

---

## Medical Policy



Blue Cross  
Blue Shield  
Blue Care Network  
of Michigan

Nonprofit corporations and independent licensees  
of the Blue Cross and Blue Shield Association

**Joint Medical Policies are a source for BCBSM and BCN medical policy information only. These documents are not to be used to determine benefits or reimbursement. When Centers for Medicare and Medicaid (CMS) coverage rules are not fully developed, this medical policy may be used by BCBSM or BCN Medicare Advantage plans 42 CFR § 422.101 (b)(6). Please reference the appropriate certificate or contract for benefit information. This policy may be updated and is therefore subject to change.**

---

**\*Current Policy Effective Date: 1/1/26**  
(See policy history boxes for previous effective dates)

### **Title: Multimarker Serum-Based Testing Related to Ovarian Cancer (e.g., OVA1®, Overa™, OvaWatch and ROMA™ testing)**

---

#### **Description**

A variety of serum biomarkers have been studied for their association with ovarian cancer. Of particular interest have been tests that integrate results from multiple analytes into a risk score to predict the presence of disease. Three tests based on this principle, OVA1, Overa (the second-generation OVA1 test), and the Risk of Ovarian Malignancy Algorithm (ROMA) have been cleared by the U.S. Food and Drug Administration. The intended use of OVA1 and Overa is as an aid to further assess whether malignancy is present in a patient with an ovarian adnexal mass who has not yet been referred to an oncologist, even when the physician's independent clinical and radiologic evaluation does not indicate malignancy. The intended use of ROMA is as an aid, in conjunction with clinical assessment, to assess whether a premenopausal or a postmenopausal woman who presents with an ovarian adnexal mass and has not yet been referred to an oncologist is at a high or low likelihood of finding malignancy on surgery.

---

#### **Medical Policy Statement**

Multimarker serum-based testing (e.g., OVA1®, Overa™ and ROMA™ tests) to identify individuals with adnexal masses who may benefit from referral to a gynecologic-oncology specialist is established. These tests may be considered a useful (but not mandatory) diagnostic option in guiding referral to a gynecologic oncologist for individuals meeting defined criteria.

OvaWatch does not provide any additional clinically relevant information in the diagnosis and/or treatment of ovarian cancer. The test is therefore experimental/investigational.

---

## **Inclusionary and Exclusionary Guidelines**

### **Inclusions:**

The proteomics-based OVA1® test, Overa™ and the ROMA™ (Risk of Ovarian Malignancy Algorithm [HE4 EIA + ARCHITECT CA 125 II]) tests are considered established. They are established when used as an aid to further assess the likelihood an adnexal mass is malignant when the physician's (other than gynecologic oncologist) independent clinical and radiological preoperative evaluations do not indicate malignancy and a referral to a gynecologic oncologist is being considered in the event of a positive test result when **ALL** of the following criteria have been met:

- The individual should be older than age 18 years; **AND**
- Ovarian adnexal mass is present; **AND**
- Surgery is planned for treatment of the mass.

### **Exclusions:**

Testing for Ova1®, Overa™ and ROMA™ that does not meet the above criteria, including, but not limited to:

- Screening for ovarian cancer;
- Selecting patients for surgery for an adnexal mass;
- Evaluation of patients with clinical or radiologic evidence of malignancy;
- OvaWatch testing.

---

**CPT/HCPCS Level II Codes** (Note: The inclusion of a code in this list is not a guarantee of coverage. Please refer to the medical policy statement to determine the status of a given procedure)

### **Established codes:**

81500            81503            0003U

### **Other codes (investigational, not medically necessary, etc.):**

0375U            0577

---

## **Background**

### **Epithelial Ovarian Cancer**

The term *epithelial ovarian cancer* collectively includes high-grade serous epithelial ovarian, fallopian tubal, and peritoneal carcinomas due to their shared pathogenesis, clinical presentation, and treatment. We use epithelial ovarian cancer to refer to this group of malignancies in the discussion that follows. There is currently no serum biomarker that can distinguish between these types of carcinomas. An estimated 19,710 women in the United States were expected to be diagnosed in 2024 with ovarian cancer, and approximately 12,740 were estimated to die of the disease.<sup>1</sup> The mortality rate depends on 3 variables: (1) patient characteristics; (2) tumor biology (grade, stage, type); and (3) treatment quality (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. Racial, ethnic, and socioeconomic disparities in management and outcomes are prominent in patients with ovarian cancer. Compared to non-Hispanic White and Asian patients, Hispanic and non-Hispanic Black patients are more likely to be diagnosed with advanced disease, and are less likely to undergo optimal primary surgery and adjuvant chemotherapy.<sup>3,4,5</sup> Patients with ovarian cancer from racial and ethnic minorities are also less likely to be enrolled in clinical trials.<sup>6</sup> These are among the contributing factors to worsened overall survival among these racial and ethnic groups.<sup>7,4,8</sup> Patients with impediments to access healthcare (e.g.,

those living in underserved areas, with low household income, and/or who are underinsured or uninsured), which frequently intersect with racial and ethnic determinants, also experience longer time to diagnosis, suboptimal treatment, and worse outcomes.<sup>9,10,11,5</sup>

Adult women presenting with an adnexal mass have an estimated 68% likelihood of having a benign lesion.<sup>12</sup> About 6% have borderline tumors; 22%, invasive malignant lesions, and 3%, metastatic disease. Clinicians generally agree that women with masses that have a high likelihood of malignancy should undergo surgical staging by gynecologic oncologists. However, women with clearly benign masses do not require referral to a specialist. Criteria and tests that help differentiate benign from malignant pelvic masses are thus desirable.

In 2016, the American College of Obstetricians and Gynecologists updated a practice bulletin that addressed criteria for referring women with adnexal masses to gynecologic oncologists.<sup>13</sup> Separate criteria were developed for premenopausal and postmenopausal women because the specificity and positive predictive value of cancer antigen 125 (CA 125) are higher in postmenopausal women. Prior guidance, which was based on expert opinion, recommended a CA 125 >200 U/mL for referring premenopausal women with an adnexal mass to a gynecologic oncologist. The current guidance advises using very elevated CA125 levels with other clinical factors such as ultrasound findings, ascites, a nodular or fixed pelvic mass, or evidence of abdominal or distant metastasis for referral. The referral criteria for postmenopausal women are similar, except that a lower threshold for an elevated CA 125 test is used (35 U/mL). The practice bulletin states that serum biomarker panels are alternatives to CA 125 levels when deciding about a gynecologic oncologist referral.

Three multimarker serum-based tests specific to ovarian cancer have been cleared by the Food and Drug Administration (FDA) with the intended use of triaging patients with adnexal masses. They are summarized in Table 1. The proposed use of the tests is to identify women with a substantial likelihood of malignant disease who may benefit from referral to a gynecologic oncology specialist. Patients with positive results may be considered candidates for treatment. The tests have been developed and evaluated only in patients with adnexal masses and planned surgical removal. Other potential uses, such as selecting patients to have surgery, screening asymptomatic patients, and monitoring treatment, have not been investigated. Furthermore, the tests are not intended to be used as stand-alone tests, but in conjunction with clinical assessment.

Other multimarker panels and longitudinal screening algorithms are under development but are not yet commercially available.<sup>14,15</sup>

**Table 1. Summary of FDA-Approved Multimarker Serum-Based Tests Specific to Ovarian Cancer**

Variables	OVA1	Overa	ROMA
Cleared	2009	2016	2011
Manufacturer	Quest Diagnostics	Vermillion	Roche Diagnostics
Biomarkers used			
CA 125 II	x	x	x
β <sub>2</sub> -microglobulin	x		
Transferrin	x	x	
Transthyretin	x		
Apolipoprotein AI	x	x	
HE4		x	x
FSH		x	

Score range	0-10	0-10	0-10
Risk categorization			
Premenopausal	<5.0: low ≥5.0: high	<5.0: low ≥5.0: high	≥1.3: high
Postmenopausal	<4.4: low >4.4: high		≥2.77: high

CA 125: cancer antigen 125; FDA: Food and Drug Administration; FSH: follicle stimulating hormone; HEA: human epididymis secretory protein 4.

