2.04.62	Multimarker Serum Testing Related to Ovarian Cancer		
Original Policy Date:	October 15, 2007	**Effective Date:**	October 1, 2025
Section:	2.0 Medicine	**Page:**	Page 1 of 22

Policy Statement

I. All uses of the OVA1, Overa, and ROMA tests are considered **investigational**, including but not limited to:
   A. Preoperative evaluation of adnexal masses to triage for malignancy
   B. Screening for ovarian cancer
   C. Selecting individuals for surgery for an adnexal mass
   D. Evaluation of individuals with clinical or radiologic evidence of malignancy
   E. Evaluation of individuals 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: Refer to Appendix A to see the policy statement changes (if any) from the previous version.

Policy Guidelines

OVA1, Overa, and ROMA tests are combinations of several separate lab tests and involve proprietary algorithms for determining risk (ie, what CPT calls multianalyte assays with algorithmic analyses [MAAAs]). Ova1Plus is a proprietary reflex process combining two FDA-cleared tests, Ova1, leveraging high sensitivity, and Overa. No separate evidence was identified for Ova1Plus and as both of the individual tests are included within the policy no additional evidence review provided at this time. OvaWatch is a multivariate index assay that provides a single risk assessment score; currently, an FDA submission is in process and evidence review will be considered if it is cleared.

Coding
See the Codes table for details.

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.

Summary of Evidence

For individuals who have adnexal mass(es) undergoing surgery for possible ovarian cancer who receive multimarker serum testing with clinical assessment preoperatively to assess ovarian cancer risk, the evidence includes studies assessing technical performance and diagnostic accuracy. Relevant outcomes are overall survival and test accuracy. OVA1 and Overa are intended for use in patients for whom clinical assessment does not clearly indicate cancer. When used in this manner, sensitivity for ovarian malignancy was 92% and specificity was 42% with OVA1; with Overa, sensitivity was 94% and specificity was 65%. ROMA is intended for use 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 specificity of 67%. However, the National Comprehensive Cancer Network guidelines recommend (category 2A) that all patients with suspected ovarian cancer should be evaluated by an experienced gynecologic oncologist. Given the National Comprehensive Cancer Network recommendation, direct evidence will be required to demonstrate that the use of U.S. Food and Drug Administration (FDA) cleared multimarker serum testing to inform decisions regarding referral to a gynecologic oncology specialist for surgery has clinical usefulness. Direct evidence of clinical usefulness is provided by studies that have compared health outcomes for patients managed with and without the FDA cleared multimarker serum testing. Because these are intervention studies, the preferred evidence would be from randomized controlled trials. No trials were identified that have evaluated whether referral based on FDA cleared multimarker serum testing improves health outcomes. The evidence is insufficient to determine that the technology results in an improvement in the net health outcome.

Additional Information

In response to requests, clinical input was received while this policy was under review in 2012. The input was mixed in support of these tests 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 the use of these tests.

Related Policies

- Serum Biomarker Human Epididymis Protein 4

Benefit Application

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

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

Regulatory Status

SB 496

SB 496 requires health plans licensed under the Knox-Keene Act ("Plans"), Medi-Cal managed care plans ("MCPS"), and health insurers ("Insurers") to cover biomarker testing for the diagnosis, treatment, appropriate management, or ongoing monitoring of an enrollee's disease or condition to guide treatment decisions, as prescribed. The bill does not require coverage of biomarker testing for screening purposes. Restricted or denied use of biomarker testing for these purposes is subject to

state and federal grievance and appeal processes. Where biomarker testing is deemed medically necessary, Plans and Insurers must ensure that the testing is provided in a way that limits disruptions in care.

FDA Clearances

In July 2009, the OVA1® test (Aspira Labs [Austin, TX]) was cleared for marketing by the FDA through the 510(k) process. OVA1® was designed as a tool to further assess the likelihood that malignancy is present when the physician's independent clinical and radiologic evaluation does not indicate malignancy.

