AR-V7-positive (median, 7.3 months). The unadjusted HR for OS was statistically significantly longer for ARS inhibitors compared with taxanes in the AR-V7-negative men (HR=1.7; 95% CI, 1.0 to 2.8) but not in AR-V7-positive men (0.6; 95% CI, 0.3 to 1.4). However, a test of interaction for AR-V7 status by treatment was not provided. In a prospective validation study of AR-V7 detection in 118 men with high-risk mCRPC starting abiraterone or enzalutamide treatment, the detection of AR-V7 in CTCs was associated with shorter PFS and OS.

ArteraAI Prostate Test

Clinically Valid

A test must detect the presence or absence of a condition, the risk of developing a condition in the future, or treatment response (beneficial or adverse).

Feng et al (2025) applied the ArteraAI MMAI model to core prostate biopsies from 420 patients with nonmetastatic castration-resistant prostate cancer (nmCRPC) to estimate clinical outcomes (metastasis-free survival [MFS], second progression-free survival [PFS2], and overall OS) for patients who either received apalutamide or placebo.[119,] The MMAI model was associated with shorter MFS, PFS2, and OS and capable of risk-stratifying patients with nmCRPC (Table 43). Moreover, patients were split into 2 categories, MMAI high-risk and MMAI non-high-risk, to evaluate these clinical outcomes resulting in the MMAI high-risk group having shorter MFS (HR, 1.47; 95% CI, 1.03 to 2.11; p=.04) compared to MMAI low-risk groups. The MMAI high-risk group demonstrated a significant improvement in MFS (HR, 0.19; 95% CI, 0.12 to 0.29; p<.005), PFS2 (HR, 0.47, 95% CI, 0.33 to 0.68; p<.005), and OS (HR, 0.6, 95% CI, 0.40 to 0.89; p=.01) for patients treated with apalutamide compared with placebo. Notable limitations include a constrained sample size from a single RCT, in which only 39% of the original patients enrolled were included, and those synonymous with a retrospective analysis, including but not limited to, clinical heterogeneity of study populations, variability in data recording, and different conditions under which measurements occurred.

Table 43. Cox Regression Results for All Patients Using Continuous MMAI Risk Score

Analysis	Variable	Endpoint	Hazard Ratio (95% CI)	p-value
Univariable	MMAI risk score	MFS	1.24 (1.04 to 1.47)	.01
	MMAI risk score	PFS2	1.20 (1.03 to 1.39)	.02
	MMAI risk score	OS	1.19 (1.01 to 1.41)	.04
Multivariable	MMAI risk score	MFS	1.72 (1.34 to 2.21)	<.005
	MMAI risk score	PFS2	1.57 (1.20 to 2.05)	<.005
	MMAI risk score	OS	1.41 (1.06 to 1.87)	.02
	MMAI risk score: Treatment[a]	MFS	0.63 (0.45 to 0.89)	.01

	MMAI risk score: Treatment[a]	PFS2	0.70 (0.51 to 0.97)	.03

HR: hazard ratio; MFS: metastasis-free survival; MMAI: multimodal artificial intelligence; OS: overall survival; PFS2: second progression-free survival.
[a]Treatment refers to both apalutamide and placebo.

Clinically Useful

A test is clinically useful if the use of the results informs management decisions that improve the net health outcome of care. The net health outcome can be improved if patients receive correct therapy, or more effective therapy, or avoid unnecessary therapy, or avoid unnecessary testing.

Direct Evidence

Direct evidence of clinical utility is provided by studies that have compared health outcomes for patients managed with and without the test. Because these are intervention studies, the preferred evidence would be from RCTs.

BSC did not identify any studies directly supporting the clinical utility of ArteraAI as all studies were from retrospective analyses on data compiled from RCTs and no study reported management changes made in response to ArteraAI Prostate Test results.

Chain of Evidence

Indirect evidence on clinical utility rests on clinical validity. Although disease-specific survival outcomes were reported in this study, the clinical utility of risk stratification is still unclear, and more studies are needed to determine how to reconcile discordant scores. It is still unclear how reclassifications would change treatment decisions and whether the test could be used to identify individuals who would benefit from additional treatment.

Section Summary: ArteraAI Prostate Test

Multiple, high-quality studies of the marketed version of the test (including current algorithms and cutoffs), in populations independent of the developmental cohort, that include the intended use population and have consistent and precise results are needed to characterize the performance characteristics.

For individuals who have nmCRPC who receive ArteraAI Prostate Test, the evidence includes 1 retrospective cohort study of clinical validity using archived samples. Relevant outcomes include OS, disease-specific survival, QOL, and treatment-related morbidity. The MMAI model was able to predict treatment effects and determine what nmCRPC patients would derive the most benefit with apalutamide treatment. Limitations of these studies are synonymous with retrospective analysis, including but not limited to, clinical heterogeneity of study populations, variability in data recording, and different conditions under which measurements occurred, etc. No study reported management changes made in response to ArteraAI Prostate Test results. Overall, ArteraAI Prostate Test demonstrated its prognostic capabilities for nmCRPC patients and the potential to predict treatment management, but further studies are needed to determine if MMAI can be used to decide specific treatment regimens that improve net health outcomes. The evidence is insufficient to determine that the technology results in an improvement in the net health outcome

Management Decision in Castration-Sensitive Prostate Cancer

Clinical Context and Test Purpose

Individuals with metastatic castration-sensitive prostate cancer (mCSPC), the purpose of MMAI algorithm, ArteraAI, is to inform a decision whether to administer androgen receptor signaling (ARS) inhibitors (e.g., abiraterone, enzalutamide), or a taxane (e.g., docetaxel).

Multiple approved therapeutic options exist for the treatment of individuals with mCSPC, which are given in conjunction with continued ADT. In particular, ARS inhibitors and taxane-based

chemotherapy have both demonstrated effectiveness in prolonging survival but head-to-head comparisons of ARS inhibitors and taxanes in RCTs are lacking. Optimal sequencing of available treatments has also not been established. Guidelines have suggested ADT using medical castration or surgical orchiectomy for initial systemic therapy, but modern approaches use ADT in combination with ARS inhibitors and/or chemotherapy as first-line treatment of mCSPC. In practice, sequencing depends on several factors such as sites and extent of disease, rates of progression, ease and convenience of administration, side effects, comorbidities, and patient preferences. A test that could inform and guide treatment decision would fill an unmet management need.

The following PICO was used to select literature to inform this review.

Populations
The relevant population of interest is individuals with mCSPC, who have undergone metastasis direct therapy (MDT) via radiation or surgery, and who are deciding between ADT alone or in combination with an ARS inhibitor or taxane.

Interventions
ArteraAI Prostate Test, described in the previous section, can be used to classify individuals who have undergone MDT and predict disease-specific outcomes.

Comparators
Clinicopathologic risk stratification along with age/life expectancy and patient preference are currently being used to make decisions about prostate cancer management. Clinical characteristics (e.g., stage, biopsy Gleason grade, serum PSA level) and demographic characteristics (e.g., age, life expectancy) are combined to classify individuals according to risk. National Comprehensive Cancer Network (NCCN) and AUA have provided treatment recommendations based on risk stratification and life expectancy.[13,37] However, during standard clinical care, physicians and individuals with mCSPC are making treatment decisions based on patient preference, disease characteristics, and comorbidities.

Outcomes
Beneficial outcomes resulting from a true test result are prolonged survival, improved QOL, and reduction in unnecessary treatment-related adverse events. Harmful outcomes resulting from a false test result are unnecessary treatments and shortened survival. The primary survival outcome of interest is OS.

Study Selection Criteria
For the evaluation of clinical validity of the ArteraAI Prostate Test studies that meet the following eligibility criteria were considered:

- Reported on the accuracy of the marketed version of the technology (including any algorithms used to calculate scores);
- Included a validation cohort independent of the development cohort;
- Included a suitable reference standard (10-year prostate cancer-specific survival or death rate);
- Patient/sample clinical characteristics were described;
- Patient/sample selection criteria were described.

ArteraAI Prostate Test
Clinically Valid
A test must detect the presence or absence of a condition, the risk of developing a condition in the future, or treatment response (beneficial or adverse).

