

## ***Multimarker and Proteomics-based Serum Testing Related to Ovarian Cancer***

**Effective:** March 1, 2019

**Next Review:** December 2019

**Last Review:** January 2019

### **IMPORTANT REMINDER**

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

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

### **DESCRIPTION**

Multimarker serum tests have been proposed as a method for identifying patients likely to have malignant or benign adnexal masses, prior to surgery.

### **MEDICAL POLICY CRITERIA**

The Overa™, OVA1®, and ROMA™ tests are considered **investigational** for all indications, including but not limited to:

- A. Preoperative evaluation of adnexal masses to triage for malignancy
- B. Screening for ovarian cancer
- C. Selecting patients for surgery for an adnexal mass
- D. Evaluation of patients with clinical or radiologic evidence of malignancy
- E. Evaluation of patients with nonspecific signs or symptoms suggesting possible malignancy
- F. Postoperative testing and monitoring to assess surgical outcome and/or to detect recurrent malignant disease following treatment

*NOTE: A summary of the supporting rationale for the policy criteria is at the end of the policy.*

## CROSS REFERENCES

1. [Analysis of Proteomic and Metabolomic Patterns for Early Detection or Assessing Risk of Cancer](#), Laboratory, Policy No. 41

## BACKGROUND

Adult women presenting with an adnexal mass have an estimated 68% likelihood of having a benign lesion.<sup>[1]</sup> About 6% have borderline tumors, 22% have invasive lesions, and 3% have metastatic disease. The mortality rate, for patients with malignant disease depends on three variables: 1) characteristics of the patient; 2) the biology of the tumor (grade, stage, and type); and 3) the quality of treatment (nature of staging, surgery and chemotherapy used).<sup>[2]</sup> In particular, comprehensive staging and completeness of tumor resection appear to have a positive impact on patient outcome.

A number of studies have evaluated the role of a variety of practice-related factors that may improve health outcomes in patients with ovarian cancer, including specialty treatment by gynecological oncologists.<sup>[3-6]</sup> These studies have suggested that this specialty treatment may result in improved outcomes, particularly in patients with advanced stage disease.

Surgery is the only way to diagnose ovarian cancer; this is because biopsy of an ovary with suspected ovarian cancer is usually not performed due to the risk of spreading cancer cells. Serum-based tests have been proposed to triage patients with malignant versus benign adnexal masses. A suggested use of the tests is to identify women who have a higher likelihood of malignant disease and may benefit from referral to a gynecologic-oncology specialist. These tests are combinations of several separate lab tests known as multi-analyte assays with algorithmic analyses (MAAA) and are performed on a blood sample by a reference laboratory using a proprietary algorithm.

The OVA1® test algorithm uses five serum biomarkers, CA 125, prealbumin, apolipoprotein A-1, beta 2 microglobulin, and transferrin. A second-generation test called Overa™ replaces prealbumin and beta 2 microglobulin with human epididymis secretory protein 4 and follicle stimulating hormone.

The Risk of Ovarian Malignancy Algorithm (ROMA™) test combines two biomarkers, human epididymis secretory protein 4 (HE4) and CA 125, along with menopausal status.

## REGULATORY STATUS

**NOTE:** On December 10, 2011, the U.S. Food and Drug Administration (FDA) published an amendment to the regulation for classifying ovarian adnexal mass assessment score test systems to restrict these devices so that a prescribed warning statement that addresses off-label risks be highlighted by a black box warning.<sup>[7]</sup> The warning is intended to mitigate the risk to health associated with off-label use as a screening test, stand-alone diagnostic test, or as a test to determine whether or not to proceed with surgery.

### **The Overa™ test**

In March 2016, the Overa™ test (Vermillion, Inc. Fremont, CA) was cleared for marketing by the FDA through the 510(k) process, where it was submitted as the OVA1 Next Generation

test. This test was predicated on the OVA1® test, to which the Overa™ test is was considered substantially equivalent.

The intended use carried a boxed warning: “PRECAUTION: The OVA1 Next Generation test should not be used without an independent clinical and imaging evaluation and is not intended to be a screening test or to determine whether a patient should proceed to surgery. Incorrect use of the OVA1 Next Generation test carries the risk of unnecessary testing, surgery, and/or delayed diagnosis.”

### **The OVA1® test**

On July 16, 2009, the OVA1® test (Vermillion, Inc. Fremont, CA) was cleared for market by the FDA as a 510(k) submission. No predicate was identified, and the review decision was based on the de novo 510(k) review process which allows novel products with moderate or low-risk profiles and without predicates which would ordinarily require premarket approval as a class III device to be down-classified in an expedited manner and brought to market with a special control as a class II device.

The intended use carried a boxed warning: “PRECAUTION: The OVA1® test should not be used without an independent clinical/radiological evaluation and is not intended to be a screening test or to determine whether a patient should proceed to surgery. Incorrect use of the OVA1® test carries the risk of unnecessary testing, surgery, and/or delayed diagnosis.”