In September 2011, the Risk of Ovarian Malignancy Algorithm (ROMA™ test; Fujirebio Diagnostics [Sequin, TX]) was cleared for marketing by the 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 a high or low likelihood of finding malignancy on surgery.

In March 2016, a second-generation test called Overa™ (also referred to as next-generation OVA1®), 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 the 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.

Black Box Warning

In December 2011, the FDA amended its regulation for classifying ovarian adnexal mass assessment score test systems. The change required that off-label risks be highlighted using a black box warning. 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 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.[16,] 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 (ie, 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.

Rationale

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 carcinoma. An estimated 21,410 women in the U.S. were estimated to be diagnosed in 2021 with ovarian cancer, and approximately 13,770 were expected to die of the disease.[1,] 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).[2] In particular, comprehensive staging and completeness of tumor resection appear to have a positive impact on patient outcomes.

Adult women presenting with an adnexal mass have an estimated 68% likelihood of having a benign lesion.[3] About 6% of women with masses have borderline tumors; 22% possess invasive malignant lesions, and 3% have metastatic disease. Surgery is the only way to diagnose ovarian cancer; this is because a biopsy of an ovary with suspected ovarian cancer is usually not performed due to the risk of spreading cancer cells. Most clinicians agree that women with masses that have a high likelihood of malignancy should undergo surgical staging by a gynecologic oncologist. However, women with clearly benign masses do *not* require a referral to see a specialist. Therefore, criteria and tests that help differentiate benign from malignant pelvic masses are 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.[4] 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 CA 125 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 U.S. Food and Drug Administration (FDA) with the intended use of triaging patients with adnexal masses (see Regulatory Status section). These tests 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 referral to a gynecologic oncologist for treatment. The tests have been developed and evaluated only in patients with adnexal masses and planned surgeries. 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; however, these are not yet commercially available.[5,6]

Table 1. Summary of FDA-Cleared 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
b_2-microglobulin	X		
Transferrin	X	X	
Transthyretin	X		
Apolipoprotein AI	X	X	
HE4		X	X
FSH		X	
Score range	0 to 10	0 to 10	0 to 10
Risk categorization			
Premenopausal	<5.0: low ≥5.0: high	<5.0: low ≥5.0: high	≥1.3: high

Variables	OVA1	Overa	ROMA
Postmenopausal	<4.4: low ≥4.4: high		≥2.77: high

CA 125: cancer antigen 125; FDA: U.S. Food and Drug Administration; FSH: follicle-stimulating hormone; HE4: human epididymis secretory protein 4; ROMA: Risk of Ovarian Malignancy Algorithm.

Literature Review

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

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

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 PICO was used to select literature to inform this review.

Populations

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. Food and Drug Administration cleared commercial multimarker serum genetic tests (e.g., OVA1, Overa, Risk of Ovarian Malignancy Algorithm [ROMA]). Multimarker serum testing related to ovarian cancer may be performed at any point when an individual presents with an ovarian adnexal mass for which surgery is planned, in conjunction with a 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.

Comparators

The comparator of interest is a 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 harms from a false-positive could be inappropriate assessment and improper management of individuals with ovarian malignancies, which could result in the following: inappropriate surgical decisions, high frequency of unnecessary further testing, and unnecessary patient anxiety. The potential harms from a false-negative could be a determination that the individual 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 individuals who have not already been identified as needing surgery for pelvic mass, or individuals 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

For the evaluation of clinical validity of the tests within this review, studies that meet the following eligibility criteria were considered:

- Reported on the accuracy of the marketed version of the technology
- Included a suitable reference standard
- Patient/sample characteristics were described
- 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).