Markowski et al (2024) applied the ArteraAI MMAI model to prostate biopsies or prostatectomy samples from a phase 3 RCT, ChemoHormonal Therapy Versus Androgen Ablation Randomized Trial

for Extensive Disease in Prostate Cancer (CHAARTED) which enrolled metastatic hormone-sensitive prostate cancer patients (mHSPC) who received either ADT alone or ADT plus docetaxel chemotherapy, with the primary objective to evaluate the prognostic ability of the MMAI algorithm in mHSPC patients.[120,] The univariable analysis results for the association between the MMAI algorithm score, as a continuous variable, and the study endpoints (OS, clinical progression [CP], and castration-resistant prostate cancer [CRPC] rate) by clinical subgroup demonstrated that the model was prognostic for overall survival OS (HR, 1.51; 95% CI, 1.33 to 1.73; p<.001), CP (sHR, 1.54; 95% CI, 1.36 to 1.74; p<.001), and CRPS (sHR, 1.63; 95% CI, 1.45 to 1.83; p<.001). Moreover, multivariable proportional-hazards models adjusted for treatment arms and clinical risk groups were created to assess the additional prognostic ability of the MMAI model over covariates of interest and demonstrated that the MMAI algorithm was prognostic for OS (HR, 1.51, 95% CI, 1.33 to 1.73; p<.001), CP (HR, 1.54; 95% CI, 1.36 to 1.74; p<.001), and CRPC (HR, 1.63, 95% CI, 1.45 to 1.83; p<.001). Kaplan-Meier curves were generated for a binary analysis comparing mHSPC patients categorized as high-risk versus low/intermediate-risk by the MMAI model and resulted in significantly worse outcomes for MMAI high-risk patients for OS (p<.001), CP (p<.001), and CRPC (p<.001). Notable limitations of this study include those associated with retrospective analysis, a small sample size from a single RCT, and applying the model to a subset of patients for which it was not developed for.

Wang et al (2025) set out to evaluate the prognostic and predictive performance of the MMAI models developed by ArteraAI in oligometastatic castration-sensitive prostate cancer (omCSPC), defined as ≤5 lesions on either conventional (computed tomography [CT] or nuclear medicine bone scan) or molecular (prostate-specific membrane antigen [PSMA] or choline positron emission tomography [PET]) imaging, from 2 RCTs. [121,] The univariable analysis of the MMAI model as a continuous variable demonstrated that patients with a MMAI high-risk score were significantly associated with worse OS (HR, 6.46; 95% CI, 1.44 to 28.9; p=.01) and shorter time to castration-resistant prostate cancer (TTCRPC; HR, 2.07; 95% CI, 1.15 to 3.72; p=.015) compared to patients with MMAI low-risk scores. A multivariable analysis to account for covariates was conducted and concluded that MMAI scores were the only variable significantly associated with OS (HR, 6.51; 95% CI, 1.32 to 32.2; p=.02) for omCSPC patients. Kaplan-Meier curves were generated to assess the endpoints of OS and TTCRPC for MMAI high- and low-risk patients and demonstrated worse OS (p=.005) and shorter TTCRPC (p=.013) for MMAI high-risk compared to low-risk patients. Additional Kaplan-Meier curves were created to evaluate the MMAI score as a biomarker to predict response to MDT within a subset of patients that enrolled in the STOMP and ORIOLE RCTs using metastasis-free survival (MFS) as the endpoint, given that there were too few OS and castration-resistance events for this subset. Patients with MMAI high-risk scores demonstrated significantly improved MFS with MDT (p=.039) compared to observation. Notable limitations of this study include those associated with retrospective analysis, a homogenous sample size that included mainly metachronous omCSPC patients, lack of appropriate comparators, and applying the model to a subset of patients for which it was not developed.

Clinically Useful

A test is clinically useful if the use of the results informs management decisions that improve the net health outcome of care. The net health outcome can be improved if patients receive correct therapy, or more effective therapy, or avoid unnecessary therapy, or avoid unnecessary testing.

Direct Evidence

Direct evidence of clinical utility is provided by studies that have compared health outcomes for patients managed with and without the test. Because these are intervention studies, the preferred evidence would be from RCTs.

No studies reporting direct evidence were identified.

Chain of Evidence
Indirect evidence on clinical utility rests on clinical validity. If the evidence is insufficient to demonstrate test performance, no inferences can be made about clinical utility. Although disease-specific survival outcomes were reported in these studies, the clinical utility of risk stratification is still unclear, and more studies are needed to determine how to reconcile discordant scores. It is still unclear how reclassifications would change treatment decisions and whether the test could be used to identify individuals who would benefit from additional treatment.

Section Summary: ArteraAI Prostate Test
For individuals who have mCSPC who receive ArteraAI Prostate Test, the evidence includes 2 retrospective cohort studies of clinical validity using archived samples. Relevant outcomes include OS, disease-specific survival, QOL, and treatment-related morbidity. MMAI was able to estimate treatment effects and determine that MMAI high-risk mCRPC patients would derive benefit from MDT when compared to observation. Limitations of these studies are synonymous with retrospective analysis, including but not limited to, clinical heterogeneity of study populations, variability in data recording, and different conditions under which measurements occurred, etc. No study reported management changes made in response to ArteraAI Prostate Test results. Overall, ArteraAI Prostate Test is prognostic for mCSPC patients and has the potential to guide treatment management, but further studies are needed to determine if MMAI can be used to decide specific treatment regimens that improve net health outcomes. The evidence is insufficient to determine that the technology results in an improvement in the net health outcome.

Supplemental Information
The purpose of the following information is to provide reference material. Inclusion does not imply endorsement or alignment with the evidence review conclusions.

Practice Guidelines and Position Statements
Guidelines or position statements will be considered for inclusion in 'Supplemental Information' if they were issued by, or jointly by, a US professional society, an international society with US representation, or National Institute for Health and Care Excellence (NICE). Priority will be given to guidelines that are informed by a systematic review, include strength of evidence ratings, and include a description of management of conflict of interest.

American Society of Clinical Oncology
In 2020, the American Society of Clinical Oncology (ASCO) published a guideline on molecular biomarkers in localized prostate cancer.[122,] The guidelines state, "Currently, there are no strong data or expert guidelines to support active surveillance in otherwise healthy men with Grade Group 3 or higher cancer; therefore, we would consider the use of genomic biomarkers only in situations in which the assay result, when considered as a whole with routine clinical factors, is likely to affect a physician's recommendation or a patient's choice for surveillance versus treatment, but they should not be used routinely."

Specific recommendations included the following:
Molecular biomarkers to identify patients with prostate cancer who are most likely to benefit from active surveillance:

- Recommendation 1.1. Commercially available molecular biomarkers (i.e. Oncotype Dx Prostate, Prolaris, Decipher, and ProMark) may be offered in situations in which the assay result, when considered as a whole with routine clinical factors, is likely to affect management. Routine ordering of molecular biomarkers is not recommended (Type: Evidence based; Evidence quality: Intermediate; Strength of recommendation: Moderate).
- Recommendation 1.2. Any additional molecular biomarkers evaluated do not have sufficient data to be clinically actionable or are not commercially available and thus should not be offered (Type: Evidence based; Evidence quality: Insufficient; Strength of recommendation: Moderate).

Molecular biomarkers to diagnose clinically significant prostate cancer:

- Recommendation 2.1. Commercially available molecular biomarkers (i.e. Oncotype Dx Prostate, Prolaris, Decipher, and ProMark) may be offered in situations in which the assay result, when considered as a whole with routine clinical factors, is likely to affect management. Routine ordering of molecular biomarkers is not recommended (Type: Evidence based; Evidence quality: Intermediate; Recommendation: Moderate).
- Recommendation 2.2. Any additional molecular biomarkers evaluated do not have sufficient data to be clinically actionable or are not commercially available and thus should not be offered (Type: Evidence based; Evidence quality: Insufficient; Strength of recommendation: Moderate).

Molecular biomarkers to guide the decision of post prostatectomy adjuvant versus salvage radiation: Recommendation 3.1. The Expert Panel recommends consideration of a commercially available molecular biomarker (e.g., Decipher Genomic Classifier) in situations in which the assay result, when considered as a whole with routine clinical factors, is likely to affect management. In the absence of prospective clinical trial data, routine use of genomic biomarkers in the postprostatectomy setting to determine adjuvant versus salvage radiation or to initiate systemic therapies should not be offered (Type: Evidence based; Evidence quality: Intermediate; Strength of recommendation: Moderate).

Recommendation 3.2. Any additional molecular biomarkers evaluated do not have sufficient data to be clinically actionable or are not commercially available and thus should not be offered (Type: Evidence based; Evidence quality: Insufficient; Strength of recommendation: Moderate).

American Urological Association and American Society for Radiation Oncology

The American Urological Association and American Society for Radiation Oncology published guidelines on clinically localized prostate cancer.[14,] The guidelines included the following statements on risk assessment:

1. "Clinicians should use clinical T stage, serum PSA, Grade Group (Gleason score), and tumor volume on biopsy to risk stratify patients with newly diagnosed prostate cancer. (Strong Recommendation; Evidence Level: Grade B)."
2. "Clinicians may selectively use tissue-based genomic biomarkers when added risk stratification may alter clinical decision-making. (Expert Opinion)."
3. "Clinicians should not routinely use tissue-based genomic biomarkers for risk stratification or clinical decision-making. (Moderate Recommendation; Evidence Level: Grade B)."