Another test, OvaWatch, is a non-invasive, serum test that combines the results of seven serum biomarker immunoassays as well as two categorical inputs, age and menopausal status, using a proprietary algorithm. OvaWatch is intended to assist in assessing the risk of ovarian cancer in women with an adnexal mass when a healthcare provider’s initial clinical assessment indicates the mass is either indeterminate or benign, who is not planned for surgery.

## SUMMARY OF EVIDENCE

The evidence for use of multimarker serum-based testing (OVA1 test, Overa or ROMA test) in conjunction with clinical assessment in patients who have adnexal masses undergoing surgery includes studies assessing the technical performance and diagnostic accuracy of the tests. Relevant outcomes are overall survival and test accuracy. The evidence is sufficient to determine that the technology results in an improvement in health outcomes.

Data has been presented that confirms the ROMA algorithm and OVA1 testing in the distinction of ovarian cancer from benign disease. The utility of this testing can be found in patients already going to surgery. The results of this testing along with the clinical picture allow for the appropriate triage of patients to a surgeon who is a generalist or to a gynecologist oncologist. The evidence is sufficient to determine that the technology results in an improvement in health outcomes.

The evidence for use of OvaWatch testing includes two studies. Limitations of this work include the largely retrospective nature of the data set and the unequal, albeit random, assignment of histologic subtypes between the training and validation data sets. It is important to note that prospective data sets are limited in size. The low likelihood of finding malignancies in patients with incidentally discovered and mostly simple cystic adnexal masses and the rare nature of ovarian cancer could have impacted the prevalence and likely the accuracy of performance metrics. The OvaWatch was designed on data that was initially collected to address a higher risk population. The influence of the higher risk patients on the biomarker values in the train/test sets have led to low sensitivity in the prospective populations, collected from women at lower risk. The evidence is insufficient to determine that the technology results in an improvement in the net health outcome.

## Rationale

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. The following is a summary of the key findings to date.

## **MULTIMARKER SERUM TESTING RELATED TO OVARIAN CANCER**

### **Clinical Context and Test Purpose**

The purpose of multimarker serum testing of individuals over age 18 with an ovarian adnexal mass for which surgery is planned and not yet referred to an oncologist is to use the test as an aid to further assess the probability that malignancy is present, even when the physician's independent clinical and radiologic evaluation does not indicate malignancy.

The following **PICOs** were used to select literature to inform this review.

### **Population**

The relevant population of interest is individuals who:

- Are over age 18
- Have ovarian adnexal mass for which surgery is planned
- Have not yet been referred to an oncologist
- A physician's independent clinical and radiologic evaluation does not indicate malignancy.

### **Interventions**

The relevant interventions are 3 U.S. FDA cleared commercial multimarker serum genetic tests (e.g., OVA1, Overa, ROMA). Multimarker serum testing for related to ovarian cancer may be performed at any point when an individual presents with an ovarian adnexal mass for which surgery is planned, to use in conjunction with physician's independent clinical and radiologic evaluation to assess the probability that malignancy is present and aid in the decision of whether a referral to an oncologist is indicated. Most patients are likely to be tested in an outpatient setting.

### **Comparators**

The comparator of interest is standard clinical assessment.

### **Outcomes**

The potential beneficial outcomes of primary interest in the case of a true negative would be the avoidance of unnecessary surgery and its associated consequences (e.g., morbidity, mortality, resource utilization, patient anxiety). The potential harm from a false-positive could be inappropriate assessment and improper management of patients with ovarian malignancies, which could result in the following: inappropriate surgical decisions, high frequency of unnecessary further testing, and unnecessary patient anxiety. The potential harm from a false negative could be a determination that the patient does not have ovarian malignancy, which would lead to a delay in surgery and tumor diagnosis.

Off-label use of the test (e.g., in patients who have not already been identified as needing surgery for pelvic mass, or patients without reference to an independent clinical and radiologic evaluation), might lead to a high frequency of unnecessary testing and surgery due to false-positive results, or to a delay in tumor diagnosis due to false-negative results.

### **Study Selection Criteria**

Below are selection criteria for studies to assess whether a test is clinically valid.

1. Reported on the accuracy of the marketed version of the technology
2. Included a suitable reference standard

3. Patient/sample characteristics were described
- 4 Patient/sample selection criteria were described.

### 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).

### OVA1 Test

Descriptions of the developmental process for the OVA1 test have been published in U.S. Food and Drug Administration (FDA) documents and in a perspective by Fung (2010).<sup>17</sup> Candidate biomarkers were selected based on initial studies using mass spectroscopy but were converted to standard immunoassays to improve analytic performance. Seven final markers were evaluated, none of which individually appeared to be highly specific to malignant ovarian disease. However, the choice of 5 of these (CA 125, prealbumin, apo A1,  $\beta$ 2-microglobulin, transferrin) produced a composite profile that did appear to have discriminatory ability. The test, as cleared by FDA, is performed on a blood sample, which is to be sent to a reference laboratory for testing using the 5 immunoassays previously described. Results of the 5 determinations are entered manually into an Excel spreadsheet used by the OvaCalc software. This software contains an algorithm that combines the 5 discrete values into a single unitless numeric score from 0.0 to 10.0.

Details of the algorithm appear proprietary, but the development is described as an empiric process, based on use of banked samples from academic partners, on a small prospective study of samples from Europe and using a designated subset of samples from the clinical study used to support submission to FDA. It appears at an undisclosed point in the developmental process as a result of interaction with FDA; separate cutpoints were developed for premenopausal and postmenopausal women.

The clinical validity was evaluated in a prospective, double-blind, clinical study using 27 enrollment sites.<sup>18</sup> The study was supported by the commercial sponsor of the test. Patients underwent a complete clinical evaluation before surgical intervention, and only patients with adnexal masses who had a planned surgical intervention were included. The study enrolled 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. The final prevalence of cancer in the population was 27%.

Using pathologic diagnosis as the criterion standard, OVA1 test performance, when combined with presurgical assessment for benign disease, was as follows (see Table 2). The method used for combining clinical assessment and OVA1 result was to consider the test positive if *either* clinical assessment or OVA1 test was positive. Thus, in practice, OVA1 testing would not be necessary if clinical assessment alone indicated cancer. Using OVA1 testing in this manner guarantees that OVA1 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 72% to 92%, and specificity decreased from 83% to 42%.

**Table 2. Clinical Validity of the OVA1 Test<sup>a</sup> Among 269 Patients Evaluated by Nongynecologic Oncologists**

Diagnostic Characteristics	Clinical Assessment Alone, %	Clinical Assessment with OVA1, %
Sensitivity	72	92
Specificity	83	42
Positive predictive value	61	37

<sup>a</sup> Confidence interval not provided

Adapted from U.S. Food and Drug Administration (FDA). 510(k) Substantial Equivalence Determination Decision Summary: OVA1™ Test (K081754) n.d.; [http://www.accessdata.fda.gov/cdrh\\_docs/reviews/K081754.pdf](http://www.accessdata.fda.gov/cdrh_docs/reviews/K081754.pdf). Accessed October 2019.

One additional 2015 study (by Grenache et al) was identified; it evaluated the diagnostic performance of the OVA-1 test.<sup>19</sup> However, it did not evaluate diagnostic performance in conjunction with clinical assessment, as the test was intended to be used. By itself, OVA1 was 97% sensitive and 55% specific. This means that with clinical assessment (as intended to be used), the test would be no worse than 97% sensitive and no better than 55% specific, but these characteristics cannot be determined from the study.

**Table 3. Summary of Key Study Characteristics**

Study; Trial	Countries	Sites	Dates	Participants	Interventions	
					<b>Active</b>	<b>Comparator</b>
Grenach (2015)	US		2009-2011	Women with an adnexal mass (n=146)	OVA1	ROMA

**Table 4. Summary of Key Study Results**

Study	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)
Grenache (2015)				
OVA1	96.8% (83.3-99.9)	54.8% (45.2-64.1)	36.6% (26.2-48.0)	98.4% (91.6-99.9)
ROMA	83.9% (66.3-94.6)	83.5% (75.4-89.8)	57.8% (42.2-72.3)	95.1% (88.8-98.4)

CI: confidence interval; PPV: positive predictive value; NPV: negative predictive value.