Review of Evidence

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).[17,] 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 for malignant ovarian disease. However, the choice of 5 of these (cancer antigen 125 [CA 125], prealbumin, apo AI, b_2-microglobulin, transferrin) produced a composite profile that did appear to have the discriminatory ability. The test, as cleared by the FDA, is performed on a blood sample, which is 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 empirical process. It is a process based on several different factors: the use of banked samples from academic partners; a small prospective study of samples from Europe; and a designated subset of samples from the clinical study used to support the submission to the FDA. It appears that at an undisclosed point in the developmental process, as a result of interaction with the FDA, separate cut points were developed for premenopausal and postmenopausal women.

The clinical validity was evaluated in a prospective, double-blind, clinical study using 27 enrollment sites.[18,] 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 a clinical assessment by nongynecologic oncologists, was as follows in Table 2. The method used for combining clinical assessment and OVA1 results 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 the 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[a] Among 269 Patients Evaluated by Nongynecologic Oncologists

Diagnostic Characteristics	Clinical Assessment Alone, %	Clinical Assessment With OVA1 Test, %
Sensitivity	72	92
Specificity	83	42
Positive predictive value	61	37
Negative predictive value	89	93

Adapted from the FDA. 510(k) Substantial Equivalence Determination Decision Summary: OVA1™ Test (K081754) n.d.; http://www.accessdata.fda.gov/cdrh_docs/reviews/K081754.pdf. Accessed October 31, 2023.
[a] Confidence intervals not provided.
One additional study (by Grenache et al [2015]) was identified; it evaluated the diagnostic performance of the OVA1 test.[19] 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	Dates	Participants	Interventions	
				Active	Comparator
Grenache (2015)[19]	U.S.	2009-2011	Women with an adnexal mass (n=146)	OVA1	ROMA

ROMA: Risk of Ovarian Malignancy Algorithm.

Table 4. Summary of Key Study Results

Study	Sensitivity (95% CI)	Specificity (95% CI)	PPV (95% CI)	NPV (95% CI)
Grenache (2015)[19]				
OVA1	96.8% (83.3 to 99.9)	54.8% (45.2 to 64.1)	36.6% (26.2 to 48.0)	98.4% (91.6 to 99.9)
ROMA	83.9% (66.3 to 94.6)	83.5% (75.4 to 89.8)	57.8% (42.2 to 72.3)	95.1% (88.8 to 98.4)

CI: confidence interval; NPV: negative predictive value; PPV: positive predictive value; ROMA: Risk of Ovarian Malignancy Algorithm.
The purpose of the limitations tables (Tables 5 and 6) 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 the conclusions on the sufficiency of the evidence supporting the position statement.

Table 5. Study Relevance Limitations

Study	Population[a]	Intervention[b]	Comparator[c]	Outcomes[d]	Follow-Up[e]
FDA(k) OVA1 Test K081754[18]	1. Some patients were not evaluated by a gynecologic oncologist; 2. Unclear how patients were recruited; 3. Enrollment was limited to patients with planned surgical intervention 4. Test sample demographics not described; reference values were determined in a sample that was 81.3% White				
Grenache et al (2015)[19]	1. Patients were not evaluated by a gynecologic oncologist; 2. 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 4. Sample demographics not described				

FDA: U.S. Food and Drug Administration.
The study limitations stated in this table are those notable in the current review; this is not a comprehensive gaps assessment.

[a] Population key: 1. Intended use population unclear; 2. Clinical context is unclear; 3. Study population is unclear; 4. Study population not representative of intended use.
[b] Intervention key: 1. Not clearly defined; 2. Version used unclear; 3. Delivery not similar intensity as comparator; 4. Not the intervention of interest.
[c] Comparator key: 1. Not clearly defined; 2. Not standard or optimal; 3.Delivery not similar intensity as intervention; 4. Not delivered effectively.
[d] 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.
[e] Follow-Up key: 1. Not sufficient duration for benefit; 2. Not sufficient duration for harms.