The American Urological Association (2018) published guidelines for castration-resistant prostate cancer.[123,] The guidelines do not mention AR-V7 assays.

National Comprehensive Cancer Network

The National Comprehensive Cancer Network (NCCN) guidelines for prostate cancer (v.1.2025) provide a table of tissue-based tests for prostate cancer prognosis.[13,]Guidelines are updated frequently; refer to the source document for current recommendations. The most recent guidelines (v.1.2025) include the following recommendations and statements related to risk-stratification and testing for biomarkers:

<u>22-gene genomic classifier (GC) (Decipher)</u>

- "RT alone may be considered for patients with a low GC score and NCCN intermediate-risk disease."
- "The addition of ST-ADT should be considered for patients with a high GC score given their increased risk of DM and significant benefit of ST-ADT on DM, irrespective of RT dose or brachytherapy boost."
- "Patients with a GC low-risk score should be counseled that the absolute benefit of LT-ADT over ST-ADT is smaller than for patients with GC high-risk scores and when accounting for

patient age, comorbidities, and patient preferences, it may be reasonable with shared decision-making to use a duration shorter than LT-ADT."
- "For patients with node-negative disease post-RP planned for early secondary RT (PSA ≤ 0.5 ng/mL) with GC low or intermediate risk, use of RT alone should be considered."
- "For patients planned for early secondary RT with a GC high-risk tumor, use of secondary RT with ADT is recommended."

ArteraAI Prostate Test
- Patients with intermediate-risk prostate cancer planning to receive RT, those with biomarker-positive disease, and especially those with unfavorable intermediate-risk disease, should be recommended for the addition of ST-ADT regardless of RT dose or type, notwithstanding contraindications to ADT. Those with biomarker (-) tumors, especially tumors with more favorable prognostic risk, may consider the use of RT alone.
- "Specific MMAI cut points have not been published to date to precisely guide specific treatment decisions. Rather, the test may be used to provide more accurate risk stratification to enable improved shared decision-making."

The discussion section in the guidelines, which is pending update as of April 2024, includes the following statements related to risk stratification:
- Patients with low or favorable intermediate disease and life expectancy greater than or equal to 10 years may consider the use of Decipher, Oncotype DX Prostate, or Prolaris during initial risk stratification. Patients with unfavorable intermediate- and high-risk disease and life expectancy greater than or equal to 10 years may consider the use of Decipher or Prolaris.
- Decipher may be considered to inform adjuvant treatment if adverse features are found after radical prostatectomy and during workup for radical prostatectomy PSA persistence or recurrence (NCCN category 2A; Simon et al [2019] category 2B).

The panel also stated that "the use of AR-V7 tests in circulating tumor cells can be considered to help guide selection of therapy in the post-abiraterone/enzalutamide metastatic castration-resistant prostate cancer setting."

Of note, in the April 2024 version of the NCCN guideline, the following footnotes were noted to be removed, but the related discussion sections are still pending update:
- "Decipher molecular assay should be considered if not previously performed to inform adjuvant treatment if adverse features are found post- RP."
- "Consider AR-V7 testing to help guide selection of therapy.

National Institute for Health and Care Excellence
In 2019 (updated 2021), the National Institute for Health and Care Excellence updated its guidance on the diagnosis and management of prostate cancer.[124,] The guidance did not address gene expression profile testing.

U.S. Preventive Services Task Force Recommendations
Not applicable.

Medicare National and Local Coverage
There is no national coverage determination. In the absence of a national coverage determination, coverage decisions are left to the discretion of local Medicare carriers.

Local coverage guidance for California is provided by the Molecular Diagnostic Services Program (MolDX®) program in the document MolDX: Prostate Cancer Genomic Classifier Assay for Men with Localized Disease and the associated Billing and Coding: MolDX: Prostate Cancer Genomic Classifier Assay for Men with Localized Disease.

MolDx considers the Decipher® Prostate Cancer Classifier Assay (identified as Decipher® reasonable and necessary to help identify men with localized Prostate Cancer and a life expectancy of at least 10 years who are good candidates for active surveillance according to the most recent National Comprehensive Cancer Network (NCCN) guidelines. Decipher® is covered for men with prostate cancer with localized or biochemically recurrent adenocarcinoma of the prostate (i.e., no clinical evidence of metastasis) who have a life expectancy of greater than or equal to 10 years if they are a candidate for and are considering (or being considered for) at least 1 of the following:

- Conservative management and yet would be eligible for definitive therapy (radical prostatectomy (RP), radiation or brachytherapy), or;
- Radiation therapy and yet would be eligible for the addition of a brachytherapy boost, or;
- Radiation therapy and yet would be eligible for the addition of short-term androgen deprivation therapy (ADT), or;
- Radiation therapy with short-term ADT yet would be eligible for the use of long-term ADT, or;
- Radiation with standard ADT yet would be eligible for systemic therapy intensification using next generation androgen signaling inhibitors or chemotherapy, or;
- Observation post-prostatectomy yet would be eligible for the addition of post-operative adjuvant radiotherapy, or;
- Salvage radiotherapy post-prostatectomy yet would be eligible for the addition of ADT.

The following criteria must also be met for coverage:

- The assay is performed on formalin-fixed paraffin embedded (FFPE) prostate biopsy tissue with at least 0.5 mm of linear tumor diameter or FFPE tissue from a prostate resection specimen, and;
- Result will be used to determine treatment according to established practice guidelines, and;
- Patient has not received pelvic radiation or ADT prior to the biopsy or prostate resection specimen, and;
- Patient is monitored for disease progression according to established standards of care.

Other genomic tests that demonstrate an equivalent analytical validity and clinical validity will be considered reasonable and necessary for the same indications. Analytical and clinical validity will be assessed as part of a thorough and comprehensive technical assessment by MolDX and will similarly attain coverage for indications that are supported by the evidence and intended use within the scope of this policy.

The following CPT and PLA Codes are included in MolDx Billing and Coding: MolDX: Prostate Cancer Genomic Classifier Assay for Men with Localized Disease:

Code	Description	TEST NAME
81541	ONCOLOGY (PROSTATE), MRNA GENE EXPRESSION PROFILING BY REAL-TIME RT-PCR OF 46 GENES (31 CONTENT AND 15 HOUSEKEEPING), UTILIZING FORMALIN-FIXED PARAFFIN-EMBEDDED TISSUE, ALGORITHM REPORTED AS A DISEASE-SPECIFIC MORTALITY RISK SCORE	Prolaris® Prostate Cancer Genomic Assay
81542	ONCOLOGY (PROSTATE), MRNA, MICROARRAY GENE EXPRESSION PROFILING OF 22 CONTENT GENES, UTILIZING FORMALIN-FIXED PARAFFIN-EMBEDDED TISSUE, ALGORITHM REPORTED AS METASTASIS RISK SCORE	Decipher Prostate Genomic Classifier.
0047U	ONCOLOGY (PROSTATE), MRNA, GENE EXPRESSION PROFILING BY REAL-TIME RT-PCR OF 17 GENES (12 CONTENT AND 5 HOUSEKEEPING), UTILIZING FORMALIN-FIXED	Oncotype DX® Genomic Prostate Score™

Code	Description	TEST NAME
	PARAFFIN-EMBEDDED TISSUE, ALGORITHM REPORTED AS A RISK SCORE	

Ongoing and Unpublished Clinical Trials

Some currently unpublished trials that might influence this review are listed in Table 44.