The purpose of the limitation tables (see Tables 5 and 6) is to display notable gaps identified in each study. This information is synthesized as a summary of the body of evidence following each table and provides conclusions on the sufficiency of evidence supporting the position statement.

**Table 5. Relevance Limitations**

Study	Population <sup>a</sup>	Intervention <sup>b</sup>	Comparator <sup>c</sup>	Outcomes <sup>d</sup>	Follow-Up <sup>e</sup>
FDA(k) OVA1 Test K081754	1. Some patients were not evaluated by a gynecologic oncologist; 3. Unclear how patients were recruited; 4. Enrollment was limited to patients with planned surgical intervention				
Grenache et al (2015)	1. Patients were not evaluated by a gynecologic oncologist; 4. Enrollment included only patients with planned surgical intervention, due to the small number of women with malignant adnexal				

masses, the strength of conclusions was limited

The evidence limitations stated in this table are those notable in the current review; this is not a comprehensive limitations assessment.  
<sup>a</sup> 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.  
<sup>b</sup> Intervention key: 1. Not clearly defined; 2. Version used unclear; 3. Delivery not similar intensity as comparator; 4. Not the intervention of interest.  
<sup>c</sup> Comparator key: 1. Not clearly defined; 2. Not standard or optimal; 3. Delivery not similar intensity as intervention; 4. Not delivered effectively.  
<sup>d</sup> Outcomes key: 1. Key health outcomes not addressed; 2. Physiologic measures, not validated surrogates; 3. No CONSORT reporting of harms; 4. Not establish and validated measurements; 5. Clinical significant difference not prespecified; 6. Clinical significant difference not supported.  
<sup>e</sup> Follow-Up key: 1. Not sufficient duration for benefit; 2. Not sufficient duration for harms.

**Table 6. Study Design and Conduct Limitations**

Study	Allocation <sup>a</sup>	Blinding <sup>b</sup>	Selective Reporting <sup>c</sup>	Follow-Up <sup>d</sup>	Power <sup>e</sup>	Statistical <sup>f</sup>
FDA(k) OVA1 Test K081754	1. Subjects were not allocated randomly			1. 10% of subjects were eliminated due to missing information or lack of sample		
Grenache et al (2015)	1. Subjects were not allocated randomly	1,2. Treatment assignment and outcome assessment were not blinded				

The evidence limitations stated in this table are those notable in the current review; this is not a comprehensive limitations assessment.  
<sup>a</sup> Allocation key: 1. Participants not randomly allocated; 2. Allocation not concealed; 3. Allocation concealment unclear; 4. Inadequate control for selection bias.  
<sup>b</sup> Blinding key: 1. Not blinded to treatment assignment; 2. Not blinded outcome assessment; 3. Outcome assessed by treating physician.  
<sup>c</sup> Selective Reporting key: 1. Not registered; 2. Evidence of selective reporting; 3. Evidence of selective publication.  
<sup>d</sup> Follow-Up key: 1. High loss to follow-up or missing data; 2. Inadequate handling of missing data; 3. High number of crossovers; 4. Inadequate handling of crossovers; 5. Inappropriate exclusions; 6. Not intent to treat analysis (per protocol for noninferiority trials).  
<sup>e</sup> Power key: 1. Power calculations not reported; 2. Power not calculated for primary outcome; 3. Power not based on clinically important difference.  
<sup>f</sup> Statistical key: 1. Intervention is not appropriate for outcome type: (a) continuous; (b) binary; (c) time to event; 2. Intervention is not appropriate for multiple observations per patient; 3. Confidence intervals and/or p values not reported; 4. Comparative treatment effects not calculated.

## Overa Test

Descriptions of the developmental process for the Overa test have been published in FDA documents.<sup>17</sup> The FDA documents do not provide details on how biomarkers were selected. The test, as cleared by the FDA, is performed on a blood sample, which is to be sent to a reference laboratory for testing using the 5 immunoassays previously described. Results of the 5 determinations are entered into a proprietary algorithm, called OvaCalc software (v4.0.0), which combines the 5 discrete values into a single unitless numeric score from 0.0 to 10.0.

Clinical validity 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>18</sup> Overa test scores were determined based on the analysis of archived serum specimens from a previous study,<sup>20</sup> and the patient was stratified into a low- or high-risk group for finding malignancy on surgery. The analysis examined whether patient referral to a gynecologic oncologist were 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 clinical assessment by nongynecologic oncologists, was as follows (see Table 7). 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. 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%.

**Table 7. Clinical Validity of the Overa Test Among 493 Patients Evaluated by Nongynecologic Oncologists**

Diagnostic Characteristics	Clinical Assessment Alone, %	Clinical Assessment with OVA1, %
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)
Prevalence	19 (92/493)	

Adapted from U.S. Food and Drug Administration (FDA). 510(k) Substantial Equivalence Determination Decision Summary: OVA1™ Next Generation Test (K150588). n.d.;  
CI: confidence interval

The purpose of the limitation’s tables (see Tables 8 and 9) is to display notable limitations identified in each study. This information is synthesized as a summary of the body of evidence following each table and provides conclusions on the sufficiency of evidence supporting the position statement.

**Table 8. Relevance Limitations**

Study	Population <sup>a</sup>	Intervention <sup>b</sup>	Comparator <sup>c</sup>	Outcomes <sup>d</sup>	Follow-Up <sup>e</sup>
FDA 510(k) OVA1 K1505881	4. 70.3% of subjects were white				

The evidence limitations stated in this table are those notable in the current review; this is not a comprehensive limitations assessment.

<sup>a</sup> 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.

<sup>b</sup> Intervention key: 1. Not clearly defined; 2. Version used unclear; 3. Delivery not similar intensity as comparator; 4. Not the intervention of interest.

<sup>c</sup> Comparator key: 1. Not clearly defined; 2. Not standard or optimal; 3. Delivery not similar intensity as intervention; 4. Not delivered effectively.

<sup>d</sup> Outcomes key: 1. Key health outcomes not addressed; 2. Physiologic measures, not validated surrogates; 3. No CONSORT reporting of harms; 4. Not establish and validated measurements; 5. Clinical significant difference not prespecified; 6. Clinical significant difference not supported.

<sup>e</sup> Follow-Up key: 1. Not sufficient duration for benefit; 2. Not sufficient duration for harms.

**Table 9. Study Design and Conduct Limitations**

Study	Selection <sup>a</sup>	Blinding <sup>b</sup>	Selective Reporting <sup>c</sup>	Delivery of Test <sup>d</sup>	Data Completeness <sup>e</sup>	Statistical <sup>f</sup>
FDA 510(k) OVA1 K1505881	1. Not described	1. Not described	1. Registration not described	1. Not described	1. Inadequate description of indeterminate and missing samples	None

The evidence limitations stated in this table are those notable in the current review; this is not a comprehensive limitations assessment.

<sup>a</sup> Allocation key: 1. Participants not randomly allocated; 2. Allocation not concealed; 3. Allocation concealment unclear; 4. Inadequate control for selection bias.

<sup>b</sup> Blinding key: 1. Not blinded to treatment assignment; 2. Not blinded outcome assessment; 3. Outcome assessed by treating physician.

<sup>c</sup> Selective Reporting key: 1. Not registered; 2. Evidence of selective reporting; 3. Evidence of selective publication.

<sup>d</sup> 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.  
<sup>e</sup> Data Completeness: 1. Inadequate description of indeterminate and missing samples; 2. High number of samples excluded; 3. High loss to follow-up or missing data  
<sup>f</sup> Statistical key: 1. Intervention is not appropriate for outcome type: (a) continuous; (b) binary; (c) time to event; 2. Intervention is not appropriate for multiple observations per patient; 3. Confidence intervals and/or p values not reported; 4. Comparative treatment effects not calculated.

## ROMA Test

Moore et al (2008) described the development of the ROMA test.<sup>21</sup> The authors studied 9 biomarkers and chose human epididymis protein 4 (HE4) and CA 125 because these markers in tandem produced the best performance. The algorithm developed was subsequently modified to include menopausal status and was independently validated.<sup>22</sup> Again, separate cutoffs were used for premenopausal and postmenopausal women.