### **The ROMA™ test**

On September 1, 2011, the Risk of Ovarian Malignancy Algorithm (ROMA™ test, Fujirebio Diagnostics, Inc.) was cleared by the FDA as a 510(k) submission. Because the OVA1 test had been found to be a class II medical device by virtue of the July 2009 clearance, ROMA was found to be substantially equivalent to that predicate device.

The intended use carried a boxed warning: “PRECAUTION: The ROMA (HE4 EIA+ARCHITECT CA 125 II) should not be used without an independent clinical/radiological evaluation and is not intended to be a screening test or to determine whether a patient should proceed to surgery. Incorrect use of the ROMA (HE4 EIA+ARCHITECT CA 125 II) carries the risk of unnecessary testing, surgery, and/or delayed diagnosis.”

## **EVIDENCE SUMMARY**

Assessment of a diagnostic technology typically focuses on three parameters: 1) technical performance; 2) diagnostic performance (sensitivity, specificity, and positive [PPV] and negative predictive value [NPV]) in appropriate populations of patients; and 3) demonstration that the diagnostic information can be used to improve patient outcomes (clinical utility).

Technical performance of a device is typically assessed with two types of studies, those that compare test measurements with a gold standard and those that compare results taken with the same device on different occasions (test-retest). Demonstration of technical performance should include an assessment of its reproducibility and precision.

Diagnostic performance is evaluated by the ability of a test to accurately diagnose a clinical condition in comparison with the gold standard. The sensitivity of a test is the ability to detect a disease when the condition is present (true-positive), while specificity indicates the ability to detect patients who are suspected of disease but who do not have the condition (true-

negative). Evaluation of diagnostic performance, therefore, requires independent assessment by the two methods in a population of patients who are suspected of disease but who do not all have the disease.

Evidence related to improvement of clinical outcomes with use of this testing assesses the data linking use of a test to changes in health outcomes (clinical utility). While in some cases, tests can be evaluated adequately using technical and diagnostic performance, when a test identifies a new or different group of patients with a disease; randomized trials are needed to demonstrate impact of the test on the net health outcome.

## **TECHNICAL PERFORMANCE**

Evidence on the technical performance of these tests was evaluated by the U.S. Food and Drug Administration (FDA) and is available through the FDA website. The final analysis indicated acceptable technical performance of Overa™, OVA1®, and ROMA™ for use in clinical care.<sup>[8-10]</sup>

Analytical performance for the OVA1® test demonstrated good test precision (coefficient of variation (CV) ranging from 1% to 7.4%, depending on the sample levels studied) and good reproducibility (CV from 2.8% to 8.9%). The test appears linear, reagent and samples stable, and there was no observed interference evaluating common endogenous substances (hemoglobin, bilirubin, etc.)

The second-generation Overa™ test, had an improved analytical performance, with an overall precision CV of 1.54% across all days and sample pools, in contrast to the OVA1®'s overall precision CV of 4.09%. The reproducibility of Overa™ was also slightly improved over the OVA1®, and it demonstrated a similar lack of interference from common endogenous substances.

Analytical performance for the ROMA also exhibited good precision with a total CV ranging from 0.49% to 7.72%, depending on both sample values and menopausal status. The reproducibility of the test was acceptable, with a CV that ranged from 0.98 to 25.9%, with highest values observed in patients with low scores, as expected. The reagents are variably stable, and users are instructed to follow package inserts for stability on each analyte used. The test was unaffected by interference with hemoglobin, bilirubin, lipids, or human anti-mouse antibodies (HAMA). However, high levels of rheumatoid factor (more than 500 IU/mL) did appear to cause elevations in test values, and testing in patients with elevated rheumatoid factor is not recommended.

## **DIAGNOSTIC PERFORMANCE**

### **Systematic Reviews**

A BlueCross BlueShield Association Technology Evaluation Center (TEC) Assessment was completed in 2012 on "Multi-analyte testing for the evaluation of adnexal masses."<sup>[11]</sup> The Assessment included evaluation of both the OVA1® and ROMA tests and their impact on health outcomes. The single existing study assessing OVA1® was selected. Studies were selected showing the diagnostic characteristics of ROMA using prespecified cutoff values that assessed diagnostic performance for all types of malignancy, and that did not include healthy subjects as non-malignant control subjects. The TEC Assessment concluded that OVA1® appears to improve sensitivity for detection of malignancy, however specificity declines so much that most patients test positive, and that ROMA does not appear to improve the

sensitivity of testing to a great extent.