Table 44. Summary of Key Trials

NCT No.	Trial Name	Planned Enrollment	Completion Date
Ongoing			
Prolaris or Decipher or Oncotype			
NCT04404894[a]	Long-Term Prospective Registry to Evaluate Treatment Decisions and Clinical Outcomes in Prostate Cancer Patients From Diverse Urology Practice Settings Following Prolaris® Testing	500	Nov 2029
Decipher			
NCT02723734	Validation Study on the Impact of Decipher Testing - VANDAAM Study	240	Mar 2025
NCT04396808	Genomics in Michigan to AdJust Outcomes in Prostate canceR (G-MAJOR): A Randomized Multi-center Study for Men With Newly Diagnosed Favorable Risk Prostate Cancer	900	Jul 2025
NCT05050084[a]	Parallel Phase III Randomized Trials of Genomic-Risk Stratified Unfavorable Intermediate Risk Prostate Cancer: De-Intensification and Intensification Clinical Trial Evaluation (GUIDANCE)	2050	Apr 2037
NCT04484818	A Phase III Double Blinded Study of Early Intervention After RADICAl ProstaTEctomy With Androgen Deprivation Therapy With or Without Darolutamide vs. Placebo in Men at Highest Risk of Prostate Cancer Metastasis by Genomic Stratification (ERADICATE)	810	May 2028
NCT04513717	Parallel Phase III Randomized Trials for High Risk Prostate Cancer Evaluating De-Intensification for Lower Genomic Risk and Intensification of Concurrent Therapy for Higher Genomic Risk With Radiation (PREDICT-RT*)	2478	Dec 2033
NCT06282588	Treatment of High-Risk Prostate Cancer Guided by Novel Diagnostic Radio- and Molecular Tracers (THUNDER): A Two-part Phase 2/ 3 Trial	493	Dec 2030
NCT05100472	Phase II Trial of Short Course Androgen Deprivation, Hypofractionated Pelvic Radiation and a Brachytherapy Boost for NCCN High-Risk Prostate Cancer With Low-Intermediate Risk Decipher Genomic Score	50	Oct 2025
NCT03495427	The Utility of PSMA-PET Imaging for Detecting Early Metastatic Prostate Cancer in Men With High GC Decipher® Test Scores: A Sub-aim of the VANDAAM Study (MCC #18523)	60	May 2029
NCT05169970	A Phase II Study of Decipher-Guided Dose Escalated Radiation Therapy In Unfavorable Intermediate Risk Prostate Cancer Patients Treated SBRT Alone Without Androgen Deprivation Therapy	215	Dec 2025
NCT02609269	Prospective Expression Analysis Using The Decipher Genomics Resource for Intelligent Discovery (GRID) and Data Sharing Progra	1,000,000	Dec 2040
NCT04541030	UAB-NCI Collaborative Study on Integrating Genomic Prostate Score With MRI Targeted Prostate Biopsies	241	Mar 2025
ArteraAI			
NCT06582446	Whole-pelvis Hypofractionated Radiotherapy Combined With Dose-escalation to the Prostate and Androgen Deprivation Therapy in Primary Localized, NCCN and MMAI High-risk Prostate Cancer - a Prospective, Single-arm, Phase II Study	30	Aug 2027

NCT No.	Trial Name	Planned Enrollment	Completion Date
NCT06772441[a]	Prostate-only, Dose-escalated Radiotherapy Plus Concomitant Androgen Deprivation Therapy in Primary Localized, NCCN High Risk and MMAI Classifier Low or Intermediate-risk Prostate Cancer - a Prospective, Single-arm, Phase II Study	30	Oct 2027
	Ge		
Unpublished			
Prolaris or Decipher or Oncotype			
NCT03152448[a]	Two-Part Prospective Study to Measure Impact of Prolaris® Testing Added to Treatment Decision Following Biopsy in Newly Diagnosed Prostate Cancer Patients to Measure Prediction of Progression/Recurrence in Men Treated at VAMC	1511	Mar 2022 (Terminated)
NCT03290508[a]	Long-Term Prospective Registry to Evaluate Treatment Decisions and Clinical Outcomes in Patients With Favorable Intermediate-Risk Localized Prostate Cancer Following Cell Cycle Progression (CCP) Testing (Prolaris® Test)	524	Jan 2022 (Terminated)
NCT03851211	Prolaris Enhanced Risk Stratification - an ecONomic and clinicAL Evaluation	100	Oct 2020 (Unknown)
NCT03511235[a]	Clinical Outcomes in Men With Prostate Cancer Who Selected Active Surveillance Using Prolaris Testing	774	Jul 2018
NCT02648919	Phase II Clinical Study of Noni Extract in Men With Very Low Risk or Low Risk Prostate Cancer	6	Dec 2018 (Terminated)
NCT02668276	The Impact of a Gene Expression Profile on Treatment Choice and Outcome Among Minority Men Newly Diagnosed With Prostate Cancer: A Randomized Trial	200	Aug 2019

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

References

1. National Cancer Institute. Cancer Stat Facts: Cancer of Any Site. https://seer.cancer.gov/statfacts/html/all.html. Accessed March 18, 2025.
2. Dall'Era MA, Cooperberg MR, Chan JM, et al. Active surveillance for early-stage prostate cancer: review of the current literature. Cancer. Apr 15 2008; 112(8): 1650-9. PMID 18306379
3. Bangma CH, Roemeling S, Schröder FH. Overdiagnosis and overtreatment of early detected prostate cancer. World J Urol. Mar 2007; 25(1): 3-9. PMID 17364211
4. Johansson JE, Andrén O, Andersson SO, et al. Natural history of early, localized prostate cancer. JAMA. Jun 09 2004; 291(22): 2713-9. PMID 15187052
5. Ploussard G, Epstein JI, Montironi R, et al. The contemporary concept of significant versus insignificant prostate cancer. Eur Urol. Aug 2011; 60(2): 291-303. PMID 21601982
6. Harnden P, Naylor B, Shelley MD, et al. The clinical management of patients with a small volume of prostatic cancer on biopsy: what are the risks of progression? A systematic review and meta-analysis. Cancer. Mar 01 2008; 112(5): 971-81. PMID 18186496
7. Brimo F, Montironi R, Egevad L, et al. Contemporary grading for prostate cancer: implications for patient care. Eur Urol. May 2013; 63(5): 892-901. PMID 23092544
8. Eylert MF, Persad R. Management of prostate cancer. Br J Hosp Med (Lond). Feb 2012; 73(2): 95-9. PMID 22504752
9. Eastham JA, Kattan MW, Fearn P, et al. Local progression among men with conservatively treated localized prostate cancer: results from the Transatlantic Prostate Group. Eur Urol. Feb 2008; 53(2): 347-54. PMID 17544572
10. Bill-Axelson A, Holmberg L, Ruutu M, et al. Radical prostatectomy versus watchful waiting in early prostate cancer. N Engl J Med. May 12 2005; 352(19): 1977-84. PMID 15888698
11. Thompson IM, Goodman PJ, Tangen CM, et al. Long-term survival of participants in the prostate cancer prevention trial. N Engl J Med. Aug 15 2013; 369(7): 603-10. PMID 23944298
12. Albertsen PC, Hanley JA, Fine J. 20-year outcomes following conservative management of clinically localized prostate cancer. JAMA. May 04 2005; 293(17): 2095-101. PMID 15870412

13. National Comprehensive Cancer Network (NCCN). NCCN Clincal Practice Guidelines in Oncology: Prostate Cancer. Version 1.2025. https://www.nccn.org/professionals/physician_gls/pdf/prostate.pdf. Accessed March 18, 2025.
14. Eastham JA, Auffenberg GB, Barocas DA, et al. Clinically Localized Prostate Cancer: AUA/ASTRO Guideline, Part I: Introduction, Risk Assessment, Staging, and Risk-Based Management. J Urol. Jul 2022; 208(1): 10-18. PMID 35536144
15. Thompson IM, Valicenti RK, Albertsen P, et al. Adjuvant and salvage radiotherapy after prostatectomy: AUA/ASTRO Guideline. J Urol. Aug 2013; 190(2): 441-9. PMID 23707439
16. Simon RM, Paik S, Hayes DF. Use of archived specimens in evaluation of prognostic and predictive biomarkers. J Natl Cancer Inst. Nov 04 2009; 101(21): 1446-52. PMID 19815849
17. Food and Drug Administration (FDA). The Public Health Evidence for FDA Oversight of Laboratory Developed Tests: 20 Case Studies. 2015; http://www.nila-usa.org/images/nila/The%20Public%20Health%20Case%20for%20FDA%20Oversight%20of%20LDTs%20110915(2)_508ed%20(1).pdf. Accessed March 18, 2025.
18. Borley N, Feneley MR. Prostate cancer: diagnosis and staging. Asian J Androl. Jan 2009; 11(1): 74-80. PMID 19050692
19. Freedland SJ. Screening, risk assessment, and the approach to therapy in patients with prostate cancer. Cancer. Mar 15 2011; 117(6): 1123-35. PMID 20960523
20. Albertsen PC. Treatment of localized prostate cancer: when is active surveillance appropriate?. Nat Rev Clin Oncol. Jul 2010; 7(7): 394-400. PMID 20440282
21. Ip S, Dahabreh IJ, Chung M, et al. An evidence review of active surveillance in men with localized prostate cancer. Evid Rep Technol Assess (Full Rep). Dec 2011; (204): 1-341. PMID 23126653
22. American Urological Association (AUA). Clinically Localized Prostate Cancer: AUA/ASTRO/SUO Guideline. 2022; https://www.auanet.org/guidelines-and-quality/guidelines/clinically-localized-prostate-cancer-aua/astro-guideline-2022. Accessed March 18, 2025.
23. Nam RK, Cheung P, Herschorn S, et al. Incidence of complications other than urinary incontinence or erectile dysfunction after radical prostatectomy or radiotherapy for prostate cancer: a population-based cohort study. Lancet Oncol. Feb 2014; 15(2): 223-31. PMID 24440474
24. Hamdy FC, Donovan JL, Lane JA, et al. 10-Year Outcomes after Monitoring, Surgery, or Radiotherapy for Localized Prostate Cancer. N Engl J Med. Oct 13 2016; 375(15): 1415-1424. PMID 27626136
25. Tosoian JJ, Mamawala M, Epstein JI, et al. Intermediate and Longer-Term Outcomes From a Prospective Active-Surveillance Program for Favorable-Risk Prostate Cancer. J Clin Oncol. Oct 20 2015; 33(30): 3379-85. PMID 26324359
26. Klotz L, Vesprini D, Sethukavalan P, et al. Long-term follow-up of a large active surveillance cohort of patients with prostate cancer. J Clin Oncol. Jan 20 2015; 33(3): 272-7. PMID 25512465
27. Wilt TJ, Brawer MK, Jones KM, et al. Radical prostatectomy versus observation for localized prostate cancer. N Engl J Med. Jul 19 2012; 367(3): 203-13. PMID 22808955
28. Wilt TJ, Jones KM, Barry MJ, et al. Follow-up of Prostatectomy versus Observation for Early Prostate Cancer. N Engl J Med. Jul 13 2017; 377(2): 132-142. PMID 28700844
29. van den Bergh RC, Korfage IJ, Roobol MJ, et al. Sexual function with localized prostate cancer: active surveillance vs radical therapy. BJU Int. Oct 2012; 110(7): 1032-9. PMID 22260273
30. Johansson E, Steineck G, Holmberg L, et al. Long-term quality-of-life outcomes after radical prostatectomy or watchful waiting: the Scandinavian Prostate Cancer Group-4 randomised trial. Lancet Oncol. Sep 2011; 12(9): 891-9. PMID 21821474
31. Wu CL, Schroeder BE, Ma XJ, et al. Development and validation of a 32-gene prognostic index for prostate cancer progression. Proc Natl Acad Sci U S A. Apr 09 2013; 110(15): 6121-6. PMID 23533275
32. Spans L, Clinckemalie L, Helsen C, et al. The genomic landscape of prostate cancer. Int J Mol Sci. May 24 2013; 14(6): 10822-51. PMID 23708091