## ROMA Compared with HE4 and CA 125

Three systematic reviews have assessed the diagnostic accuracy of ROMA in comparison with CA 125 and HE4 through metaanalysis.<sup>23,24,25</sup> Study characteristics are summarized in Table 10. Across analyses, there was little variability in estimates of sensitivity and specificity, and the area under the receiver operating characteristic (AUROC) (Table 11). ROMA sensitivities (range 85.3% to 87.3%) were higher than those for CA 125 (range 76.3% to 84.0%) and HE4 (range 68.2% to 76.3%). HE4 was associated with higher specificities (range 85.1% to 93.6%) than both ROMA (range 79.0% to 85.5%) and CA 125 (range 73.0% to 82.5%). ROMA, CA 125, and HE4 all showed excellent discrimination, based on AUROCs of 0.91 to 0.92 for ROMA, 0.86 to 0.89 for CA 125, and 0.87 to 0.91 for HE4.

A sensitivity analysis conducted by Suri et al (2021)<sup>23</sup>, found ROMA had better diagnostic accuracy in postmenopausal women (sensitivity 88%, specificity 83%) than premenopausal women (sensitivity 80%, specificity 80%), and better discrimination (AUROC 0.94 [SE 0.01] and 0.88 [SE 0.01], respectively). The review found no evidence of publication bias, nor did it find differential results when analyses were limited to blinded studies.

**Table 10. Characteristics of Systematic Reviews That Compared ROMA with HE4 and CA 125**

Study	Tests Evaluated (No. Studies)	Reference Standard	Study Populations Included	Study Designs Included
Wang et al (2014)	CA 125 (28), HE4 (28), and ROMA (14)	Pathologic diagnosis	Women with ovarian cancer and benign gynecologic disease	Blinded and unblinded
Dayyani et al (2016)	CA 125 (6), HE4 (6), and ROMA (6)	Pathologic diagnosis	Women with ovarian cancer	All
Suri et al (2021)	CA 125 (26), HE4 (25), and ROMA (22)	Pathologic diagnosis	Women with ovarian cancer or benign ovarian mass	Blinded and unblinded; sensitivity analysis limited to blinded studies

CA 125: cancer antigen 125; HE4: human epididymis secretory protein 4

**Table 11. Meta-Analytic Findings for Diagnostic Performance of the ROMA Test vs. HE4 and CA 125**

Test	Study	No. Studies	Sensitivity (95% CI)	Specificity (95% CI)	AUROC (95% CI)
<b>ROMA</b>					
	Suri 2021 <sup>23</sup> ,	22	86.0 (84.0 to 87.0)	79.0 (78.0 to 80.0)	0.91 (95% CI NR; SE 0.01)
	Dayyani 2016 <sup>24</sup> ,	6	87.3 (75.2 to 94.0)	85.5 (71.9 to 93.2)	0.92 (0.86 to 0.96)

Wang 201425,	14	85.3 (81.2 to 88.6)	82.4 (77.4 to 86.5)	0.91 (0.88 to 0.93)
<b>CA 125</b>				
Suri 202123,	26	84.0 (82.0 to 85.0)	73.0 (72.0 to 74.0)	0.86 (95% CI NR; SE 0.02)
Dayyani 201624,	6	79.6 (66.3 to 88.5)	82.5 (82.5 to 91.9)	0.88 (0.77 to 0.95)
Wang 201425,	28	76.3 (72.0 to 80.1)	82.1 (76.6 to 86.5)	0.89 (0.86-0.92)
<b>HE4</b>				
Suri 202123,	25	73.0 (71.0 to 75.0)	90.0 (89.0 to 91.0)	0.91 (95% CI NR; SE 0.01)
Dayyani 201624,	6	68.2 (69.3 to 90.1)	85.1 (71.6 to 92.8)	0.90 (0.84 to 0.94)
Wang 201425,	28	76.3 (72.0 to 80.1)	93.6 (90.0 to 95.9)	0.87 (0.84-0.90)

AUROC: area under the receiver operating characteristic; CA 125: cancer antigen 125; CI: confidence interval; HE4: human epididymis secretory protein 4; NR: not reported; ROMA: Risk of Ovarian Malignancy Algorithm; SE: standard error.

Since the Wang et al (2014) and Dayyani et al (2016) meta-analyses, multiple studies have compared the use of the ROMA test to HE4 and CA 125 in various subgroups based on menopausal status, the cutoff value used, and different racial/ethnic background.<sup>26-33</sup> These studies demonstrate that ROMA's sensitivity (range, 54.5% to 93%) and specificity (range, 75% to 96%) can vary importantly depending on variation in these factors. For example, in a few recent studies of racial/ethnic subpopulations, ROMA's sensitivity dramatically declined and was lowest when used in a sample of 274 African American women (54.5%; 95% CI 33.7-75.3)<sup>31</sup> and when distinguishing between malignant/borderline vs. benign or between malignant and borderline/benign in a sample of 177 premenopausal Korean women (46.4% and 52.6%, respectively).<sup>30</sup> On the other hand, specificity was highest (95.9%) in a subgroup of 104 postmenopausal women when using a "new optimal cutoff value" of 33.4% instead of 29.9%.<sup>28</sup>

### ROMA compared with Other Risk Indices

Two systematic reviews have compared ROMA to other tests for detection of ovarian cancer (Table 12).<sup>34,35</sup> Chacon et al (2019) conducted a meta-analysis comparing ROMA with Risk Malignancy Index (RMI, a model incorporating menopausal status, ultrasound findings, and serum CA 125 level) for detecting ovarian cancer.<sup>34</sup> Among the 2662 women included in the meta-analysis, 50 percent were premenopausal and 50 percent were postmenopausal. Mean ovarian cancer prevalence was 29% in premenopausal women and 51% in postmenopausal women. The majority of studies were conducted at a single center. Although pooled sensitivities for ROMA (Table 13) were similar to those reported in previous systematic reviews that compared ROMA to HE4 and CA 125, specificities for ROMA were somewhat lower in this meta-analysis (range of 82 to 85% in Wang et al 2014 and the Dayyani et al 2016 meta-analyses, compared with 75 to 78%). However, findings from this meta-analysis should be interpreted with caution due to important limitations including a high-risk of selection bias in most studies and significant unexplained statistical heterogeneity.

Davenport et al (2022) conducted a meta-analysis comparing commonly-used tests, including ROMA, RMI, International Ovarian Tumor Analysis Logistic Regression Model 2 (LR2, a model incorporating menopausal status and ultrasound findings), and Assessment of Different Neoplasias in the adnexa (ADNEX), a model incorporating menopausal status, CA 125, type of center (referral center for gynecologic oncology vs. other), and ultrasound findings.<sup>35</sup> The analysis included 59 studies, 42 of which evaluated ROMA; 32,059 patients (9545 cases of ovarian cancer) were included. Mean ovarian cancer prevalence ranged from 16% to 27% in premenopausal patients and 38% to 55% in postmenopausal women. In general, ROMA and other tests had higher sensitivity than RMI, but carried lower specificity,

particularly in premenopausal women (Table 12). This analysis carries important limitations, including high risk of selection bias, index test- and reference standard-related biases, and heterogeneity.

