Multigene Expression Assay for Predicting Recurrence in Colon Cancer

I. Description

Gene expression profiling (GEP) tests have been developed for use as prognostic markers in stage 2 or 3 colon cancer to help identify patients who are at high risk for recurrent disease and could be candidates for adjuvant chemotherapy.

For individuals who have stage 2 or 3 colon cancer who receive GEP testing, the evidence includes development and validation studies and 1 decision-impact study. Relevant outcomes are disease-specific survival, test accuracy and validity, and change in disease status. The available evidence has shown that GEP tests for colon cancer can improve risk prediction, particularly the risk of recurrence in patients with stage 2 or 3 colon cancer. However, evidence to date is insufficient to permit conclusions on how GEP classification compares with other approaches for identifying recurrence risk in stage 2 or 3 patients. The indirect chain of evidence that demonstrates GEP testing would improve health outcomes is weak. Studies showing management changes as a consequence of testing do not demonstrate whether such changes improve outcomes. The evidence is insufficient to determine the effects of the technology on health outcomes.

II. Criteria/Guidelines

Gene expression assays for determining the prognosis of stage 2 or 3 colon cancer following surgery are not covered as they are not known to be effective in improving health outcomes.

III. Policy Guidelines

Genetic Counseling

Genetic counseling is primarily aimed at patients who are at risk for inherited disorders, and experts recommend formal genetic counseling in most cases when genetic testing for an inherited
condition is considered. The interpretation of the results of genetic tests and the understanding of risk factors can be very difficult and complex. Therefore, genetic counseling will assist individuals in understanding the possible benefits and harms of genetic testing, including the possible impact of the information on the individual’s family. Genetic counseling may alter the utilization of genetic testing substantially and may reduce inappropriate testing. Genetic counseling should be performed by an individual with experience and expertise in genetic medicine and genetic testing methods.

Coding

CPT code specific to Oncotype DX Colon Cancer Assay –

- 81525 Oncology (colon), mRNA, gene expression profiling by real-time RT-PCR of 12 genes (7 content and 5 housekeeping), utilizing formalin-fixed paraffin-embedded tissue, algorithm reported as a recurrence score

For the other tests in the policy, if the test is a multianalyte assay with algorithmic analysis (MAAA), it would be reported with the unlisted MAAA code – 81599. Otherwise, it would likely be reported using an unlisted code such as:

- 84999 Unlisted chemistry procedure
- 88299 Unlisted cytogenetic study

IV. Background

Of patients with stage II colon cancer, 75–80% are cured by surgery alone, and the absolute benefit of chemotherapy for the overall patient population is small. Patients most likely to benefit from chemotherapy are difficult to identify by standard clinical and pathologic risk factors. Genomic tests are intended to facilitate identifying stage 2 patients most likely to experience recurrence after surgery and most likely to benefit from additional treatment.

Colorectal cancer is classified as stage 2 (also called Dukes B) when it has spread outside the colon and/or rectum to nearby tissue but is not detectable in lymph nodes (stage 3 disease, also called Dukes C) and has not metastasized to distant sites stage 4 disease). Primary treatment is surgical resection of the primary cancer and colonic anastomosis. After surgery prognosis is good, with survival rates of 75% to 80% at 5 years. A 2008 meta-analysis of 50 studies of adjuvant therapy versus surgery alone in all stage II patients found statistically significant, though small, absolute benefit of chemotherapy for disease-free survival (DFS) but not for overall survival. Therefore, adjuvant chemotherapy with 5-fluorouracil or capecitabine is recommended only for resected patients with high-risk stage II disease (ie those with poor prognostic features). However, clinical and pathologic features used to identify high-risk disease are not well-established, and patients for whom benefits of adjuvant chemotherapy would most likely outweigh harms cannot be identified with certainty. The current system relies on a variety of factors including tumor substage IIB (T4A tumors that invade the muscularis propria and extend into pericolorectal tissues) or IIC (T4B tumors that invade or are adherent to other organs or structures), obstruction or bowel perforation at initial diagnosis, an inadequately low number of sampled lymph nodes at surgery (12 or less), histologic features of aggressiveness, a high preoperative carcinoembryonic antigen level, and indeterminate or positive resection margins.
For patients with stage 3 colon cancer, current guidelines from the National Comprehensive Cancer Network recommend “6 months of adjuvant chemotherapy after primary surgical treatment.” However, some have questioned the benefit of adjuvant chemotherapy in subsets of patients with stage 3 disease (eg, stage 3A) whose predicted survival may actually exceed that of some stage 2 patients (eg, stage 2C).