Wang (2014) published a meta-analysis of studies evaluating the diagnostic accuracy of the ROMA algorithm and comparing it to the performance of single markers HE4 and CA 125.<sup>[12]</sup> To be included in the meta-analysis, studies had to investigate both HE4 and CA 125 or calculate ROMA, include women with ovarian cancer and benign gynecologic disease, use pathology diagnosis as the reference standard, and collect blood samples before treatment was initiated. A total of 32 studies met these inclusion criteria; 6 of these were conducted in the United States. Findings of the overall pooled analysis of diagnostic accuracy are presented in Table 1. Findings were similar when diagnostic performance in premenopausal women and postmenopausal women were evaluated separately. ROMA had similar or higher sensitivity than HE4 and CA 125, and HE4 had the highest specificity.

**Table 1. Diagnostic Performance of ROMA compared with HE4 and CA 125 from Wang (2014): Meta-analysis findings**

	No. Studies	Sensitivity % (95% CI)	Specificity %, 95% CI
<b>ROMA</b>	14	85.3 (81.2-88.6)	82.4 (77.4-86.5)
<b>HE4</b>	28	76.3 (72.0-80.1)	93.6 (90.0-95.9)
<b>CA 125</b>	28	79.2 (74.0-83.6)	82.1 (76.6- 86.5)

Dayyani (2016) conducted a meta-analysis comparing ROMA with HE4 and CA 125 in patients with suspected ovarian cancer.<sup>[13]</sup> Six studies met the inclusion criteria, four of which were included in the Wang (2014) meta-analysis. Two studies were published in 2014 or later. ROMA had higher area under the curve (AUC) values than either HE4 or CA 125 alone (0.921, 95% confidence interval [CI] 0.855 to 0.960, vs 0.899, 95% CI 0.835 to 0.943, and 0.883, 95% CI 0.771 to 0.950, for HE4 and CA 125, respectively). Findings of the pooled analysis of diagnostic accuracy are shown in Table 2.

**Table 2. Diagnostic Performance of ROMA compared with HE4 and CA 125 from Dayyani (2016): Meta-analysis findings**

	No. Studies	Sensitivity % (95% CI)	Specificity %, 95% CI
<b>ROMA</b>	6	87.3 (75.2-94.0)	85.5 (71.9-93.2)
<b>HE4</b>	6	68.2 (69.3-90.1)	85.1 (71.6-92.8)
<b>CA 125</b>	6	79.6 (66.3-88.5)	82.5 (66.2- 91.9)

## Clinical Studies

### OVA1®

Diagnostic performance of the OVA1® test was evaluated in a prospective, double-blind clinical study using 27 demographically mixed subject enrollment sites.<sup>[8]</sup> Patients underwent a complete clinical evaluation prior to surgical intervention, and only patients with planned surgical intervention were included in the study. The pre-surgical process for identifying patients for surgery and for establishing a preliminary diagnosis as benign or malignant were not specifically described but were noted to be “based on a variety of clinical assessments.” The study did require at least one imaging test be performed within 12 weeks of surgery. Presumably, use of this somewhat non-standardized diagnostic methodology provides information on how the test works in conjunction with real-world decision making. The study enrolled a total of 743 patients with 146 subjects used in the training set and 516 in the testing

set. Seventy-four patients were excluded because of missing information or samples. All patients had adnexal masses and were scheduled for surgery. The final prevalence of cancer in the population was 27%.

Using pathologic diagnosis as the gold standard, test performance, when combined with presurgical assessment for benign disease, was as follows in the hands of non-gynecological oncologists:

	Clinical assessment alone	Clinical assessment with OVA1
Sensitivity	72%*	92%
Specificity	83%	42%
Positive predictive value	61%	37%
Negative predictive value	89%	93%

\* Confidence intervals not provided.

OVA1® appeared to improve sensitivity for detection of malignancy; however, specificity declined so much that most patients tested positive.

Bristow (2013) reported on a prospective non-randomized study of 494 patients evaluated for multivariate index assay (OVA1®), CA 125-II, and clinical impression.<sup>[14]</sup> Patients were all scheduled to undergo surgery for an adnexal mass and all were recruited from non-gynecological oncology practices. Authors sought to assess the OVA1 test in determining the need for gynecological oncology referral by comparing OVA1® to clinical assessment and CA 125-II in identifying women with ovarian cancer. Sensitivity, specificity, positive predictive value, negative predictive value, and likelihood ratios were calculated to estimate the performance of OVA1®, CA 125, clinical impression, and OVA1 with clinical impression. For ovarian malignancies, authors reported a sensitivity of 95.7% when combined with clinical impression. The negative predictive value was reported at 98.1%. However, both clinical impression and CA 125-II were more accurate in identifying benign disease. As in the previous study, although sensitivity improved with the OVA1® test, specificity declined and was reported as 53.5% with OVA1® alone and 50.7% with OVA1® combined with clinical impression compared to 92.5% with clinical impression and 86.1% and 94.5% for CA 125-II (using two different cut-off values) in predicting disease.