Table 6. Study Design and Conduct Limitations

Study	Selection[a]	Blinding[b]	Delivery of Test[c]	Selective Reporting[d]	Data Completeness[e]	Statistical[f]
FDA(k) OVA1 Test K081754[18,]			1. Not described	1. Registration not described	1. 10% of subjects were eliminated due to missing information or lack of sample	
Grenache et al (2015)[19,]		1,2. Treatment assignment and outcome assessment were not blinded			1. Inadequate description of indeterminate and missing samples	

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

Overa Test

Descriptions of the developmental process for the Overa test have been published in FDA documents.[U.S. Food and Drug Administration (FDA). 510(k) Su.... df. Accessed October 28, 2024.] 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.[18,] Overa test scores were determined based on the analysis of archived serum specimens from a previous study,[20,] 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 was supported when a 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 in Table 7. The method used for combining clinical assessment and Overa test results was to consider the test positive if *either* clinical assessment or Overa test was positive. Thus, in practice, 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 the 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, %	Dual Assessment With Overa Test, %
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 the FDA. 510(k) Substantial Equivalence Determination Decision Summary: OVA1™ Next Generation Test (K150588).
CI: confidence interval.
The purpose of the limitations tables (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 the conclusions on the sufficiency of the evidence supporting the position statement.

Table 8. Study Relevance Limitations

Study	Population[a]	Intervention[b]	Comparator[c]	Outcomes[d]	Follow-Up[e]
FDA 510(k) OVA1 Next Generation K150588[18],	4. 70.3% of subjects were white				

FDA: U.S. Food and Drug Administration.
The study limitations stated in this table are those notable in the current review; this is not a comprehensive gaps assessment.
[a] Population key: 1. Intended use population unclear; 2. Clinical context is unclear; 3. Study population is unclear; 4. Study population not representative of intended use.
[b] Intervention key: 1. Not clearly defined; 2. Version used unclear; 3. Delivery not similar intensity as comparator; 4. Not the intervention of interest.
[c] Comparator key: 1. Not clearly defined; 2. Not standard or optimal; 3.Delivery not similar intensity as intervention; 4. Not delivered effectively.
[d] 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.
[e] Follow-Up key: 1. Not sufficient duration for benefit; 2. Not sufficient duration for harms.

Table 9. Study Design and Conduct Limitations

Study	Selection[a]	Blinding[b]	Delivery of Test[c]	Selective Reporting[d]	Data Completeness[e]	Statistical[f]
FDA 510(k) OVA1 Next Generation K150588[18],	1. Not described	1. Not described	1. Not described	1. Registration not described	1. Inadequate description of indeterminate and missing samples	

FDA: U.S. Food and Drug Administration.
The study limitations stated in this table are those notable in the current review; this is not a comprehensive gaps assessment.
[a] Selection key: 1. Selection not described; 2. Selection not random or consecutive (ie, convenience).
[b] Blinding key: 1. Not blinded to results of reference or other comparator tests.
[c] Test Delivery key: 1. Timing of delivery of index or reference test not described; 2. Timing of index and comparator tests not same; 3. Procedure for interpreting tests not described; 4. Expertise of evaluators not

described.
[d] Selective Reporting key: 1. Not registered; 2. Evidence of selective reporting; 3. Evidence of selective publication.
[e] Data Completeness key: 1. Inadequate description of indeterminate and missing samples; 2. High number of samples excluded; 3. High loss to follow-up or missing data.
[f] Statistical key: 1. Confidence intervals and/or p values not reported; 2. Comparison with other tests not reported.

ROMA Test

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

ROMA compared with CA 125 and HE4

Three systematic reviews have assessed the diagnostic accuracy of ROMA in comparison with CA 125 and HE4 through meta analysis.[23,24,25,] 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)[23,] 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 CA 125 and HE4

Study	Tests evaluated (No. Studies)	Reference Standard	Study Populations Included	Study Designs Included
Suri et al (2021)[23,]	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
Dayyani et al (2016)[24,]	CA 125 (6), HE4 (6), and ROMA (6)	Pathologic diagnosis	Women with ovarian cancer	All
Wang et al (2014)[25,]	CA 125 (28), HE4 (28), and ROMA (14)	Pathologic diagnosis	Women with ovarian cancer and benign gynecologic disease	Blinded and unblinded

CA 125: cancer antigen 125; HE4: human epididymis secretory protein 4; ROMA: Risk of Ovarian Malignancy Algorithm.