Of interest, a recent review has noted that microsatellite instability (MSI) and mismatch repair (MMR) deficiency in colon cancer may represent confounding factors to be considered in treatment. These factors may identify a minority (15% to 20%) of the population with improved DFS who may derive no benefit or may exhibit deleterious effects from adjuvant fluorouracil/leucovorin-based treatments. Patient MSI and MMR status may be critically important in how to study, interpret, and use a particular gene expression profiling test.

Regulatory Status
Clinical laboratories may develop and validate tests in-house and market them as a laboratory service; laboratory-developed tests (LDTs) must meet the general regulatory standards of the Clinical Laboratory Improvement Amendments (CLIA). Multigene expression assay testing for predicting recurrent colon cancer is available under the auspices of CLIA. Laboratories that offer LDTs must be licensed by CLIA for high-complexity testing. To date, the U.S. Food and Drug Administration has chosen not to require any regulatory review of this test.

V. Rationale
Validation of genotyping to improve treatment outcomes is a multistep process. In general, important steps in the validation process address the following:

- **Analytic validity**: measures technical performance (ie, whether the test accurately and reproducibly detects the gene markers of interest).
- **Clinical validity**: measures the strength of associations between selected genetic markers and clinical status.
- **Clinical utility**: determines whether the use of genotyping for specific genetic markers to guide treatment decisions improves patient outcomes (eg, survival or adverse event rates) compared with standard treatment without genotyping.

An updated literature update covers the period through July 10, 2016.

Analytical Validity
Many gene expression profiling (GEP) assays have been developed and reported for use as prognostic markers in stage II colon cancer since 2004. Four are currently offered commercially in the United States (ColoPrint 18-Gene Colon Cancer Recurrence Assay, GeneFx Colon, OncoDefender-CRC, Oncotype DX Colon Recurrence Score). Information on specimen type, sample handling, and technique used for GEP has been reported for many of these assays.

Clinical Validity
ColonPRS: Van Laar (2010) reported on a 163-gene expression test using data from 232 colon cancer patients across all stages (I to IV) of disease (training set). Patients were stratified into high- and low-risk groups, and a second validation was performed in 33 stage II and 27 stage III patients (test set). Among stage II patients, 5-year disease-free survival (DFS) was statistically and significantly prolonged in low-risk compared with high-risk patients; but among stage III patients, 5-year DFS did not differ statistically between low- and high-risk groups. ColonPRS is marketed for research use only. The test was originally marketed by Signal Genetics. Direct contact with Signal Genetics in 2015 confirmed that it is not currently marketing ColonPRS.

ColoPrint 18-Gene Colon Cancer Recurrence Assay: Salazar et al (2011) described the development of an 18-gene expression test called the ColoPrint 18-Gene Colon Cancer Recurrence Assay. A total of 188 samples were prospectively collected from patients with colorectal cancer (CRC). RNA was isolated from fresh tissue frozen in liquid nitrogen, labeled and hybridized to customized whole-genome oligonucleotide high-density microarrays. A cross-validation procedure was performed on 33,834 gene probes that showed variation across the training samples. They were scored for their association with 5-year distant metastasis-free survival. From this pool of genes, an optimal set of 18 nonredundant probes was identified and used to construct classification scores for the test. Results were dichotomized into a 2-category low- and high-risk and low-risk scoring system.