33. Schoenborn JR, Nelson P, Fang M. Genomic profiling defines subtypes of prostate cancer with the potential for therapeutic stratification. Clin Cancer Res. Aug 01 2013; 19(15): 4058-66. PMID 23704282
34. Huang J, Wang JK, Sun Y. Molecular pathology of prostate cancer revealed by next-generation sequencing: opportunities for genome-based personalized therapy. Curr Opin Urol. May 2013; 23(3): 189-93. PMID 23385974
35. Yu YP, Song C, Tseng G, et al. Genome abnormalities precede prostate cancer and predict clinical relapse. Am J Pathol. Jun 2012; 180(6): 2240-8. PMID 22569189
36. Agell L, Hernández S, Nonell L, et al. A 12-gene expression signature is associated with aggressive histological in prostate cancer: SEC14L1 and TCEB1 genes are potential markers of progression. Am J Pathol. Nov 2012; 181(5): 1585-94. PMID 23083832
37. Thompson I, Thrasher JB, Aus G, et al. Guideline for the management of clinically localized prostate cancer: 2007 update. J Urol. Jun 2007; 177(6): 2106-31. PMID 17509297
38. Kattan MW, Eastham JA, Wheeler TM, et al. Counseling men with prostate cancer: a nomogram for predicting the presence of small, moderately differentiated, confined tumors. J Urol. Nov 2003; 170(5): 1792-7. PMID 14532778
39. Cooperberg MR, Freedland SJ, Pasta DJ, et al. Multiinstitutional validation of the UCSF cancer of the prostate risk assessment for prediction of recurrence after radical prostatectomy. Cancer. Nov 15 2006; 107(10): 2384-91. PMID 17039503
40. Chen RC, Chang P, Vetter RJ, et al. Recommended patient-reported core set of symptoms to measure in prostate cancer treatment trials. J Natl Cancer Inst. Jul 2014; 106(7). PMID 25006192
41. Cuzick J, Berney DM, Fisher G, et al. Prognostic value of a cell cycle progression signature for prostate cancer death in a conservatively managed needle biopsy cohort. Br J Cancer. Mar 13 2012; 106(6): 1095-9. PMID 22361632
42. Cuzick J, Stone S, Fisher G, et al. Validation of an RNA cell cycle progression score for predicting death from prostate cancer in a conservatively managed needle biopsy cohort. Br J Cancer. Jul 28 2015; 113(3): 382-9. PMID 26103570
43. Lin DW, Crawford ED, Keane T, et al. Identification of men with low-risk biopsy-confirmed prostate cancer as candidates for active surveillance. Urol Oncol. Jun 2018; 36(6): 310.e7-310.e13. PMID 29655620
44. Montironi R, Mazzuccheli R, Scarpelli M, et al. Gleason grading of prostate cancer in needle biopsies or radical prostatectomy specimens: contemporary approach, current clinical significance and sources of pathology discrepancies. BJU Int. Jun 2005; 95(8): 1146-52. PMID 15877724
45. Tward JD, Schlomm T, Bardot S, et al. Personalizing Localized Prostate Cancer: Validation of a Combined Clinical Cell-cycle Risk (CCR) Score Threshold for Prognosticating Benefit From Multimodality Therapy. Clin Genitourin Cancer. Aug 2021; 19(4): 296-304.e3. PMID 33608228
46. Crawford ED, Scholz MC, Kar AJ, et al. Cell cycle progression score and treatment decisions in prostate cancer: results from an ongoing registry. Curr Med Res Opin. Jun 2014; 30(6): 1025-31. PMID 24576172
47. Shore N, Concepcion R, Saltzstein D, et al. Clinical utility of a biopsy-based cell cycle gene expression assay in localized prostate cancer. Curr Med Res Opin. Apr 2014; 30(4): 547-53. PMID 24320750
48. Shore ND, Kella N, Moran B, et al. Impact of the Cell Cycle Progression Test on Physician and Patient Treatment Selection for Localized Prostate Cancer. J Urol. Mar 2016; 195(3): 612-8. PMID 26403586
49. Schaink A, Li C, Wells D, et al. Prolaris Cell Cycle Progression Test for Localized Prostate Cancer: A Health Technology Assessment. Ont Health Technol Assess Ser. 2017; 17(6): 1-75. PMID 28572867
50. Klein EA, Cooperberg MR, Magi-Galluzzi C, et al. A 17-gene assay to predict prostate cancer aggressiveness in the context of Gleason grade heterogeneity, tumor multifocality, and biopsy undersampling. Eur Urol. Sep 2014; 66(3): 550-60. PMID 24836057