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

Test	Study	No. Studies	Sensitivity (95% CI)	Specificity (95% CI)	AUROC (95% CI)
ROMA					
	Suri 2021[23,]	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[24,]	6	87.3 (75.2 to 94.0)	85.5 (71.9 to 93.2)	0.92 (0.86 to 0.96)
	Wang 2014[25,]	14	85.3 (81.2 to 88.6)	82.4 (77.4 to 86.5)	0.91 (0.88 to 0.93)
CA 125					
	Suri 2021[23,]	26	84.0 (82.0 to 85.0)	73.0 (72.0 to 74.0)	0.86 (95% CI NR; SE 0.02)

Test	Study	No. Studies	Sensitivity (95% CI)	Specificity (95% CI)	AUROC (95% CI)
	Dayyani 2016[24,]	6	79.6 (66.3 to 88.5)	82.5 (82.5 to 91.9)	0.88 (0.77 to 0.95)
	Wang 2014[25,]	28	76.3 (72.0 to 80.1)	82.1 (76.6 to 86.5)	0.89 (0.86-0.92)
HE4					
	Suri 2021[23,]	25	73.0 (71.0 to 75.0)	90.0 (89.0 to 91.0)	0.91 (95% CI NR; SE 0.01)
	Dayyani 2016[24,]	6	68.2 (69.3 to 90.1)	85.1 (71.6 to 92.8)	0.90 (0.84 to 0.94)
	Wang 2014[25,]	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 individual 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 backgrounds.[26,27,28,29,30,31,32,33,] 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 to 75.3)[31,] and when distinguishing between malignant/borderline versus benign or between malignant and borderline/benign in a sample of 177 premenopausal Korean women (46.4% and 52.6%, respectively).[30,] 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%.[28,]

ROMA compared with Other Risk Indices

Two systematic reviews have compared ROMA to other tests for detection of ovarian cancer (Table 12).[34,35,] 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.[34,] 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.[35,] 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 13). 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)[34,]	2011-2018	8	Patients in whom both ROMA and RMI were calculated for predicting malignancy in adnexal masses	2662 (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)[35,]	2009-2019	59	Patients with signs or symptoms suspicious for 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 borderline 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 index test 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

Test	Sensitivity (95% CI), %		Specificity (95% CI), %	
Chacon et al (2019)[34,]	*Premenopausal*	*Postmenopausal*	*Premenopausal*	*Postmenopausal*
ROMA	80% (70 to 88%)	87% (78 to 93%)	78% (69 to 85%)	75% (66 to 83%)

Test	Sensitivity (95% CI), %		Specificity (95% CI), %	
RMI	73% (62 to 81%)	77% (65 to 86%)	89% (83 to 93%)	85% (73 to 92%)
Davenport et al (2022)[35,]	*Premenopausal*	*Postmenopausal*	*Premenopausal*	*Postmenopausal*
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: risk malignancy 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).[36,] 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% (Table 14).

Table 14. Summary of Key Study Characteristics

Study; Trial	Countries	Sites	Dates	Participants	Comparison	
					ROMA Group	Comparator
Moore (2014)[36,]	U.S.	13	2009-2010	Women with an ovarian cyst or pelvic mass (n=461)	ICRA + ROMA	ICRA

ICRA: Initial Cancer Risk Assessment; ROMA: Risk of Ovarian Malignancy Algorithm.

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).[36,]
CI: confidence interval; ROMA: Risk of Ovarian Malignancy Algorithm.

The purpose of the limitations tables (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 the conclusions on the sufficiency of the evidence supporting the position statement.