In a small independent validation study using a patient cohort of 206 patients, 60% of patients were identified as low risk and 40% as high risk. The population studied, however, had a mix of patients with different disease stages with only 56% representing stage II tumors. In the evaluation of patients with stage II disease, 63.2% were classified as low risk (5-year recurrence-free survival [RFS], 90.9%) and 36.8% were classified as high risk (5-year RFS, 73.9%).

Maak et al (2013) conducted a subsequent validation study in fresh frozen tumor samples from 135 patients who had undergone curative resection for stage II colon cancer. Mismatch repair (MMR) status, clinical parameters, and follow-up data (median 8.4 years) were collected. Five-year distant metastasis-free survival was 95% for patients classified as low risk by ColoPrint and 80% for patients classified as high risk. Information about net reclassification and clinical utility was not provided. To date, larger validation studies have been published only in abstract form.

In 2015, Kopetz et al23 reported a pooled analysis of patients with stage II colon cancer from independent cohorts in the United States, Spain, Italy, Austria, and Germany. Of 416 patients in the pooled dataset, 124 (30%) received adjuvant 5-fluorouracil (5-FU)–based chemotherapy. Investigators compared the prognostic ability of ColoPrint with National Comprehensive Cancer Network (NCCN) risk prediction based on clinicopathologic factors (T4; high-grade tumor; lymphovascular or perineural invasion; perforation or obstruction; <12 lymph nodes examined; and positive margins). ColoPrint classified 263 (63%) patients as low risk and 153 (37%) patients as high risk. NCCN classified 236 (57%) patients as low risk and 180 (43%) patients as high risk. At median follow-up of 81 months (range, 56–178 months), 5-year recurrence risks in ColoPrint low- and high-risk groups were 10% (95% confidence interval [CI], 7% to 14%) and 21% (95% CI, 14% to 28%), respectively. In NCCN low- and high-risk groups, 5-year recurrence risks were 13% (95% CI, 9% to
18%) and 15% (95% CI, 10% to 20%), respectively. Statistical comparison of the risk models (eg, using a likelihood ratio test and/or receiver operating characteristic [ROC] curves) and comparison of classifications by survival outcomes (ie, reclassification analysis) were not provided. Further, a 5-year recurrence risk as high as 14% in patients classified as low risk by ColoPrint may be too high for some patients to consider foregoing chemotherapy.

Genefx Colon: Kennedy et al (2011) reported on the development of a 634-probe set signature. A training set of 215 patients (142 low risk, 73 high risk) was identified based on 5-year DFS. The assay was performed using DNA-microarray analysis of formalin-fixed paraffin-embedded (FFPE) samples. Cross-validation studies were used to select an optimal transcript signature for prognostic classification.

Independent validation was performed on 144 patients enriched for recurrence (85 low-risk, 59 high-risk) using the threshold score identified in the training set. The signature in this convenience sample of patients predicted disease recurrence with a hazard ratio (HR) of 2.53 (p<0.001) in the high-risk group. The signature also predicted cancer-related death with an HR of 2.21 (p=0.001) in the high-risk group. The authors noted that additional retrospective validation of the test in a large cohort of stage II colon cancer samples collected as part of a clinical trial was planned.

OncoDefender - CRC: Lenehan et al (2012) reported on their development of a 5-gene test, OncoDefender. A total of 417 cancer-associated genes were preselected for study in archived FFPE primary adenocarcinoma tissues of 74 patients with CRC (15 with stage I disease and 59 with stage II disease; 60 with colon, 14 with rectal cancer). Patients were divided into a training set and a test set. Cross validation was performed to estimate the ability of the classifier to generalize to unseen samples. The most important feature of gene fitness was the area under the ROC curve for each gene.