Further analysis of combined data from the Bristow (2013) study<sup>[14]</sup> and another published OVA1® study, the OVA500 trial<sup>[15]</sup> were reported in two additional articles; one analyzed the performance of clinical assessment with versus without OVA1® testing<sup>[16]</sup> and the other “was undertaken to better understand the impact of ovarian imaging on the clinical interpretation of the MIA score.”<sup>[17]</sup> The sensitivity, specificity, PPV, and NPV were reported for a number of comparisons and combinations of OVA1® testing with other clinical/imaging findings. However, the authors noted that both the OVA1® and OVA500 trials were designed to measure accuracy in prediction of malignancy rather than the test’s effect on patient referral for sub-specialty care.

Grenache (2015) evaluated the diagnostic performance of both the OVA1® and ROMA tests in a prospective case series of 146 women with an adnexal mass.<sup>[18]</sup> Although performance characteristics of both tests were determined and compared, the study did not evaluate diagnostic performance in conjunction with clinical assessment, as the both tests were intended to be used. OVA1® was 97% sensitive and 55% specific, and ROMA was 87% sensitive and 83% specific. This means that with clinical assessment (as intended to be used), the OVA1® test would be no worse than 97% sensitive and no better than 55% specific, but

cannot be determined from the study. The same conclusions can be drawn regarding the ROMA test.

## ROMA

Since the publication of the Wang (2014) and Dayyani (2016) meta-analyses, multiple studies have described the use of Roma populations of women in whom decisions to pursue surgery had been made, including Al Musalhi (2016, n=213 cases),<sup>[19]</sup> Cho (2015, n=90 cases),<sup>[20]</sup> Terlikowska (2016, n=224 cases),<sup>[21]</sup> and Minar (2017, n=267).<sup>[22]</sup>

Diagnostic performance of the ROMA test was evaluated for FDA approval in a prospective, blinded clinical trial using 13 demographically mixed subject enrollment sites with company sponsorship.<sup>[9]</sup> Patients all presented with an adnexal mass and were scheduled to undergo surgery. An Initial Cancer Risk Assessment (ICRA) was performed to determine the detection of benign versus malignant lesions before testing. The prevalence of cancer was 15%.

Using pathologic diagnosis as the gold standard, test performance, when combined with presurgical assessment for benign disease, was as follows in the hands of a mixed population of generalist and specialist physicians:

	ICRA alone	ICRA with ROMA testing
Sensitivity (95% CI)	77% (66% to 86%)	91% (81% to 96%)
Specificity (95% CI)	84% (80% to 88%)	67% (61% to 71%)
Positive predictive value (95% CI)	46% (17% to 56%)	33% (26% to 40%)
Negative predictive value (95% CI)	96% (93% to 97%)	98% (95% to 99%)

Both tests, when added to pre-testing clinical assessment, produced a fall in the positive predictive value of diagnosis with a small increase in the negative predictive value. The changes observed in the negative predictive value were of uncertain statistical and clinical significance.

It is important to note that all of the above literature assessed ROMA as a stand-alone test and did not evaluate diagnostic performance in conjunction with clinical assessment, as the test was intended to be used. Therefore, the ability to draw conclusions regarding the test's diagnostic performance is limited.

Leist (2019) compared the performance of ROMA with that of the risk of malignancy index (RMI) in a prospective study of 784 women from nine Swedish hospitals who were scheduled to have a pelvic mass removed.<sup>[23]</sup> HE4 and CA125 were also evaluated. There were no significant differences between the RMI and ROMA using a fixed specificity of 75%. A study by Moore (2014) did evaluate ROMA in conjunction with clinical assessment, using either positive clinical assessment or positive ROMA as a positive test (similar to the recommended usage for OVA1®).<sup>[24]</sup> Using this method of combining tests guarantees a higher sensitivity and lower specificity for the combined test than for either test alone. Used in this way, ROMA would only need to be evaluated in patients with a negative clinical assessment. In this study, 461 women were enrolled, of whom a total of 86 (19%) had a malignancy. Combined assessment improved sensitivity from 77.9% to 89.7%, but worsened specificity from 84.3% to 67.2%.

Nikolova (2017) published a case-control study that compared ROMA, Copenhagen Index (CPH-I), HE4, CA 125, RMI, and Morphology Index (MI) for differentiation between ovarian endometriosis and epithelial ovarian cancer.<sup>[25]</sup> The study included 164 patients: 37 with ovarian endometriosis, 57 with other benign pelvic masses, 11 with ovarian cancer, and 59

controls. The sensitivity, specificity and accuracy to distinguish endometriosis from ovarian cancer for HE4 were: 81.82%, 100%, and 95.83%; for CA 125: 81.82%, 48.65%, and 56.25%; for ROMA: 90.91%, 83.78%, and 85.42%; for CPH-I: 81.82%, 97.30%, and 93.75%; for RMI: 90.91%, 35.14%, and 47.92%; and for MI: 100%, 75.68%, and 81.25%, respectively. The authors concluded that HE4 and CPH-1 had the best ability to discriminate these two conditions, and that MI had the highest sensitivity for ovarian cancer.