51. Cullen J, Rosner IL, Brand TC, et al. A Biopsy-based 17-gene Genomic Prostate Score Predicts Recurrence After Radical Prostatectomy and Adverse Surgical Pathology in a Racially Diverse Population of Men with Clinically Low- and Intermediate-risk Prostate Cancer. Eur Urol. Jul 2015; 68(1): 123-31. PMID 25465337
52. Whalen MJ, Hackert V, Rothberg MB, et al. Prospective correlation between likelihood of favorable pathology on the 17-Gene Genomic Prostate Score and actual pathological outcomes at radical prostatectomy. Urol Pract. Sep 2016;3(5):379-386.
53. Van Den Eeden SK, Lu R, Zhang N, et al. A Biopsy-based 17-gene Genomic Prostate Score as a Predictor of Metastases and Prostate Cancer Death in Surgically Treated Men with Clinically Localized Disease. Eur Urol. Jan 2018; 73(1): 129-138. PMID 28988753
54. Salmasi A, Said J, Shindel AW, et al. A 17-Gene Genomic Prostate Score Assay Provides Independent Information on Adverse Pathology in the Setting of Combined Multiparametric Magnetic Resonance Imaging Fusion Targeted and Systematic Prostate Biopsy. J Urol. Sep 2018; 200(3): 564-572. PMID 29524506
55. Cooperberg MR, Simko JP, Cowan JE, et al. Validation of a cell-cycle progression gene panel to improve risk stratification in a contemporary prostatectomy cohort. J Clin Oncol. Apr 10 2013; 31(11): 1428-34. PMID 23460710
56. McShane LM, Altman DG, Sauerbrei W, et al. Reporting recommendations for tumor marker prognostic studies. J Clin Oncol. Dec 20 2005; 23(36): 9067-72. PMID 16172462
57. Epstein JI, Allsbrook WC, Amin MB, et al. The 2005 International Society of Urological Pathology (ISUP) Consensus Conference on Gleason Grading of Prostatic Carcinoma. Am J Surg Pathol. Sep 2005; 29(9): 1228-42. PMID 16096414
58. Brand TC, Zhang N, Crager MR, et al. Patient-specific Meta-analysis of 2 Clinical Validation Studies to Predict Pathologic Outcomes in Prostate Cancer Using the 17-Gene Genomic Prostate Score. Urology. Mar 2016; 89: 69-75. PMID 26723180
59. Albala D, Kemeter MJ, Febbo PG, et al. Health Economic Impact and Prospective Clinical Utility of Oncotype DX® Genomic Prostate Score. Rev Urol. 2016; 18(3): 123-132. PMID 27833462
60. Eure G, Germany R, Given R, et al. Use of a 17-Gene Prognostic Assay in Contemporary Urologic Practice: Results of an Interim Analysis in an Observational Cohort. Urology. Sep 2017; 107: 67-75. PMID 28454985
61. Badani KK, Kemeter MJ, Febbo PG, et al. The impact of a biopsy based 17-Gene Genomic Prostate Score on treatment recommendations in men with newly diagnosed clinically prostate cancer who are candidates for active surveillance. Urol Pract. 2015;2(4):181-189. PMID not Indexed in Pubmed
62. Canfield SK, M.J.; Febbo, P.G.; Hornberger, J. Balancing confounding and generalizability using observational, real-world data: 17-gene genomic prostate score assay effect on active surveillance. Rev Urol. 2018;20(2):69-76.
63. Canfield S, Kemeter MJ, Hornberger J, et al. Active Surveillance Use Among a Low-risk Prostate Cancer Population in a Large US Payer System: 17-Gene Genomic Prostate Score Versus Other Risk Stratification Methods. Rev Urol. 2017; 19(4): 203-212. PMID 29472824
64. Nguyen AM, Carter GC, Wilson LM, et al. Real-world utilization, patient characteristics, and treatment patterns among men with localized prostate cancer tested with the 17-gene genomic prostate score® (GPS TM) assay. Prostate. Jul 2024; 84(10): 922-931. PMID 38666513
65. Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. Med Decis Making. 2006; 26(6): 565-74. PMID 17099194
66. Cooperberg MR, Broering JM, Carroll PR. Risk assessment for prostate cancer metastasis and mortality at the time of diagnosis. J Natl Cancer Inst. Jun 16 2009; 101(12): 878-87. PMID 19509351
67. Ross AE, Iwata KK, Elsouda D, et al. Transcriptome-Based Prognostic and Predictive Biomarker Analysis of ENACT: A Randomized Controlled Trial of Enzalutamide in Men Undergoing Active Surveillance. JCO Precis Oncol. Apr 2024; 8: e2300603. PMID 38635932

68. Berlin A, Murgic J, Hosni A, et al. Genomic Classifier for Guiding Treatment of Intermediate-Risk Prostate Cancers to Dose-Escalated Image Guided Radiation Therapy Without Hormone Therapy. Int J Radiat Oncol Biol Phys. Jan 01 2019; 103(1): 84-91. PMID 30170099
69. Nguyen PL, Shin H, Yousefi K, et al. Impact of a Genomic Classifier of Metastatic Risk on Postprostatectomy Treatment Recommendations by Radiation Oncologists and Urologists. Urology. Jul 2015; 86(1): 35-40. PMID 26142578
70. Tosoian JJ, Birer SR, Jeffrey Karnes R, et al. Performance of clinicopathologic models in men with high risk localized prostate cancer: impact of a 22-gene genomic classifier. Prostate Cancer Prostatic Dis. Dec 2020; 23(4): 646-653. PMID 32231245
71. Zhu A, Proudfoot JA, Davicioni E, et al. Use of Decipher Prostate Biopsy Test in Patients with Favorable-risk Disease Undergoing Conservative Management or Radical Prostatectomy in the Surveillance, Epidemiology, and End Results Registry. Eur Urol Oncol. Jul 06 2024. PMID 38972832
72. Blume-Jensen P, Berman DM, Rimm DL, et al. Development and clinical validation of an in situ biopsy-based multimarker assay for risk stratification in prostate cancer. Clin Cancer Res. Jun 01 2015; 21(11): 2591-600. PMID 25733599
73. Esteva A, Feng J, van der Wal D, et al. Prostate cancer therapy personalization via multi-modal deep learning on randomized phase III clinical trials. NPJ Digit Med. Jun 08 2022; 5(1): 71. PMID 35676445
74. Spratt DE, Tang S, Sun Y, et al. Artificial Intelligence Predictive Model for Hormone Therapy Use in Prostate Cancer. NEJM Evid. Aug 2023; 2(8): EVIDoa2300023. PMID 38320143
75. Gerrard P, Zhang J, Yamashita R, et al. Analytical Validation of a Clinical Grade Prognostic and Classification Artificial Intelligence Laboratory Test for Men with Prostate Cancer. AI in Precision Oncology 2024 1:2, 119-126.
76. Spratt DE, Liu VYT, Jia AY, et al. Meta-analysis of Individual Patient-level Data for a Multimodal Artificial Intelligence Biomarker in High-risk Prostate Cancer: Results from Six NRG/RTOG Phase 3 Randomized Trials. Eur Urol. Oct 2024; 86(4): 369-371. PMID 39025748
77. Ross AE, Zhang J, Huang HC, et al. External Validation of a Digital Pathology-based Multimodal Artificial Intelligence Architecture in the NRG/RTOG 9902 Phase 3 Trial. Eur Urol Oncol. Oct 2024; 7(5): 1024-1033. PMID 38302323
78. Tward JD, Huang HC, Esteva A, et al. Prostate Cancer Risk Stratification in NRG Oncology Phase III Randomized Trials Using Multimodal Deep Learning With Digital Histopathology. JCO Precis Oncol. Oct 2024; 8: e2400145. PMID 39447096
79. D'Amico AV, Whittington R, Malkowicz SB, et al. Biochemical outcome after radical prostatectomy, external beam radiation therapy, or interstitial radiation therapy for clinically localized prostate cancer. JAMA. Sep 16 1998; 280(11): 969-74. PMID 9749478
80. Fossati N, Karnes RJ, Boorjian SA, et al. Long-term Impact of Adjuvant Versus Early Salvage Radiation Therapy in pT3N0 Prostate Cancer Patients Treated with Radical Prostatectomy: Results from a Multi-institutional Series. Eur Urol. Jun 2017; 71(6): 886-893. PMID 27484843
81. Hwang WL, Tendulkar RD, Niemierko A, et al. Comparison Between Adjuvant and Early-Salvage Postprostatectomy Radiotherapy for Prostate Cancer With Adverse Pathological Features. JAMA Oncol. May 10 2018; 4(5): e175230. PMID 29372236
82. Buscariollo DL, Drumm M, Niemierko A, et al. Long-term results of adjuvant versus early salvage postprostatectomy radiation: A large single-institutional experience. Pract Radiat Oncol. 2017; 7(2): e125-e133. PMID 28274403
83. Freedland SJ, Rumble RB, Finelli A, et al. Adjuvant and salvage radiotherapy after prostatectomy: American Society of Clinical Oncology clinical practice guideline endorsement. J Clin Oncol. Dec 01 2014; 32(34): 3892-8. PMID 25366677
84. Stephenson AJ, Scardino PT, Kattan MW, et al. Predicting the outcome of salvage radiation therapy for recurrent prostate cancer after radical prostatectomy. J Clin Oncol. May 20 2007; 25(15): 2035-41. PMID 17513807
85. Stephenson AJ, Scardino PT, Eastham JA, et al. Postoperative nomogram predicting the 10-year probability of prostate cancer recurrence after radical prostatectomy. J Clin Oncol. Oct 01 2005; 23(28): 7005-12. PMID 16192588