External validation was performed on 251 patients with stage I and II colon cancer obtained from an international study set. Patient dropout from the set of archived samples used was substantial; only 264 (55%) of 484 patients with lymph-node negative CRC satisfied the initial clinicopathologic screening. This included a mix of patients with both rectal and colon cancer (stages I and II). The test appeared to distinguish patients at high- versus low-risk of recurrence with an HR of 1.63 (p=0.031). Sensitivity and specificity of OncoDefender was compared to with NCCN guidelines and showed similar sensitivity (69% vs. 73%), with improved specificity (48% vs 26%). However, isolated performance of the test in patients with stage II colon cancer was not reported, and several NCCN high-risk findings (bowel obstruction/perforation and lymphovascular invasion) demonstrated higher HRs than observed with the molecular signature. The study alluded to but did not directly address clinical utility.

Oncotype DX® Colon Recurrence Score: O’Connell et al (2010) described the development of a 12-gene expression test, Oncotype DX® Colon Recurrence Score. A total of 761 candidate genes of possible prognostic value for recurrence or of possible predictive value for treatment were examined by correlating the genes in tumor samples with clinical outcomes in 1,851 patients who had surgery with or without adjuvant 5-FU-based chemotherapy. Gene expression was quantified
from microdissected, FFPE primary colon cancer tissue. Of 761 candidate genes, multivariate analysis (including disease severity, stage, and nodal involvement) reduced the gene set to a 7-gene prognostic signature and a separate 6-gene predictive signature. Five reference genes also are included in the assay.

External validation of the algorithm was reported in 2011 in an independent study using FFPE primary tumor samples from patients with stage II colon cancer who had participated in the Quick and Simple and Reliable (QUASAR) study of adjuvant chemotherapy versus surgery alone. The relation between the 7-gene recurrence score and risk of recurrence was statistically significant, with 3-year risk of recurrence for predefined low-, intermediate-, and high-risk groups of 12%, 18%, and 22%, respectively. No relation between a 6-gene treatment score and benefit from chemotherapy was identified.

Venook et al (2013) conducted a validation study using tumor tissue from 690 patients with stage 2 colon cancer who had participated in the Cancer and Leukemia Group B (CALGB) 9581 trial. CALGB 9581 randomized 1713 patients with stage 2 colon cancer to treatment with edrecolomab, an experimental monoclonal antibody, or observation; DFS and overall survival did not differ between treatment groups. Venook et al selected samples stratified by treatment group from those who had tumor tissue available (40% of the original patient sample). They used recurrence score cut points of 29 and 39 to determine low-, intermediate-, and high-risk groups; these values differ from the cut points of 30 and 41 validated in the QUASAR study previously described. Estimated 5-year recurrence risk was 12% (95% CI, 10% to 15%), 15% (95% CI, 12% to 17%), and 18% (95% CI, 14% to 22%) in the low-, intermediate-, and high-risk groups, respectively. In multivariate analysis, every 25-unit change in recurrence score was associated with recurrence independent of tumor stage, tumor grade, MMR status, presence or absence of lymphovascular invasion, and number of nodes assessed.

Yothers et al (2013) conducted a validation study using tumor tissue from 264 patients with stage 2 colon cancer who had participated in the National Surgical Adjuvant Breast and Bowel Project (NSABP) C-07 trial. NSABP C-07 randomized 2409 patients with stage 2 (28%) or stage 3 (72%) colon cancer to adjuvant chemotherapy with 5-FU plus leucovorin (FULV) or oxaliplatin plus FULV (FLOX). Yothers et al randomly selected 50% of patients who had tissue available (total of 892 tissue samples), 264 of whom (30%) had stage 2 cancer. For these patients, estimated 5-year recurrence risks adjusted for treatment (FULV vs FLOX) were 9% (95% CI, 6% to 13%) in the Oncotype-defined low-risk group, 13% (95% CI, 8% to 17%) in the intermediate-risk group, and 18% (95% CI, 12% to 25%) in the high-risk group. Five-year recurrence risk was reduced in high-risk patients who received oxaliplatin compared with those who did not (Kaplan-Meier estimated 5-year recurrence risk, 9% [95% CI, 3% to 25%] FLOX vs 23% [95% CI, 12% to 42%] FULV), but this difference was not observed in low- or intermediate-risk patients. However, confidence intervals for these estimates were wide due to small numbers of patients and events in each risk group. For all stage 3 patients in any risk class, adjusted 5-year recurrence risk estimates exceeded 15%.