### Overa™

In a study submitted to the FDA, the clinical validity of the Overa™ test was evaluated in a nonconcurrent prospective study of 493 preoperatively collected serum specimens from premenopausal and postmenopausal women presenting with an adnexal mass requiring surgical intervention.<sup>[10]</sup> Overa™ test scores were determined based on the analysis of archived serum specimens from a previous study,<sup>[14]</sup> and patients were stratified into a low- or high-risk groups for finding malignancy on surgery. The analysis examined whether patient referral to a gynecologic oncologist was supported when dual assessment was determined to be positive (either Overa™ or clinical assessment was positive, or both were positive). A dual assessment was considered negative when both Overa™ and clinical assessment were negative.

Using pathologic diagnosis as the criterion standard, Overa™ test performance, when combined with a clinical assessment by nongynecologic oncologists, was as follows:

	Clinical assessment alone	Dual assessment with Overa™
Sensitivity (95% CI)	74% (64% to 82%)	94% (87% to 97%)
Specificity (95% CI)	93% (90% to 95%)	65% (60% to 70%)
Positive predictive value (95% CI)	70% (62% to 77%)	38% (35% to 41%)
Negative predictive value (95% CI)	94% (92% to 96%)	98% (95% to 99%)

The method used for combining clinical assessment and Overa™ test result was to consider the test positive if either clinical assessment or Overa™ test was positive. Thus, in practice, the Overa™ testing would not be necessary if clinical assessment alone indicated cancer. Using Overa™ testing in this manner guarantees that Overa™ testing will be more sensitive and less specific than clinical assessment alone, even if it has no better than chance capability of detecting ovarian cancer. Sensitivity improved from 74% to 94%, and specificity decreased from 93% to 65%.

### **Section Summary**

There are a limited number of studies comparing diagnostic accuracy of the ROMA, OVA1®, and Overa™ proteomic tests. Studies have found that ROMA has similar or lower accuracy to other risk prediction measures that use components of the standard workup, such as the RMI and the LR2 measures. Use of these tests in addition to other clinical assessment tools appears to increase sensitivity but decreases specificity. Further prospective studies are needed for both assays to understand their proper role in patient care.

### **CLINICAL UTILITY**

The ideal study design to evaluate clinical utility of proteomics-based testing is a randomized controlled trial comparing patient management decisions (e.g., referral patterns) and/or health outcomes (e.g., mortality) in patients managed with proteomic tests with those managed according to best current clinical practices.



## Systematic Review

The 2012 TEC Assessment<sup>[11]</sup> concluded that studies of OVA1® and ROMA showed improvements in sensitivity and worsening of specificity with the use of the tests in conjunction with clinical assessment., and that there are problems in concluding that this results in improved health outcomes. The clinical assessment performed in the studies was not well characterized. The evidence regarding the effect of OVA1® and ROMA on health outcomes is indirect, and based on studies of diagnostic performance of the tests in patients undergoing surgery for adnexal masses. The authors noted that there were no prospective studies on the use of these tests in patients who presented with an adnexal mass or studies that reported the impact of testing on referral patterns or the impact on health outcomes.

## Randomized Controlled Trials

No RCTs on the effects of these tests on patient management decisions and healthcare outcomes compared with current assessment were identified.

## Nonrandomized Studies

No outcome studies have been performed using the OVA1®, Overa™, or ROMA test. Kaijser (2014) published a retrospective cohort study included 101 newly diagnosed cases of biopsy-proven invasive ovarian cancer, which provided information relevant to outcomes.<sup>[26]</sup> Blood samples obtained before treatment were analyzed; HE4 and CA 125 levels were measured and the ROMA algorithm was calculated. Median overall survival in the study cohort was 3.7 years. In a multivariate analysis controlling for confounding variables, neither HE4 levels nor ROMA were independently associated with progression-free survival (PFS) or disease-specific survival (DSS). For example, for ROMA and the outcome of PFS, the adjusted hazard ratio (HR) for each 10% increase in risk was 0.98(95% CI) 0.88 to 1.11). Patients were not prospectively managed according to their HE4 levels or ROMA score and thus the actual impact of these tests on PFS and DSS cannot be determined from this study. It is not clear what impact either test would have on long-term health outcomes or referral patterns to specialty physicians. The use of proteomic testing to triage patients for malignancy may be only one of many factors in decision making about where treatment should be delivered.