86. Cooperberg MR, Hilton JF, Carroll PR. The CAPRA-S score: A straightforward tool for improved prediction of outcomes after radical prostatectomy. Cancer. Nov 15 2011; 117(22): 5039-46. PMID 21647869
87. Cuzick J, Swanson GP, Fisher G, et al. Prognostic value of an RNA expression signature derived from cell cycle proliferation genes in patients with prostate cancer: a retrospective study. Lancet Oncol. Mar 2011; 12(3): 245-55. PMID 21310658
88. Bishoff JT, Freedland SJ, Gerber L, et al. Prognostic utility of the cell cycle progression score generated from biopsy in men treated with prostatectomy. J Urol. Aug 2014; 192(2): 409-14. PMID 24508632
89. Swanson GP, Lenz L, Stone S, et al. Cell-cycle risk score more accurately determines the risk for metastases and death in prostatectomy patients compared with clinical features alone. Prostate. Mar 2021; 81(4): 261-267. PMID 33475174
90. Koch MO, Cho JS, Kaimakliotis HZ, et al. Use of the cell cycle progression (CCP) score for predicting systemic disease and response to radiation of biochemical recurrence. Cancer Biomark. Jun 07 2016; 17(1): 83-8. PMID 27314296
91. Freedland SJ, Gerber L, Reid J, et al. Prognostic utility of cell cycle progression score in men with prostate cancer after primary external beam radiation therapy. Int J Radiat Oncol Biol Phys. Aug 01 2013; 86(5): 848-53. PMID 23755923
92. Klein EA, Yousefi K, Haddad Z, et al. A genomic classifier improves prediction of metastatic disease within 5 years after surgery in node-negative high-risk prostate cancer patients managed by radical prostatectomy without adjuvant therapy. Eur Urol. Apr 2015; 67(4): 778-86. PMID 25466945
93. Sommariva S, Tarricone R, Lazzeri M, et al. Prognostic Value of the Cell Cycle Progression Score in Patients with Prostate Cancer: A Systematic Review and Meta-analysis. Eur Urol. Jan 2016; 69(1): 107-15. PMID 25481455
94. Morgan TM, Daignault-Newton S, Spratt DE, et al. Impact of Gene Expression Classifier Testing on Adjuvant Treatment Following Radical Prostatectomy: The G-MINOR Prospective Randomized Cluster-crossover Trial. Eur Urol. Feb 2025; 87(2): 228-237. PMID 39379238
95. Feng FY, Huang HC, Spratt DE, et al. Validation of a 22-Gene Genomic Classifier in Patients With Recurrent Prostate Cancer: An Ancillary Study of the NRG/RTOG 9601 Randomized Clinical Trial. JAMA Oncol. Apr 01 2021; 7(4): 544-552. PMID 33570548
96. Den RB, Yousefi K, Trabulsi EJ, et al. Genomic classifier identifies men with adverse pathology after radical prostatectomy who benefit from adjuvant radiation therapy. J Clin Oncol. Mar 10 2015; 33(8): 944-51. PMID 25667284
97. Den RB, Feng FY, Showalter TN, et al. Genomic prostate cancer classifier predicts biochemical failure and metastases in patients after postoperative radiation therapy. Int J Radiat Oncol Biol Phys. Aug 01 2014; 89(5): 1038-1046. PMID 25035207
98. Cooperberg MR, Davicioni E, Crisan A, et al. Combined value of validated clinical and genomic risk stratification tools for predicting prostate cancer mortality in a high-risk prostatectomy cohort. Eur Urol. Feb 2015; 67(2): 326-33. PMID 24998118
99. Ross AE, Feng FY, Ghadessi M, et al. A genomic classifier predicting metastatic disease progression in men with biochemical recurrence after prostatectomy. Prostate Cancer Prostatic Dis. Mar 2014; 17(1): 64-9. PMID 24145624
100. Karnes RJ, Bergstralh EJ, Davicioni E, et al. Validation of a genomic classifier that predicts metastasis following radical prostatectomy in an at risk patient population. J Urol. Dec 2013; 190(6): 2047-53. PMID 23770138
101. Erho N, Crisan A, Vergara IA, et al. Discovery and validation of a prostate cancer genomic classifier that predicts early metastasis following radical prostatectomy. PLoS One. 2013; 8(6): e66855. PMID 23826159
102. Ross AE, Johnson MH, Yousefi K, et al. Tissue-based Genomics Augments Post-prostatectomy Risk Stratification in a Natural History Cohort of Intermediate- and High-Risk Men. Eur Urol. Jan 2016; 69(1): 157-65. PMID 26058959

103. Freedland SJ, Choeurng V, Howard L, et al. Utilization of a Genomic Classifier for Prediction of Metastasis Following Salvage Radiation Therapy after Radical Prostatectomy. Eur Urol. Oct 2016; 70(4): 588-596. PMID 26806658
104. Glass AG, Leo MC, Haddad Z, et al. Validation of a Genomic Classifier for Predicting Post-Prostatectomy Recurrence in a Community Based Health Care Setting. J Urol. Jun 2016; 195(6): 1748-53. PMID 26626216
105. Klein EA, Haddad Z, Yousefi K, et al. Decipher Genomic Classifier Measured on Prostate Biopsy Predicts Metastasis Risk. Urology. Apr 2016; 90: 148-52. PMID 26809071
106. Spratt DE, Dai DLY, Den RB, et al. Performance of a Prostate Cancer Genomic Classifier in Predicting Metastasis in Men with Prostate-specific Antigen Persistence Postprostatectomy. Eur Urol. Jul 2018; 74(1): 107-114. PMID 29233664
107. Karnes RJ, Choeurng V, Ross AE, et al. Validation of a Genomic Risk Classifier to Predict Prostate Cancer-specific Mortality in Men with Adverse Pathologic Features. Eur Urol. Feb 2018; 73(2): 168-175. PMID 28400167
108. Ross AE, Den RB, Yousefi K, et al. Efficacy of post-operative radiation in a prostatectomy cohort adjusted for clinical and genomic risk. Prostate Cancer Prostatic Dis. Sep 2016; 19(3): 277-82. PMID 27136742
109. Lobo JM, Dicker AP, Buerki C, et al. Evaluating the clinical impact of a genomic classifier in prostate cancer using individualized decision analysis. PLoS One. 2015; 10(3): e0116866. PMID 25837660
110. West TA, Kiely BE, Stockler MR. Estimating scenarios for survival time in men starting systemic therapies for castration-resistant prostate cancer: a systematic review of randomised trials. Eur J Cancer. Jul 2014; 50(11): 1916-24. PMID 24825113
111. Scher HI, Graf RP, Schreiber NA, et al. Nuclear-specific AR-V7 Protein Localization is Necessary to Guide Treatment Selection in Metastatic Castration-resistant Prostate Cancer. Eur Urol. Jun 2017; 71(6): 874-882. PMID 27979426
112. Armstrong AJ, Halabi S, Luo J, et al. Prospective Multicenter Validation of Androgen Receptor Splice Variant 7 and Hormone Therapy Resistance in High-Risk Castration-Resistant Prostate Cancer: The PROPHECY Study. J Clin Oncol. May 01 2019; 37(13): 1120-1129. PMID 30865549
113. Sanda MG, Cadeddu JA, Kirkby E, et al. Clinically Localized Prostate Cancer: AUA/ASTRO/SUO Guideline. Part I: Risk Stratification, Shared Decision Making, and Care Options. J Urol. Mar 2018; 199(3): 683-690. PMID 29203269
114. Sanda MG, Cadeddu JA, Kirkby E, et al. Clinically Localized Prostate Cancer: AUA/ASTRO/SUO Guideline. Part II: Recommended Approaches and Details of Specific Care Options. J Urol. Apr 2018; 199(4): 990-997. PMID 29331546
115. Li EV, Ren Y, Griffin J, et al. An Artificial Intelligence-Digital Pathology Algorithm Predicts Survival After Radical Prostatectomy From the Prostate, Lung, Colorectal, and Ovarian Cancer Trial. J Urol. Jan 22 2025: 101097JU0000000000004435. PMID 39841869
116. Antonarakis ES, Lu C, Wang H, et al. AR-V7 and resistance to enzalutamide and abiraterone in prostate cancer. N Engl J Med. Sep 11 2014; 371(11): 1028-38. PMID 25184630
117. Scher HI, Lu D, Schreiber NA, et al. Association of AR-V7 on Circulating Tumor Cells as a Treatment-Specific Biomarker With Outcomes and Survival in Castration-Resistant Prostate Cancer. JAMA Oncol. Nov 01 2016; 2(11): 1441-1449. PMID 27262168
118. Scher HI, Graf RP, Schreiber NA, et al. Assessment of the Validity of Nuclear-Localized Androgen Receptor Splice Variant 7 in Circulating Tumor Cells as a Predictive Biomarker for Castration-Resistant Prostate Cancer. JAMA Oncol. Sep 01 2018; 4(9): 1179-1186. PMID 29955787
119. Feng FY, Smith MR, Saad F, et al. Digital Pathology-Based Multimodal Artificial Intelligence Scores and Outcomes in a Randomized Phase III Trial in Men With Nonmetastatic Castration-Resistant Prostate Cancer. JCO Precis Oncol. Jan 2025; 9: e2400653. PMID 39889245
120. Markowski MC, Ren Y, Tierney M, et al. Digital Pathology-based Artificial Intelligence Biomarker Validation in Metastatic Prostate Cancer. Eur Urol Oncol. Dec 10 2024. PMID 39665917