**Table 12. Characteristics of Systematic Reviews of ROMA Compared with Other Risk Indices**

Study	Dates	Studies	Participants	N (Range)	Design	Risk of Bias
Chacon et al (2019)	2011-2018	8	Patients in whom both ROMA and RMI, were calculated for predicting malignancy in adnexal masses	2,662 (50-1061)	Prospective (7) and retrospective (1) cohort studies	Based on QUADAS-2 assessment, risk of bias was “high in most studies”, due to “selection bias in that they had selected only women who underwent surgery”
Davenport et al (2022)	2009-2019	59	Patients with signs or symptoms suspicious of ovarian cancer in whom 1 or more of ROMA, RMI, LR2, or ADNEX were calculated	32,059 (36-2403)	Prospective (28), retrospective (21), or unclear (9)	Based on QUADAS-2 assessment, risk of bias was: Participant selection domain : high or unclear for applicability in 92% of studies "because study participants did not obviously represent symptomatic women" Index test domain: low risk in 79% of ROMA studies "either because of the prospective nature of studies, or the objective nature of the index test", but high risk for applicability in 100% of studies "because ultrasound was conducted by specialist sonographers or their level of specialization was unclear" Reference standard and target condition domain : unclear (46% of studies) or high risk (3% of studies) of bias "either because minimum length of follow-up for index negatives was not reported at 6 months, or because there was concern that the reference standard outcome was ascertained with knowledge of the index test result", and high or unclear risk for applicability in 85% of studies "because borderline tumors had been excluded from

						analysis or classification of border line tumors for estimation of test accuracy was unclear" Flow and timing domain : unclear risk in 54% of studies "most commonly because of no information about the interval between the index test and the reference standard" and high risk in 22% of studies "because not all participants receiving an indextest received a reference standard"
--	--	--	--	--	--	--

ADNEX, Assessment of Different NEoplasias in the adneXa; LR2, International Ovarian Tumor Analysis Logistic Regression Model 2; QUADAS-2: Quality Assessment of Diagnostic Accuracy Studies 2; ROMA: Risk of Ovarian Malignancy Algorithm; RMI: risk malignancy index.

**Table 13. Diagnostic Performance of ROMA Compared with Other Risk Indices**

Chacon et al (2019)	<i>Premenopausal</i>	<i>Postmenopausal</i>	<i>Premenopausal</i>	<i>Postmenopausal</i>
ROMA	80% (70 to 88%)	87% (78 to 93%)	78% (69 to 85%)	75% (66 to 83%)
RMI	73% (62 to 81%)	77% (65 to 86%)	89% (83 to 93%)	85% (73 to 92%)
Davenport et al (2022)	<i>Premenopausal</i>	<i>Postmenopausal</i>	<i>Premenopausal</i>	<i>Postmenopausal</i>
ROMA	77.4% (72.7 to 81.5%)	90.3% (87.5 to 92.6%)	84.3% (81.2 to 87.0%)	81.5% (76.5 to 85.5%)
RMI	57.2% (50.3 to 63.8%)	78.4% (74.6 to 81.7%)	92.5% (90.3 to 94.2%)	85.4% (82.0 to 88.2%)
LR2	83.3% (74.7 to 89.5%)	94.8% (92.3 to 96.6%)	90.4% (84.6 to 94.1%)	60.6% (50.5 to 69.9%)
ADNEX	95.5% (91.0 to 97.8%)	97.6% (95.6 to 98.7%)	77.8% (67.4 to 85.5%)	55.0% (42.8 to 66.6%)

ADNEX, Assessment of Different NEoplasias in the adneXa; CI: confidence interval; LR2, International Ovarian Tumor Analysis Logistic Regression Model 2; RMI: riskmalignancy index; ROMA: Risk of Ovarian Malignancy Algorithm.

### ROMA in Conjunction with Clinical Assessment

The FDA labeling for ROMA, unlike that for OVA1, does not indicate how ROMA is to be used in conjunction with clinical assessment. All previously cited literature assessed ROMA as a stand-alone test for ovarian cancer and did not provide a comparison with clinical assessment alone. The study by Moore et al (2014) evaluated ROMA in conjunction with clinical assessment, using either a positive clinical assessment or a positive ROMA as a positive test (similar to the recommended usage for OVA1).<sup>36</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 given to patients with a negative clinical assessment. In this study, 461 women were enrolled, of whom 86 (19%) had a malignancy. Combined assessment improved sensitivity from 77.9% to 89.7%, but specificity worsened from 84.3% to 67.2% (see Table 14).

**Table 14. Summary of Key Study Characteristics**

Study	Countries	Sites	Dates	Participants	Comparison
					ROMA Group   Comparator

Moore (2014)	U.S.	13	2009-2010	Women with an ovarian cyst or pelvic mass (n=461)	ICRA+ROMA	ICRA
--------------	------	----	-----------	---	-----------	------

ICRA: Initial Cancer Risk Assessment.

**Table 15. Diagnostic Performance of the ROMA Test for All Malignancy**

Diagnostic Characteristics	Clinical Assessment Alone (95% CI), %	Clinical Assessment with ROMA (95% CI), %
Sensitivity	77.9 (66.2 to 87.1)	89.7 (79.9 to 95.8)
Specificity	84.3 (80.2 to 87.8)	67.2 (62.2 to 71.9)
Positive predictive value	47.3 (37.8 to 57.0)	33.2 (26.4 to 40.5)
Negative predictive value	95.5 (92.6 to 97.4)	97.3 (94.5 to 98.9)

Adapted from Moore et al (2014)<sup>28</sup>  
CI: confidence interval.

The purpose of the limitation tables (see Tables 16 and 17) is to display notable limitations identified in each study. This information is synthesized as a summary of the body of evidence following each table and provides conclusions on the sufficiency of the evidence supporting the position statement.

**Table 16. Relevance Limitations**

Study	Population <sup>a</sup>	Intervention <sup>b</sup>	Comparator <sup>c</sup>	Outcomes <sup>d</sup>	Follow-Up <sup>e</sup>
Moore et al (2014)	4. 84.8% of subjects were white; 60.4% of subjects were EOC grade 3; 66.7% had stage III epithelial ovarian cancer				

The evidence limitations stated in this table are those notable in the current review; this is not a comprehensive limitations assessment.

<sup>a</sup> 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.

<sup>b</sup> Intervention key: 1. Not clearly defined; 2. Version used unclear; 3. Delivery not similar intensity as comparator; 4. Not the intervention of interest.

<sup>c</sup> Comparator key: 1. Not clearly defined; 2. Not standard or optimal; 3. Delivery not similar intensity as intervention; 4. Not delivered effectively.

<sup>d</sup> Outcomes key: 1. Key health outcomes not addressed; 2. Physiologic measures, not validated surrogates; 3. No CONSORT reporting of harms; 4. Not establish and validated measurements; 5. Clinical significant difference not prespecified; 6. Clinical significant difference not supported.

<sup>e</sup> Follow-Up key: 1. Not sufficient duration for benefit; 2. Not sufficient duration for harms.

**Table 17. Study Design and Conduct Limitations**

Study	Selection <sup>a</sup>	Blinding <sup>b</sup>	Selective Reporting <sup>c</sup>	Delivery of Test <sup>d</sup>	Data Completeness <sup>e</sup>	Statistical <sup>f</sup>
Moore et al (2014)					1. Inadequate description of indeterminate and missing samples	

The evidence limitations stated in this table are those notable in the current review; this is not a comprehensive limitations assessment.

<sup>a</sup> Allocation key: 1. Participants not randomly allocated; 2. Allocation not concealed; 3. Allocation concealment unclear; 4. Inadequate control for selection bias.

<sup>b</sup> Blinding key: 1. Not blinded to treatment assignment; 2. Not blinded outcome assessment; 3. Outcome assessed by treating physician.

<sup>c</sup> Selective Reporting key: 1. Not registered; 2. Evidence of selective reporting; 3. Evidence of selective publication.

<sup>d</sup> 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.

<sup>e</sup> 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.

### **Section Summary: Clinically Valid**

Evidence for the clinical validity for the OVA1 and Overa tests include prospective, double-blind studies that have evaluated the clinical validity of these tests in predicting the likelihood of malignancy in women who are planning to have surgery for an adnexal mass. These tests have not been studied for ovarian cancer screening. The prospective studies showed that, in patients with an adnexal mass who had a planned surgical intervention, the use of OVA1 and Overa in conjunction with a clinical assessment by nongynecologic oncologists increased the sensitivity but decreased the specificity compared with clinical assessment alone. When used with clinical assessment in this manner, the sensitivity to ovarian malignancy was 92%, and the specificity was 42%. ROMA is intended for use in conjunction with clinical assessment, but no specific method has been defined. One study, which used clinical assessment and ROMA results, showed a sensitivity of 90% and a specificity of 67%. Multiple meta-analyses have reported less than 90% sensitivity and specificity with ROMA testing.