Reimers et al (2014) conducted a retrospective study using prospectively collected tumor specimens from the Dutch total mesenteric excision (TME) trial in patients with resectable colon cancer. The Dutch TME trial compared TME with and without radiotherapy in 1861 patients.
Reimers used available tumor tissue from 569 stage 2 and 3 patients randomized to surgery alone; only 297 (52%) specimens were included in their analysis. Among patients with stage 2 rectal cancer (n=130), Oncotype® DX classified 63 (49%) patients as low risk, 37 (28%) patients as intermediate risk, and 30 (23%) patients as high risk. At median follow-up of 12 years (range, 1-14 years), 5-year Kaplan-Meier recurrence risk estimates in the low-, intermediate-, and high-risk groups were 12% (95% CI, 6% to 24%), 29% (95% CI, 17% to 47%), and 53% (95% CI, 35% to 73%), respectively. DX risk classification and estimated recurrence risks for patients with stage 3 rectal cancer were not reported.

Section Summary: Clinical Validity

Several validation studies of GEP for colon cancer have reported that testing provides prognostic information on the risk of recurrence. For patients with a low-risk Oncotype DX score, 5-year recurrence risk might be as high as 24%. This is likely too high to reasonably withhold adjuvant treatment in patients with stage II rectal cancer who would otherwise receive it. Also, the increase in recurrence risk for a high-risk score is modest, and it is uncertain whether the degree of increase is sufficient to intensify management. Some studies have reported that GEP testing offers prognostic information in a multivariate analysis. However, no studies have compared GEP testing to other methods of risk stratification for this population. The evidence is insufficient to determine whether GEP testing provides incremental prognostic information over the standard prognostic work-up.

Clinical Utility

Brenner et al (2016) published a retrospective study of the association between Oncotype DX recurrence score and management decisions. There were 269 patients from 1 health plan included who had stage II colon cancer, MMR proficient status, and Oncotype DX recurrence scores. The primary outcome measures were changes in management that occurred following Oncotype DX testing. Patients were classified as having either an increase in the intensity of surveillance/treatment, a decrease in the intensity of surveillance/treatment, or no change. A change in management following testing was found for 102 (38%) of 269 patients. Of the 102 patients with management changes, there were 76 patients in whom the intensity of management was decreased and 26 in whom it was increased. More patients who had a low recurrence score had a decrease in intensity of management, and more patients with a high recurrence score had an increase in intensity. This type of study does not determine whether patient outcomes are improved as a consequence of the changes in management.

Cartwright et al (2014) and Srivastava et al (2014) have also published studies showing the effect of Oncotype DX results on treatment recommendations made according to traditional risk classifiers in patients with stage II colon cancer. However, these studies did not assess survival or recurrence outcomes. Currently, there is no published information on how the use of GEP results impacts patient outcomes. Absent information showing a direct effect on outcomes or establishing a strong chain of evidence that testing has a positive net effect on outcomes, the clinical utility of testing remains unclear.
A Technical Brief, published by the Agency for Healthcare Research and Quality in 2012, reviewed the clinical evidence for GEP in predicting outcomes, including benefit from adjuvant chemotherapy, in patients with stage II colon cancer. The four commercially available assays reviewed herein were included in the brief. No prospective studies were identified that assessed change in net health outcome with use of a GEP assay, and no studies were identified that used a net reclassification analysis and subsequently evaluated the impact of the reclassification on net health outcome. Additionally, evidence was limited on the reproducibility of test findings, indications for GEP testing in stage II patients, and whether results of GEP assays can stratify patients into groups with clinically meaningful differences in recurrence risk.