## Section summary

Although current studies show improvements in sensitivity and worsening of specificity with the use of the OVA1®, Overa™, and ROMA tests in conjunction with clinical assessment, there are problems in concluding that this results in improved health outcomes. The clinical assessment performed in the studies is not well characterized. In addition, there is indirect evidence from studies of diagnostic accuracy which suggest that the ROMA test would not improve the accuracy of triage compared to existing measures, and is unlikely to improve the accuracy of referral to a specialist and is therefore, unlikely to improve outcomes.

## PRACTICE GUIDELINE SUMMARY

### THE SOCIETY FOR GYNECOLOGIC ONCOLOGY (SGO)

In May 2013, the Society for Gynecologic Oncology (SGO) issued the following consensus based statement on multiplex serum testing for women with pelvic masses:<sup>[27]</sup>

“Blood levels of five proteins in women with a known ovarian mass have been reported

to change when ovarian cancer is present. Tests measuring these proteins may be useful in identifying women who should be referred to a gynecologic oncologist. Recent data have suggested that such tests, along with physician clinical assessment, may improve detection rates of malignancies among women with pelvic masses planning surgery. Results from such tests should not be interpreted independently, nor be used in place of a physician's clinical assessment. Physicians are strongly encouraged to reference the American Congress of Obstetricians and Gynecologists' 2011 Committee Opinion "The Role of the Obstetrician-Gynecologist in the Early Detection of Epithelial Ovarian Cancer" to determine an appropriate care plan for their patients."

The SGO Position Statements regarding the use of the OVA1 or ROMA test is not based upon scientific evidence and does not recommend the use of the OVA1 test as part of their referral guidelines for women who present with an adnexal mass.

### **THE AMERICAN CONGRESS OF OBSTETRICIANS AND GYNECOLOGISTS (ACOG)<sup>[28]</sup>**

ACOG addresses the use of the serum marker testing in their 2016 practice bulletin on the Evaluation and Management of Adnexal Masses. This document states:

Serum biomarker panels may be used as an alternative to CA 125 level alone in determining the need for referral to or consultation with a gynecologic oncologist when an adnexal mass requires surgery. These biomarker panels are not recommended for use in the initial evaluation of an adnexal mass, but may be helpful in assessing which women would benefit from referral to a gynecologic oncologist. Trials that have evaluated the predictive value of these panels show potential for improved specificity, especially for evaluation of premenopausal women. However, comparative research has not yet defined the best testing approach.

### **THE NATIONAL CANCER INSTITUTE (NCI)**

The NCI included a discussion of proteomics in their publication on the Genetics of Breast and Ovarian Cancer.<sup>[29]</sup> The following statement is included in their discussion on proteomics:

"These (proteomic) studies have generally been small case-control studies that are limited by sample size and the number of early-stage cancer cases included. Further evaluation is needed to determine whether any additional markers identified in this fashion have clinical utility for the early detection of ovarian cancer in the unselected clinical population of interest." (Level of Evidence: 5; Opinions of respected authorities are based on clinical experience, descriptive studies, or reports of expert committees.)

### **NATIONAL COMPREHENSIVE CANCER NETWORK (NCCN)**

The NCCN ovarian cancer guidelines (v2.2018)<sup>[30]</sup> include the following statement:

"It has been suggested that specific biomarkers (serum HE4 and CA-125) along with an algorithm (Risk of Ovarian Malignancy Algorithm [ROMA]) may be useful for determining whether a pelvic mass is malignant or benign. The FDA has approved the use of HE4 and CA-125 for estimating the risk of ovarian cancer in women with a pelvic mass. Currently, the NCCN Panel does not recommend the use of these biomarkers for determining the status of an undiagnosed pelvic mass."

Regarding the OVA1 test, the guidelines state:

“The OVA1 test uses 5 markers (including transthyretin, apolipoprotein A1, transferrin, beta-2 microglobulin, and CA-125) to assess who should undergo surgery by an experienced gynecologic oncologist and who can have surgery in the community. [...] Based on data documenting an increased survival, NCCN Guidelines Panel Members recommend that all patients should undergo surgery by an experienced gynecologic oncologist (category 1).”

## SUMMARY

There is not enough research to show that multimarker serum testing to determine the risk of ovarian cancer can improve health outcomes for women with pelvic masses. Clinical guidelines based on research do not currently recommend the use of these tests for patients with pelvic masses. Therefore, all uses of these tests, including use as a screening tool for ovarian cancer, are considered investigational.