### **Clinical 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.

The ideal study design to evaluate the clinical utility of multimarker serum-based test would be a randomized controlled trial comparing health outcomes (e.g., mortality) in patients managed using the tests with those managed according to best current clinical practices. According to the chain of logic, greater numbers of people referred for initial surgical treatment with ovarian cancer should result in improved overall health outcomes. No randomized or nonrandomized studies with these comparisons were identified.

OVA1, Overa, and ROMA, when used in conjunction with clinical assessment may improve the sensitivity for detection of malignancy, the specificity declines. In studies using either positive ROMA or clinical assessment as a positive test, sensitivity improved-but it was still less than 90%. It is uncertain whether there is meaningful clinical benefit from using a test that avoids a high number of referrals and does not contain sensitive data (even though incrementally better).

### **Section Summary: Clinically Useful**

As no trials were identified that have compared health outcomes for patients managed with and without the use of FDA-cleared multimarker serum-based tests, there is no direct evidence of clinical usefulness. It is uncertain whether discrimination is sufficient to alter decision-making based on clinical assessment alone, thus offering a meaningful benefit to patients. Therefore, the chain of evidence supporting improved outcomes is incomplete.

### **OvaWatch Testing**

OvaWatch is intended to assist in assessing the risk of ovarian cancer when a healthcare provider's initial clinical assessment indicates that an adnexal mass is either indeterminate or benign. In 2022, Reilly et al examined the potential use of a noninvasive machine learning tool to accurately assess the risk of ovarian malignancy in patients with pelvic masses.<sup>37</sup> MIA3G is a deep feedforward neural network for ovarian cancer risk assessment, using seven protein biomarkers along with age and menopausal status as input features. The algorithm was developed on a heterogeneous data set of 1,067 serum specimens from women with adnexal masses (prevalence = 31.8%). It was subsequently validated on a cohort almost twice that size (N = 2,000). In the analytical validation data set (prevalence = 4.9%), MIA3G

demonstrated a sensitivity of 89.8% and a specificity of 84.02%. The positive predictive value was 22.45%, and the negative predictive value was 99.38%. When stratified by cancer type and stage, MIA3G achieved sensitivity of 94.94% for epithelial ovarian cancer, 76.92% for early-stage cancer, and 98.04% for late-stage cancer.

Reilly et al (2023) reported on the clinical validation of OvaWatch as conservative management of adnexal masses.<sup>38</sup> In retrospective, low prevalence ( $N = 1,453$ , 1.5% malignancy rate) data from patients that received an independent physician assessment of benign, OvaWatch has a sensitivity of 81.8% [95% confidence interval (CI) 65.1–92.7] for identifying a histologically confirmed malignancy, and a negative predictive value (NPV) of 99.7%. OvaWatch identified 18/22 malignancies missed by physician assessment. A prospective data set had 501 patients where 106 patients with adnexal mass went for surgery. The prevalence was 2% (10 malignancies). The sensitivity of OvaWatch for malignancy was 40% (95% CI: 16.8–68.7%), and the specificity was 87% (95% CI: 83.7–89.7) when patients were included in the analysis who did not go to surgery and were evaluated as benign. The NPV remained 98.6% (95% CI: 97.0–99.4%). An independent analysis set with a high prevalence (45.8%) the NPV value was 87.8% (95% CI: 95% CI: 75.8–94.3%).<sup>37</sup>

## SUPPLEMENTAL INFORMATION

### Clinical Input Received through Physician Specialty Societies and Academic Medical Centers

Results of clinical input received by BCBSA in 2012 from subject matter experts revealed mixed support for the use of this test as a tool for triaging patients with an adnexal mass. Reviewers agreed that the evidence was insufficient to determine the impact of these tests on referral patterns. For indications other than triaging patients with an adnexal mass, there was a lack of support for use of these tests.

## PRACTICE GUIDELINES AND POSITION STATEMENTS

### American College of Obstetricians and Gynecologists<sup>38</sup>

In 2017, with reaffirmation in 2024, the American College of Obstetricians and Gynecologists (ACOG) opinion on the role of the obstetrician-gynecologist in the early detection of epithelial ovarian cancer addressed. The opinion states that multimarker panels lack strong evidence for use in asymptomatic women without adnexal masses and do not improve early detection and survival rates in average-risk women. The Society for Gynecologic Oncology endorsed ACOG opinion in 2016, an ACOG Practice Bulletin addressing the evaluation and management of adnexal masses made a level B recommendation (based on limited or inconsistent scientific evidence) that consultation with or referral to a gynecologic oncologist is recommended for premenopausal or postmenopausal with an elevated score on a formal risk assessment test such as the multivariate index assay, risk of malignancy index, or the Risk of Ovarian Malignancy Algorithm, or 1 of the ultrasound-based scoring systems from the International Ovarian Tumor Analysis group. A level C recommendation (based on consensus and expert opinion) was given to using serum biomarker panels as an alternative to cancer antigen 125 (CA 125) level to decide about the referral to a gynecologic oncologist for an adnexal mass requiring surgery. There was no mention of the OvaWatch test in this committee's opinion.

### National Institute for Health and Care Excellence<sup>39</sup>

The National Institute for Health and Care Excellence issued guidance in 2011 (last updated October 2023) on the recognition and management of ovarian cancer. The guideline does not provide any recommendations regarding additional serum marker testing besides testing for serum CA 125 levels in women with symptoms suggestive of ovarian cancer.

## National Comprehensive Cancer Network<sup>40</sup>

The NCCN guidelines on ovarian cancer (v.3.2024) includes the following statement: “It has been suggested that specific biomarkers (serum HE4 [human epididymis secretory protein 4] and CA 125 [cancer antigen 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 [Food and Drug Administration] 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 NCCN guidelines states: “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).”

## U.S. Preventive Services Task Force Recommendations<sup>41</sup>

In 2018, the U.S. Preventive Services Task Force recommended against screening women for ovarian cancer (D recommendation). The task force has not addressed multimarker serum -based testing related to ovarian cancer.

## Ongoing and Unpublished Clinical Trials

Some currently ongoing and unpublished trials that might influence this review are listed in Table 18.

Table 18. Summary of Key Trials

NCT No.	Trial Name	Planned Enrollment	Completion Date
<b>Ongoing</b>			
NCT03010124	Prognostic and predictive biomarkers in ovarian cancers (OvBIOMARK)	350	Sep 2030
<b>Unpublished</b>			
NCT03837327	Clinical Validation of the InterVenn Ovarian CAncer Liquid Biopsy (VOCAL)	1025	Dec 2024

NCT: national clinical trial

---

## Regulatory Status

On July 16, 2009, the OVA1® test (Aspire Labs) was cleared for market by the U.S. Food and Drug Administration (FDA) through the 510(k) process. The intended use of OVA1® is as an aid to further assess the likelihood that malignancy is present when the physician’s independent clinical and radiological evaluation does not indicate malignancy.

In March 2016, a second-generation test called Overa™, in which 2 of the 5 biomarkers in OVA1® are replaced with human epididymis secretory protein 4 and follicle stimulating hormone, was cleared for marketing by FDA through the 510(k) process. Similar to OVA1®, Overa™ generates a low or high risk of malignancy on a scale from 0 to 10.

On September 1, 2011, the Risk of Ovarian Malignancy Algorithm (ROMA™ test, Fujirebio Diagnostics) was cleared for marketing by FDA through the 510(k) process. The intended use of ROMA™ is as an aid, in conjunction with clinical assessment, in assessing whether a

premenopausal or postmenopausal woman who presents with an ovarian adnexal mass is at high or low likelihood of finding malignancy on surgery.  
FDA product code: ONX.