Table 16. Study Relevance Limitations

Study	Population[a]	Intervention[b]	Comparator[c]	Outcomes[d]	Follow-Up[e]
Moore et al (2014)[36,]	4. 84.8% of subjects were white; 60.4% of subjects were EOC grade 3; 66.7% had stage III epithelial ovarian cancer				

EOC: epithelial ovarian cancer.
The study limitations stated in this table are those notable in the current review; this is not a comprehensive gaps assessment.
[a] Population key: 1. Intended use population unclear; 2. Clinical context is unclear; 3. Study population is unclear; 4. Study population not representative of intended use.
[b] Intervention key: 1. Not clearly defined; 2. Version used unclear; 3. Delivery not similar intensity as comparator; 4. Not the intervention of interest.
[c] Comparator key: 1. Not clearly defined; 2. Not standard or optimal; 3.Delivery not similar intensity as intervention; 4. Not delivered effectively.
[d] 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.
[e] Follow-Up key: 1. Not sufficient duration for benefit; 2. Not sufficient duration for harms.

Table 17. Study Design and Conduct Limitations

Study	Selection[a]	Blinding[b]	Delivery of Test[c]	Selective Reporting[d]	Data Completeness[e]	Statistical[f]
Moore et al (2014)[36],					1. Inadequate description of indeterminate and missing samples	

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

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%. Multiplemeta-analyseshave reported less than 90% sensitivity and specificity with ROMA testing.

Clinically Useful

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

The ideal study design to evaluate the clinical utility of a 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 persons with ovarian cancer referred for surgery by a gynecologic oncology

specialist should result in improved overall health outcomes. No randomized or nonrandomized studies with these comparisons were identified.

Although OVA1, Overa, and ROMA, when used in conjunction with clinical assessment, 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). Because there is no established or recommended method for using ROMA in conjunction with clinical assessment, diagnostic performance characteristics are uncertain because it would vary depending on how it is used.

It is also uncertain whether the incremental yield of malignancy resulting from the use of the tests would result in improved patient outcomes. Although prior studies revealed an improvement of outcomes when women with ovarian cancer were initially managed by gynecologic oncologists, it is uncertain whether improved outcomes would occur in the additional cases detected by the use of these tests. These additional cancer cases may differ from other cases detected by clinical assessment alone. If they tend to be earlier stage cancers or biologically less aggressive cancers, initial treatment by a gynecologic oncologist may not provide incremental benefit.

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.

Supplemental Information

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

Clinical Input From Physician Specialty Societies and Academic Medical Centers

While the various physician specialty societies and academic medical centers may collaborate with and make recommendations during this process, through the provision of appropriate reviewers, input received does not represent an endorsement or position statement by the physician specialty societies or academic medical centers, unless otherwise noted.

In response to requests, input was received while this policy was under review in 2012. The input was mixed in support of these tests 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 the use of these tests.

Practice Guidelines and Position Statements

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

American College of Obstetricians and Gynecologists

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 using multimarker serum testing.[37,] 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 this 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.[13,] 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.

National Institute for Health and Care Excellence

In 2011, the National Institute for Health and Care Excellence issued guidance on the identification and management of ovarian cancer.[38,] 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

The National Comprehensive Cancer Network (NCCN) guideline on ovarian cancer (v.3.2024) includes the following statement[39,]:

The FDA has approved the use of ROMA, OVA1, and OVERA for estimating the risk for ovarian cancer in women with an adnexal mass for which surgery is planned, and have not been referred to an oncologist. Although the American Congress of Obstetricians and Gynecologists (ACOG) has suggested that ROMA and OVA1 may be useful for deciding which patients to refer to a gynecologic oncologist, other professional organizations have been non-committal. Not all studies have found that multi-biomarker assays improve all metrics (ie, sensitivity, specificity, positive predictive value, negative predictive value) for prediction of malignancy compared with other methods (eg, imaging, single-biomarker tests, symptom index/clinical assessment). Currently, the NCCN Panel does not recommend the use of these biomarker tests for determining the status of an undiagnosed adnexal/pelvic mass.