121. Wang JH, Deek MP, Mendes AA, et al. Validation of an artificial intelligence-based prognostic biomarker in patients with oligometastatic Castration-Sensitive prostate cancer. Radiother Oncol. Jan 2025; 202: 110618. PMID 39510141
122. Eggener SE, Rumble RB, Armstrong AJ, et al. Molecular Biomarkers in Localized Prostate Cancer: ASCO Guideline. J Clin Oncol. May 01 2020; 38(13): 1474-1494. PMID 31829902
123. Lowrance WT, Murad MH, Oh WK, et al. Castration-Resistant Prostate Cancer: AUA Guideline Amendment 2018. J Urol. Dec 2018; 200(6): 1264-1272. PMID 30086276
124. National Institute for Health and Care Excellence (NICE). Prostate cancer: diagnosis and management [NG131]. 2019 (Updated 2021); https://www.nice.org.uk/guidance/ng131. Accessed March 18, 2025.
125. Centers for Medicare & Medicaid Services. (2025, July 3). Local Coverage Determination (LCD): MolDX: Prostate Cancer Genomic Classifier Assay for Men with Localized Disease (L38341). U.S. Department of Health & Human Services. https://www.cms.gov/medicare-coverage-database/view/lcd.aspx?lcdid=38341&ver=13&=. Accessed September 17, 2025.
126. Centers for Medicare & Medicaid Services. (2021, January 1). Local Coverage Determination (LCD) Reference Article: Billing and Coding: MolDX: Prostate Cancer Genomic Classifier Assay for Men with Localized Disease (A57236). U.S. Department of Health & Human Services. https://www.cms.gov/medicare-coverage-database/view/article.aspx?articleId=57236&ver=18. Accessed September 17, 2025.

Documentation for Clinical Review

Please provide the following documentation:

- History and physical and/or consultation notes including:
  - Clinical stage (TNM)
  - NCCN risk category
  - Biopsy Gleason score
  - PSA level/density
  - Number of biopsy cores with presence of disease including cancer involvement
  - Pertinet comorbidities
- Test requested and reason for test
- Projected life expectancy
- Documentation that individual is a candidate for active surveillance or definitive therapy
- Prior treatment (if applicable) including prostatectomy if applicable

Post Service (in addition to the above, please include the following):

- Laboratory report/results

Coding

The list of codes in this Medical Policy is intended as a general reference and may not cover all codes. Inclusion or exclusion of a code(s) does not constitute or imply member coverage or provider reimbursement policy.

Type	Code	Description
CPT®	0005U	Oncology (prostate) gene expression profile by real-time RT-PCR of 3 genes (ERG, PCA3, and SPDEF), urine, algorithm reported as risk score
	0047U	Oncology (prostate), mRNA, gene expression profiling by real-time RT-PCR of 17 genes (12 content and 5 housekeeping), utilizing formalin-fixed paraffin-embedded tissue, algorithm reported as a risk score
	0376U	Oncology (prostate cancer), image analysis of at least 128 histologic features and clinical factors, prognostic algorithm determining the risk of distant metastases, and prostate cancer-specific mortality, includes

Type	Code	Description
		predictive algorithm to androgen deprivation-therapy response, if appropriate (ArteraAI Prostate Test from Artera Inc)
	0497U	Oncology (prostate), mRNA gene-expression profiling by real-time RT-PCR of 6 genes (FOXM1, MCM3, MTUS1, TTC21B, ALAS1, and PPP2CA), utilizing formalin-fixed paraffin-embedded (FFPE) tissue, algorithm reported as a risk score for prostate cancer
	81541	Oncology (prostate), mRNA gene expression profiling by real-time RT-PCR of 46 genes (31 content and 15 housekeeping), utilizing formalin-fixed paraffin-embedded tissue, algorithm reported as a disease-specific mortality risk score
	81542	Oncology (prostate), mRNA, microarray gene expression profiling of 22 content genes, utilizing formalin-fixed paraffin-embedded tissue, algorithm reported as metastasis risk score
HCPCS	None	

Policy History

This section provides a chronological history of the activities, updates and changes that have occurred with this Medical Policy.

Effective Date	Action
03/01/2016	BCBSA Medical Policy adoption
02/01/2017	Policy revision without position change
01/01/2018	Policy revision without position change Coding update
05/01/2018	Coding update
11/01/2018	Policy revision with position change
01/01/2019	Policy revision without position change
03/01/2020	Coding update
04/01/2020	Annual review. No change to policy statement.
01/01/2021	Annual review. No change to policy statement. Literature review updated.
12/01/2021	Administrative update. Policy statement updated.
02/01/2021	Annual review. Literature review updated. Policy statement and literature updated.
01/01/2023	Annual review. No change to policy statement. Literature review updated.
10/01/2025	Policy reactivated. Previously archived from 08/01/2023 to 09/30/2025. Policy title changed from Gene Expression Profiling and Protein Biomarkers for Prostate Cancer Management to current one. Policy statement clarification.

Definitions of Decision Determinations

Healthcare Services: For the purpose of this Medical Policy, Healthcare Services means procedures, treatments, supplies, devices, and equipment.

Medically Necessary: Healthcare Services that are Medically Necessary include only those which have been established as safe and effective, are furnished under generally accepted professional standards to treat illness, injury or medical condition, and which, as determined by Blue Shield of California, are: (a) consistent with Blue Shield of California medical policy; (b) consistent with the symptoms or diagnosis; (c) not furnished primarily for the convenience of the patient, the attending Physician or other provider; (d) furnished at the most appropriate level which can be provided safely and effectively to the member; and (e) not more costly than an alternative service or sequence of

services at least as likely to produce equivalent therapeutic or diagnostic results as to the diagnosis or treatment of the member's illness, injury, or disease.

Investigational or Experimental: Healthcare Services which do not meet ALL of the following five (5) elements are considered investigational or experimental:

A. The technology must have final approval from the appropriate government regulatory bodies.
   - This criterion applies to drugs, biological products, devices and any other product or procedure that must have final approval to market from the U.S. Food and Drug Administration ("FDA") or any other federal governmental body with authority to regulate the use of the technology.
   - Any approval that is granted as an interim step in the FDA's or any other federal governmental body's regulatory process is not sufficient.
   - The indications for which the technology is approved need not be the same as those which Blue Shield of California is evaluating.

B. The scientific evidence must permit conclusions concerning the effect of the technology on health outcomes.
   - The evidence should consist of well-designed and well-conducted investigations published in peer-reviewed journals. The quality of the body of studies and the consistency of the results are considered in evaluating the evidence.
   - The evidence should demonstrate that the technology can measure or alter the physiological changes related to a disease, injury, illness, or condition. In addition, there should be evidence, or a convincing argument based on established medical facts that such measurement or alteration affects health outcomes.

C. The technology must improve the net health outcome.
   - The technology's beneficial effects on health outcomes should outweigh any harmful effects on health outcomes.

D. The technology must be as beneficial as any established alternatives.
   - The technology should improve the net health outcome as much as, or more than, established alternatives.

E. The improvement must be attainable outside the investigational setting.
   - When used under the usual conditions of medical practice, the technology should be reasonably expected to satisfy Criteria C and D.

Feedback

Blue Shield of California is interested in receiving feedback relative to developing, adopting, and reviewing criteria for medical policy. Any licensed practitioner who is contracted with Blue Shield of California or Blue Shield of California Promise Health Plan is welcome to provide comments, suggestions, or concerns. Our internal policy committees will receive and take your comments into consideration. Our medical policies are available to view or download at www.blueshieldca.com/provider.

For medical policy feedback, please send comments to: MedPolicy@blueshieldca.com

Questions regarding the applicability of this policy should be directed to the Prior Authorization Department at (800) 541-6652, or the Transplant Case Management Department at (800) 637-2066 ext. 3507708 or visit the provider portal at www.blueshieldca.com/provider.

Disclaimer: Blue Shield of California may consider published peer-reviewed scientific literature, national guidelines, and local standards of practice in developing its medical policy. Federal and state law, as well as member health services contract language, including definitions and specific contract provisions/exclusions, take precedence over medical policy and must be considered first in determining covered services. Member health

services contracts may differ in their benefits. Blue Shield reserves the right to review and update policies as appropriate.

Appendix A

POLICY STATEMENT	
BEFORE	**AFTER** **Blue font: Verbiage Changes/Additions**
Reactivated Policy **Policy Statement:** N/A	**Gene Expression Profiling, Protein Biomarkers, and Multimodal Artificial Intelligence for Prostate Cancer Management 2.04.111** **Policy Statement:** I. Use of gene expression analysis, protein biomarkers, and multimodal artificial intelligence (MMAI) to guide management of prostate cancer is considered **investigational** in all situations. Note: For individuals enrolled in health plans subject to the Biomarker Testing Law (Health & Safety Code Section 1367.667 and the Insurance Code Section 10123.209), Centers for Medicare & Medicaid Services (CMS) Local Coverage Determination (LCD) may also apply. Please refer to the Medicare National and Local Coverage section of this policy and to MolDX: Prostate Cancer Genomic Classifier Assay for Men with Localized Disease for reference.