In the absence of direct evidence, an indirect chain of evidence could demonstrate clinical utility if all links in the chain are intact. An indirect chain of evidence for clinical utility of GEP testing involves the following series of questions:

- Does GEP testing provide prognostic information?
  Yes. Patients with a low recurrence score have a decreased risk of recurrence and patients with a high risk score have a higher risk of recurrence. However, the degree of difference in risk conferred by the test is not large.
- Does GEP testing provide incremental prognostic information compared to the standard clinical workup for prognosis?
  Uncertain. No well-done studies have compared the prognostic information from GEP testing to standard prognostic information provided by clinical and pathologic workup.
- Does the incremental prognostic information lead to classifying patients into different groups for which management differs?
  No. There are no well-defined treatment protocols that differ according to risk of recurrence. The intensity of surveillance and management may be impacted by results of GEP testing, but the evidence to demonstrate this is weak and not definitive.
- Do the changes in management resulting from GEP testing lead to improvements in health outcomes?
  No. There is no evidence that different treatment protocols for patients with different risks of recurrence improve outcomes.

Section Summary: Clinical Utility

Some studies have reported management changes following GEP testing. However, these studies do not report clinical outcomes and cannot determine whether GEP testing improves health outcomes. An indirect chain of evidence does not demonstrate clinical utility because multiple links in the chain are not intact.

Ongoing and Unpublished Clinical Trials
Some currently unpublished trials that might influence this review are listed in Table 1.

Table 1. Summary of Key Trials
Summary of Evidence

For individuals who have stage II or III colon cancer who receive gene expression profiling (GEP) testing, the evidence includes development and validation studies and 1 decision-impact study. Relevant outcomes are disease-specific survival, test accuracy and validity, and change in disease status. The available evidence has shown that GEP tests for colon cancer can improve risk prediction, particularly the risk of recurrence in patients with stage II or III colon cancer. However, evidence to date is insufficient to permit conclusions on how GEP classification compares with other approaches for identifying recurrence risk in stage II or III patients. The indirect chain of evidence that demonstrates GEP testing would improve health outcomes is weak. Studies showing management changes as a consequence of testing do not demonstrate whether such changes improve outcomes. The evidence is insufficient to determine the effects of the technology on health outcomes.

Supplemental Information

Practice Guidelines and Position Statements

Current clinical practice guidelines from the National Comprehensive Cancer Network (v.2.2016) on colon cancer state that “there is insufficient data to recommend the use of multigene assays to determine adjuvant therapy” in patients with stage II or III colon cancer.

U.S. Preventive Services Task Force Recommendations

Not applicable.

Medicare National Coverage

There is no national coverage determination (NCD). In the absence of an NCD, coverage decisions are left to the discretion of local Medicare carriers.

Important Reminder

The purpose of this Medical Policy is to provide a guide to coverage. This Medical Policy is not intended to dictate to providers how to practice medicine. Nothing in this Medical Policy is intended to discourage or prohibit providing other medical advice or treatment deemed appropriate by the treating physician.
Benefit determinations are subject to applicable member contract language. To the extent there are any conflicts between these guidelines and the contract language, the contract language will control.

This Medical Policy has been developed through consideration of the medical necessity criteria under Hawaii's Patients' Bill of Rights and Responsibilities Act (Hawaii Revised Statutes §432E-1.4), generally accepted standards of medical practice and review of medical literature and government approval status. HMSA has determined that services not covered under this Medical Policy will not be medically necessary under Hawaii law in most cases. If a treating physician disagrees with HMSA’s determination as to medical necessity in a given case, the physician may request that HMSA reconsider the application of the medical necessity criteria to the case at issue in light of any supporting documentation.

VI. References