In addition, the guideline states "based on data documenting increased survival, the NCCN Guidelines Panel recommends that all patients with suspected ovarian malignancies (especially those with an adnexal mass) should undergo evaluation by an experienced gynecologic oncologist prior to surgery."

U.S. Preventive Services Task Force Recommendations

In 2018, the U.S. Preventive Services Task Force recommended against screening asymptomatic women for ovarian cancer (D recommendation). [40,] The Task Force has not addressed multimarker serum testing related to ovarian cancer.

Medicare National Coverage

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

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
Ongoing			
NCT03837327	Clinical Validation of the InterVenn Ovarian CAncer Liquid Biopsy (VOCAL)	1025	Jan 2024

References

1. Surveillance Epidemology and End Results (SEER) Program. SEER Stat Fact Sheets: Ovarian Cancer. n.d.; https://seer.cancer.gov/statfacts/html/ovary.html. Accessed October 30, 2024.
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): 422-36. 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, Tiyayon J, Siricharoenthai 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 October 31, 2024.
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. Accessed October 28, 2024.
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, Ammalli 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. National Center for Clinical Excellence (NICE). Ovarian cancer: recognition and initial management [CG122]. 2011; https://www.nice.org.uk/guidance/cg122. Accessed October 21, 2024.
39. National Comprehensive Cancer Network (NCCN). NCCN Clinical Practice Guidelines in Oncology: Ovarian Cancer Including Fallopian Tub Cancer and Primary Peritoneal Cancer. Version 3.2024. https://www.nccn.org/professionals/physician_gls/pdf/ovarian.pdf. Accessed October 31, 2024.
40. 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

Documentation for Clinical Review

- No records required

Coding

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

Type	Code	Description
CPT®	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
	81500	Oncology (ovarian), biochemical assays of two proteins (CA-125 and HE4), utilizing serum, with menopausal status, algorithm reported as a risk score
	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
HCPCS	None	

Policy History

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

Effective Date	Action
10/15/2007	BCBSA Medical Policy adoption

Effective Date	Action
10/01/2010	Policy Revision with title change from Analysis of Proteomic Patterns in Serum to Identify Cancer
03/25/2011	Administrative Review
02/22/2013	Coding Update
09/27/2013	Policy revision with position change effective 12/19/2013
12/19/2013	Policy revision with position change
06/30/2015	Coding update
01/01/2016	Coding update
09/01/2016	Policy title change from Proteomic Pattern Analysis in Serum to Identify Cancer. Policy revision without position change.
02/01/2017	Policy title change from Proteomics-Based Testing Related to Ovarian Cancer. Policy revision without position change.
02/01/2018	Policy revision without position change
03/01/2019	Policy revision without position change
02/01/2020	Annual review. No change to policy statement. Literature review updated.
02/01/2021	Annual review. No change to policy statement. Literature review updated.
02/01/2022	Annual review. No change to policy statement. Literature review updated.
02/01/2023	Annual review. Policy statement and literature review updated.
06/01/2023	Coding update
10/01/2025	Policy reactivated. Previously archived from 08/01/2023 to 09/30/2025

Definitions of Decision Determinations

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

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

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

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

B. The scientific evidence must permit conclusions concerning the effect of the technology on health outcomes.

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

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

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

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

Feedback

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

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

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

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

Appendix A

POLICY STATEMENT	
BEFORE	**AFTER** **Blue font: Verbiage Changes/Additions**
Reactivated Policy **Policy Statement:** N/A	**Multimarker Serum Testing Related to Ovarian Cancer 2.04.62** **Policy Statement:** I. All uses of the OVA1, Overa, and ROMA tests are considered **investigational**, including but not limited to: A. Preoperative evaluation of adnexal masses to triage for malignancy B. Screening for ovarian cancer C. Selecting individuals for surgery for an adnexal mass D. Evaluation of individuals with clinical or radiologic evidence of malignancy E. Evaluation of individuals 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