### **Black Box Warning**

On December 10, 2011, the FDA amended its regulation for classifying ovarian adnexal mass assessment score test systems. The change required off-label risks be highlighted by using a black box warning.<sup>16</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. Considering the history and currently unmet medical needs for ovarian cancer testing, the FDA concluded that there is a risk of off-label use of this device.<sup>7</sup> To address this risk, the FDA requires that manufacturers provide notice concerning the risks of off-label uses in the labeling, advertising, and promotional material of ovarian adnexal mass assessment score test systems. Manufacturers must address the following risks:

- Women without adnexal pelvic masses (i.e., for cancer "screening") are not part of the intended use population for the ovarian adnexal mass assessment score test systems. Public health risks associated with false-positive results for ovarian cancer screening tests are well described in the medical literature and include morbidity or mortality associated with unneeded testing and surgery. The risk from false-negative screening results also includes morbidity and mortality due to failure to detect and treat ovarian malignancy.
  - Analogous risks, adjusted for prevalence and types of disease, arise if test results are used to determine the need for surgery in patients who are known to have ovarian adnexal masses.
  - If used outside the "OR" rule that is described in this special control guidance, results from ovarian adnexal mass assessment score test systems pose a risk for morbidity and mortality due to nonreferral for oncologic evaluation and treatment.
- 

## **Government Regulations**

### **National:**

There is no national coverage determination on combined ovarian cancer biomarker tests. Requests would be reviewed on an individual consideration basis. Codes 81500, 81503 and 0003U have a fee schedule attached for January 2023.

### **Local:**

There is no WPS local coverage determination on combined ovarian cancer biomarker tests. Requests would be reviewed on an individual consideration basis.

*(The above Medicare information is current as of the review date for this policy. However, the coverage issues and policies maintained by the Centers for Medicare & Medicare Services [CMS, formerly HCFA] are updated and/or revised periodically. Therefore, the most current CMS information may not be contained in this document. For the most current information, the reader should contact an official Medicare source.)*

---

## **Related Policies**

N/A

---

## **References**

1. Surveillance Epidemiology and End Results (SEER) Program. SEER Stat Fact Sheets: Ovarian Cancer. n.d.; <https://seer.cancer.gov/statfacts/html/ovary.html>. Accessed April 2025.
2. du Bois A, Rochon J, Pfisterer J, et al. Variations in institutional infrastructure, physician specialization and experience, and outcome in ovarian cancer: a systematic review. *Gynecol Oncol*. Feb 2009;112(2):422-436. PMID 18990435
3. Matthews BJ, Qureshi MM, Fiascone SJ, et al. Racial disparities in non-recommendation of adjuvant chemotherapy in stage II-III ovarian cancer. *Gynecol Oncol*. Jan 2022; 164(1): 27-33. PMID 34785030
4. Zhang C, Zhang C, Wang Q, et al. Differences in Stage of Cancer at Diagnosis, Treatment, and Survival by Race and Ethnicity Among Leading Cancer Types. *JAMA Netw Open*. Apr 01, 2020; 3(4): e202950. PMID 32267515
5. Joslin CE, Brewer KC, Davis FG, et al. The effect of neighborhood-level socioeconomic status on racial differences in ovarian cancer treatment in a population-based analysis in Chicago. *Gynecol Oncol*. Nov 2014; 135(2): 285-91. PMID 25173584
6. Mattei LH, Robb L, Banning K, et al. Enrollment of Individuals from Racial and Ethnic Minority Groups in Gynecologic Cancer Precision Oncology Trials. *Obstet Gynecol*. Oct 01, 2022; 140(4): 654-661. PMID 36075065
7. Zhong P, Yang B, Pan F, et al. Temporal trends in Black-White disparities in cancer surgery and cancer-specific survival in the United States between 2007 and 2015. *Cancer Med*. Feb 2023; 12(3): 3509-3519. PMID 35968573
8. Ellis L, Canchola AJ, Spiegel D, et al. Racial and Ethnic Disparities in Cancer Survival: The Contribution of Tumor, Sociodemographic, Institutional, and Neighborhood Characteristics. *J Clin Oncol*. Jan 01, 2018; 36(1): 25-33. PMID 29035642
9. Albright BB, Nasioudis D, Craig S, et al. Impact of Medicaid expansion on women with gynecologic cancer: a difference-in-difference analysis. *Am J Obstet Gynecol*. Feb 2021; 224(2): 195.e1-195.e17. PMID 32777264
10. Bodurtha Smith AJ, Applebaum J, Fader AN. Association of the Affordable Care Act's Medicaid Expansion With 1-Year Survival Among Patients with Ovarian Cancer. *Obstet Gynecol*. Jun 01, 2022; 139(6): 1123-1129. PMID 35675609
11. Smick AH, Holbert M, Neff R. Association of Physician Densities and Gynecologic Cancer Outcomes in the United States. *Obstet Gynecol*. Nov 01, 2022; 140(5): 751-757. PMID 36201771
12. 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 Obstet Gynecol*. Jul 2010; 36(1): 81-7. PMID 20217895
13. Eskander R, Berman M, Keder L. Practice Bulletin No. 174: Evaluation and Management of Adnexal Masses. *Obstet Gynecol*. Nov 2016; 128(5): e210-e226. PMID 27776072
14. Simmons AR, Clarke CH, Badgwell DB, et al. Validation of a Biomarker Panel and Longitudinal Biomarker Performance for Early Detection of Ovarian Cancer. *Int J Gynecol Cancer*. Jul 2016; 26(6): 1070-7. PMID 27206285
15. Yanaranop M, Tiyyayon J, Siricharoenchai S, et al. Rajavithi-ovarian cancer predictive score (R-OPS): A new scoring system for predicting ovarian malignancy in women presenting with a pelvic mass. *Gynecol Oncol*. Jun 2016; 141(3): 479-484. PMID 26996662
16. Guidance for Industry and FDA Staff - Class II Special Controls Guidance Document: Ovarian Adnexal Mass Assessment Score Test System. <https://www.fda.gov/regulatory-information/search-fda-guidance-documents/class-ii-special-controls-guidance-document-ovarian-adnexal-mass-assessment-score-test-system>. Updated February 27, 2018. Accessed March 2025.

17. Fung ET. A recipe for proteomics diagnostic test development: the OVA1 test, from biomarker discovery to FDA clearance. *Clin Chem*. Feb 2010; 56(2): 327-9. PMID 20110452
18. U.S. Food and Drug Administration (FDA). 510(k) Substantial Equivalence Determination Decision Summary: OVA1TM Test (K081754) n.d.; [https://www.accessdata.fda.gov/cdrh\\_docs/reviews/K081754.pdf](https://www.accessdata.fda.gov/cdrh_docs/reviews/K081754.pdf). Accessed March 2025.
19. Grenache DG, Heichman KA, Werner TL, et al. Clinical performance of two multi-marker blood tests for predicting malignancy in women with an adnexal mass. *Clin Chim Acta*. Jan 01, 2015; 438: 358-63. PMID 25283731
20. Bristow RE, Smith A, Zhang Z, et al. Ovarian malignancy risk stratification of the adnexal mass using a multivariate index assay. *Gynecol Oncol*. Feb 2013; 128(2): 252-9. PMID 23178277
21. Moore RG, Brown AK, Miller MC, et al. The use of multiple novel tumor biomarkers for the detection of ovarian carcinoma in patients with a pelvic mass. *Gynecol Oncol*. Feb 2008; 108(2): 402-8. PMID 18061248
22. Moore RG, Miller MC, Disilvestro P, et al. Evaluation of the diagnostic accuracy of the risk of ovarian malignancy algorithm in women with a pelvic mass. *Obstet Gynecol*. Aug 2011; 118(2 Pt 1): 280-288. PMID 21775843
23. Suri A, Perumal V, Ammali P, et al. Diagnostic measures comparison for ovarian malignancy risk in Epithelial ovarian cancer patients: a meta-analysis. *Sci Rep*. Aug 27, 2021; 11(1): 17308. PMID 34453074
24. 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. *Int J Gynecol Cancer*. Nov 2016; 26(9): 1586-1593. PMID 27540691
25. Wang J, Gao J, Yao H, et al. Diagnostic accuracy of serum HE4, CA125 and ROMA in patients with ovarian cancer: a meta-analysis. *Tumour Biol*. Jun 2014; 35(6): 6127-38. PMID 24627132
26. 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 Med J*. Sep 2016; 31(5): 336-44. PMID 27602187
27. 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. *J Korean Med Sci*. Dec 2015; 30(12): 1777-83. PMID 26713052
28. Terlikowska KM, Dobrzycka B, Witkowska AM, et al. Preoperative HE4, CA125 and ROMA in the differential diagnosis of benign and malignant adnexal masses. *J Ovarian Res*. Jul 19, 2016; 9(1): 43. PMID 27436085
29. Shen Y, Zhao L, Lu S. Diagnostic performance of HE4 and ROMA among Chinese women. *Clin Chim Acta*. Jan 2020; 500: 42-46. PMID 31626761
30. Shin KH, Kim HH, Kwon BS, et al. Clinical Usefulness of Cancer Antigen (CA) 125, Human Epididymis 4, and CA72-4 Levels and Risk of Ovarian Malignancy Algorithm Values for Diagnosing Ovarian Tumors in Korean Patients with and Without Endometriosis. *Ann Lab Med*. Jan 2020; 40(1): 40-47. PMID 31432638
31. Dunton C, Bullock RG, Fritsche H. Multivariate Index Assay Is Superior to CA125 and HE4 Testing in Detection of Ovarian Malignancy in African American Women. *Biomark Cancer*. 2019; 11: 1179299X19853785. PMID 31236012