## REFERENCES

1. Van Holsbeke, C, Van Belle, V, Leone, FP, et al. Prospective external validation of the 'ovarian crescent sign' as a single ultrasound parameter to distinguish between benign and malignant adnexal pathology. *Ultrasound in obstetrics & gynecology : the official journal of the International Society of Ultrasound in Obstetrics and Gynecology*. 2010 Jul;36(1):81-7. PMID: 20217895
2. du Bois, A, Rochon, J, Pfisterer, J, Hoskins, WJ. Variations in institutional infrastructure, physician specialization and experience, and outcome in ovarian cancer: a systematic review. *Gynecologic oncology*. 2009 Feb;112(2):422-36. PMID: 18990435
3. Vernooij, F, Heintz, P, Witteveen, E, van der Graaf, Y. The outcomes of ovarian cancer treatment are better when provided by gynecologic oncologists and in specialized hospitals: a systematic review. *Gynecologic oncology*. 2007 Jun;105(3):801-12. PMID: 17433422
4. Giede, KC, Kieser, K, Dodge, J, Rosen, B. Who should operate on patients with ovarian cancer? An evidence-based review. *Gynecologic oncology*. 2005 Nov;99(2):447-61. PMID: 16126262
5. Tingulstad, S, Skjeldestad, FE, Hagen, B. The effect of centralization of primary surgery on survival in ovarian cancer patients. *Obstetrics and gynecology*. 2003 Sep;102(3):499-505. PMID: 12962932
6. Grabowski, JP, Harter, P, Buhrmann, C, et al. Re-operation outcome in patients referred to a gynecologic oncology center with presumed ovarian cancer FIGO I-IIIa after sub-standard initial surgery. *Surgical oncology*. 2012 Mar;21(1):31-5. PMID: 20875732
7. Medical Devices: Ovarian adnexal mass assessment score test system; Labeling; Black box restrictions. 21 CFR Part 866, Federal Register 2011;76(251):82128-82123.
8. U.S. Food and Drug Administration. 510(k) Substantial Equivalence Determination Decision Summary: OVA1™ Test (K081754). [cited 01/09/2019]; Available from: [http://www.accessdata.fda.gov/cdrh\\_docs/reviews/K081754.pdf](http://www.accessdata.fda.gov/cdrh_docs/reviews/K081754.pdf)
9. U.S. Food and Drug Administration. 510(k) Substantial Equivalence Determination Decision Summary: ROMA™ Test (K103358). [cited 01/09/2019]; Available from: [http://www.accessdata.fda.gov/cdrh\\_docs/reviews/K103358.pdf](http://www.accessdata.fda.gov/cdrh_docs/reviews/K103358.pdf)

10. U.S. Food and Drug Administration. 510(k) Substantial Equivalence Determination Decision Summary: Overa™ Test (K150588). [cited 01/09/2019]; Available from: [http://www.accessdata.fda.gov/cdrh\\_docs/pdf15/k150588.pdf](http://www.accessdata.fda.gov/cdrh_docs/pdf15/k150588.pdf)
11. TEC Assessment 2012. "Multi-analyte Testing for the Evaluation of Adnexal Masses." BlueCross BlueShield Association Technology Evaluation Center, Vol. 27, Tab TBA.
12. Wang, J, Gao, J, Yao, H, Wu, Z, Wang, M, Qi, J. Diagnostic accuracy of serum HE4, CA125 and ROMA in patients with ovarian cancer: a meta-analysis. *Tumour biology : the journal of the International Society for Oncodevelopmental Biology and Medicine*. 2014 Jun;35(6):6127-38. PMID: 24627132
13. Dayyani, F, Uhlig, S, Colson, B, et al. Diagnostic Performance of Risk of Ovarian Malignancy Algorithm Against CA125 and HE4 in Connection With Ovarian Cancer: A Meta-analysis. *International journal of gynecological cancer : official journal of the International Gynecological Cancer Society*. 2016 Nov;26(9):1586-93. PMID: 27540691
14. Bristow, RE, Smith, A, Zhang, Z, et al. Ovarian malignancy risk stratification of the adnexal mass using a multivariate index assay. *Gynecologic oncology*. 2013 Feb;128(2):252-9. PMID: 23178277
15. Ueland, FR, Desimone, CP, Seamon, LG, et al. Effectiveness of a multivariate index assay in the preoperative assessment of ovarian tumors. *Obstetrics and gynecology*. 2011;117:1289-97. PMID: 21606739
16. Longoria, TC, Ueland, FR, Zhang, Z, et al. Clinical performance of a multivariate index assay for detecting early-stage ovarian cancer. *American journal of obstetrics and gynecology*. 2014 Jan;210(1):78 e1-9. PMID: 24055582
17. Goodrich, ST, Bristow, RE, Santoso, JT, et al. The effect of ovarian imaging on the clinical interpretation of a multivariate index assay. *American journal of obstetrics and gynecology*. 2014 Jul;211(1):65 e1- e11. PMID: 24530816
18. Grenache, DG, Heichman, KA, Werner, TL, Vucetic, Z. Clinical performance of two multi-marker blood tests for predicting malignancy in women with an adnexal mass. *Clinica chimica acta; international journal of clinical chemistry*. 2015 Jan 1;438:358-63. PMID: 25283731
19. Al Musalhi, K, Al Kindi, M, Al Aisary, F, et al. Evaluation of HE4, CA-125, Risk of Ovarian Malignancy Algorithm (ROMA) and Risk of Malignancy Index (RMI) in the Preoperative Assessment of Patients with Adnexal Mass. *Oman medical journal*. 2016 Sep;31(5):336-44. PMID: 27602187
20. Cho, HY, Park, SH, Park, YH, et al. Comparison of HE4, CA125, and Risk of Ovarian Malignancy Algorithm in the Prediction of Ovarian Cancer in Korean Women. *Journal of Korean medical science*. 2015 Dec;30(12):1777-83. PMID: 26713052
21. Terlikowska, KM, Dobrzycka, B, Witkowska, AM, et al. Preoperative HE4, CA125 and ROMA in the differential diagnosis of benign and malignant adnexal masses. *Journal of ovarian research*. 2016 Jul 19;9(1):43. PMID: 27436085
22. Minar, L, Felsing, M, Cermakova, Z, Zlamal, F, Bienertova-Vasku, J. Comparison of the Copenhagen Index versus ROMA for the preoperative assessment of women with ovarian tumors. *International journal of gynaecology and obstetrics: the official organ of the International Federation of Gynaecology and Obstetrics*. 2017 Oct 31. PMID: 29086914
23. Liest, AL, Omran, AS, Mikiver, R, Rosenberg, P, Uppugunduri, S. RMI and ROMA are equally effective in discriminating between benign and malignant gynecological tumors: A prospective population-based study. *Acta obstetrica et gynecologica Scandinavica*. 2019 Jan;98(1):24-33. PMID: 30216407