32. Han KH, Park NH, Kim JJ, et al. The power of the Risk of Ovarian Malignancy Algorithm considering menopausal status: a comparison with CA 125 and HE4. *J Gynecol Oncol*. Nov 2019; 30(6): e83. PMID 31576682
33. Carreras-Dieguez N, Glickman A, Munmany M, et al. Comparison of HE4, CA125, ROMA and CPH-I for Preoperative Assessment of Adnexal Tumors. *Diagnostics (Basel)*. Jan 17, 2022; 12(1). PMID 35054393
34. Chacón E, Dasí J, Caballero C, et al. Risk of Ovarian Malignancy Algorithm versus Risk Malignancy Index-I for Preoperative Assessment of Adnexal Masses: A Systematic Review and Meta-Analysis. *Gynecol Obstet Invest*. 2019; 84(6): 591-598. PMID 31311023
35. Davenport C, Rai N, Sharma P, et al. Menopausal status, ultrasound and biomarker tests in combination for the diagnosis of ovarian cancer in symptomatic women. *Cochrane Database Syst Rev*. Jul 26, 2022; 7(7): CD011964. PMID 35879201
36. Moore RG, Hawkins DM, Miller MC, et al. Combining clinical assessment and the Risk of Ovarian Malignancy Algorithm for the prediction of ovarian cancer. *Gynecol Oncol*. Dec 2014; 135(3): 547-51. PMID 25449569
37. Committee Opinion No. 716: The Role of the Obstetrician-Gynecologist in the Early Detection of Epithelial Ovarian Cancer in Women at Average Risk. *Obstet Gynecol*. Sep 2017; 130(3): e146-e149. PMID 28832487
38. American College of Obstetricians Gynecologists Committee on Gynecologic Practice. Practice bulletin number 174: Evaluation and Management of Adnexal Masses. <https://www.acog.org/clinical/clinical-guidance/practice-bulletin/articles/2016/11/evaluation-and-management-of-adnexal-masses>. Accessed March 2025.
39. National Center for Clinical Excellence (NICE). Ovarian cancer: recognition and initial management [CG122]. 2011; <https://www.nice.org.uk/guidance/cg122>. Accessed March 2025.
40. National Comprehensive Cancer Network (NCCN). NCCN Clinical Practice Guidelines in Oncology: Ovarian Cancer Including Fallopian Tube Cancer and Primary Peritoneal Cancer. Version 3.2024. [https://www.nccn.org/professionals/physician\\_gls/pdf/ovarian.pdf](https://www.nccn.org/professionals/physician_gls/pdf/ovarian.pdf). Accessed March 2025.
41. Grossman DC, Curry SJ, Owens DK, et al. Screening for Ovarian Cancer: US Preventive Services Task Force Recommendation Statement. *JAMA*. Feb 13, 2018; 319(6): 588-594. PMID 29450531
42. Blue Cross Blue Shield Association. Proteomics-based testing Related to Ovarian Cancer. Medical Policy Reference Manual. Policy #2.04.62. Issue 4:2015. Original policy date 4/8/10. Last review date January 2025.
43. HAYES GTE Report. OVA1 Ovarian Tumor Triage Test. Lansdale, PA: HAYES, Inc. August 21, 2014. Last updated August 2015.
44. HAYES Medical Technology Directory. Architect HE4 Assay for Ovarian Cancer Detection. Lansdale, PA: HAYES, Inc. Published December 2012. Last updated December 2015.
45. BCBSA TEC Assessment Program. Multianalyte Testing for the Evaluation of Adnexal Masses. Volume 27, Number 8, April 2013.

*The articles reviewed in this research include those obtained in an Internet based literature search for relevant medical references through March 2025, the date the research was completed.*

## Joint BCBSM/BCN Medical Policy History

Policy Effective Date	BCBSM Signature Date	BCN Signature Date	Comments
9/1/10	7/22/10	6/29/10	Joint policy established
7/1/12	5/15/12	5/15/12	Policy updated; status changed from experimental and investigational to established. Added information regarding the ROMA™ test. Title changed from “Evaluation of Ovarian (Adnexal) Masses (OVA1) by Proteomics-based Testing to Evaluation of Ovarian (Adnexal) Masses By Proteomics-Based Testing (e.g., OVA1™ and ROMA™ Testing). Added additional references and rationale to include ROMA testing. Reformatted to match BCBSA policy.
9/1/13	4/16/13	4/23/13	Routine update. Policy status changed to experimental/ investigational based on new evidence and BCBSA policy position change. References updated.
1/1/15	10/21/14	11/7/14	Policy status changed back to established for patients meeting FDA guidelines.
7/1/16	4/19/16	6/23/16	Routine policy maintenance. Updated references and rationale section. No change in policy status.
7/1/17	4/18/17	4/18/17	Policy title changed to Multimarker Serum Based Testing Related to Ovarian Cancer to mirror BCBSA policy of same title. Code 0003U added to policy. Added references and rationale. No change in policy status.
7/1/18	4/17/18	4/17/18	Rationale updated, added reference # 1, 10, 12, 16 and 27. Added Overa to policy, policy statement and title.
7/1/19	4/16/19		Routine policy maintenance. No change in policy status.
7/1/20	4/14/20		Routine policy maintenance. Added references 23-27. No change in policy status.

7/1/21	4/20/21		Routine policy maintenance. No change in policy status.
7/1/22	4/19/22		Routine policy maintenance, no change in policy status.
7/1/23	4/26/23		Reviewed materials sent from Aspira Women's Health. Added references 35 & 36. Code 0375U added as E/I effective 4/1/23. No change in policy status. Vendor managed: N/A. (ds)
7/1/24	4/16/24		Rationale updated, reference #28 added. No change in policy status. Vendor managed: N/A (ds)
7/1/25	4/15/25		Minor changes in MPS and inclusion section made. No change in policy status. Vendor managed: N/A (ds)
1/1/26	10/21/25		Code update: added code 0577U as E/I. Formatting changes made. Vendor managed: N/A (ds)

Next Review Date: 2<sup>nd</sup> Qtr. 2026

**BLUE CARE NETWORK BENEFIT COVERAGE**  
**POLICY: MULTIMARKER SERUM-BASED TESTING RELATED TO OVARIAN CANCER**  
**(E.G., OVA1® , OVERA™ , OVAWATCH AND ROMA™ TESTING)**

**I. Coverage Determination:**

<b>Commercial HMO (includes Self-Funded groups unless otherwise specified)</b>	Covered; criteria apply
<b>BCNA (Medicare Advantage)</b>	See government section
<b>BCN65 (Medicare Complementary)</b>	Covered if primary Medicare covers the service.

**II. Administrative Guidelines:**

- The member's contract must be active at the time the service is rendered.
- The service must be authorized by the member's PCP except for Self-Referral Option (SRO) members seeking Tier 2 coverage.
- Services must be performed by a BCN-contracted provider, if available, except for Self-Referral Option (SRO) members seeking Tier 2 coverage.
- Payment is based on BCN payment rules, individual certificate and certificate riders.
- Appropriate copayments will apply. Refer to certificate and applicable riders for detailed information.
- CPT - HCPCS codes are used for descriptive purposes only and are not a guarantee of coverage.