24. Moore, RG, Hawkins, DM, Miller, MC, et al. Combining clinical assessment and the Risk of Ovarian Malignancy Algorithm for the prediction of ovarian cancer. *Gynecologic oncology*. 2014 Dec;135(3):547-51. PMID: 25449569
25. Nikolova, T, Zivadinovic, R, Evtimovska, N, et al. Diagnostic performance of human epididymis protein 4 compared to a combination of biophysical and biochemical markers to differentiate ovarian endometriosis from epithelial ovarian cancer in premenopausal women. *The journal of obstetrics and gynaecology research*. 2017 Dec;43(12):1870-9. PMID: 29027715
26. Kaijser, J, Van Belle, V, Van Gorp, T, et al. Prognostic value of serum HE4 levels and risk of ovarian malignancy algorithm scores at the time of ovarian cancer diagnosis. *International journal of gynecological cancer : official journal of the International Gynecological Cancer Society*. 2014 Sep;24(7):1173-80. PMID: 24987915
27. Society of Gynecologic Oncologists. Multiplex Serum Testing for Women with Pelvic Mass (May 2013). [cited 01/09/2019]; Available from: <https://www.sgo.org/newsroom/position-statements-2/multiplex-serum-testing-for-women-with-pelvic-mass/>
28. The American College of Obstetricians and Gynecologists Practice Bulletin: Evaluation and Management of Adnexal Masses. [cited 1/4/2018]; Available from: <http://www.acog.org/Resources-And-Publications/Practice-Bulletins/Committee-on-Practice-Bulletins-Gynecology/Evaluation-and-Management-of-Adnexal-Masses>
29. National Cancer Institute. PDQ Genetics of Breast and Ovarian Cancer. 2014. [cited 01/09/2019]; Available from: <http://www.cancer.gov/cancertopics/pdq/genetics/breast-and-ovarian/HealthProfessional/page4>
30. National Comprehensive Cancer Network (NCCN). Clinical Practice Guidelines in Oncology™. Ovarian Cancer Including Fallopian Tube Cancer and Primary Peritoneal Cancer. v.2.2018. [cited 01/09/2019]; Available from: [http://www.nccn.org/professionals/physician\\_gls/pdf/ovarian.pdf](http://www.nccn.org/professionals/physician_gls/pdf/ovarian.pdf)
31. BlueCross BlueShield Association Medical Policy Reference Manual "Proteomics-based testing for the Evaluation of Ovarian (Adnexal) Masses." Policy No. 2.04.62

## CODES

Codes	Number	Description
CPT	81500	Oncology (ovarian), biochemical assays of two proteins (CA-125 and HE4), utilizing serum, with menopausal status, algorithm reported as a risk score (This code is for reporting the ROMA™ test)
	81503	Oncology (ovarian), biochemical assays of five proteins (CA-125, apolipoprotein A1, beta-2 microglobulin, transferrin and pre-albumin), utilizing serum, algorithm reported as a risk score (This code is for reporting the OVA1™ test)
	0003U	Oncology (ovarian) biochemical assays of five proteins (apolipoprotein A-1, CA 125 II, follicle stimulating hormone, human epididymis protein 4, transferrin), utilizing serum, algorithm reported as a likelihood score
HCPCS	None	

**Date of Origin:** November 2012