Regence Oregon and Utah | Regence Idaho and select counties of Washington

Medical Policy Manual

Topic: Proteomics-based Testing Related to Ovarian Cancer

Date of Origin: November 2012

Section: Laboratory

Last Reviewed Date: December 2013

Policy No: 60

Effective Date: March 1, 2014

IMPORTANT REMINDER

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

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

DESCRIPTION

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

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

Proteomic tests have been proposed to triage patients with malignant versus benign adnexal masses. A suggested use of the tests is to identify women who have a higher likelihood of malignant disease and may benefit from referral to a gynecologic-oncology specialist. These tests are combinations of several separate lab tests known as multi-analyte assays with algorithmic analyses (MAAA) and are performed on a blood sample by a reference laboratory using a proprietary algorithm. The OVA1™ test algorithm uses five biomarkers, CA-125, prealbumin, apolipoprotein A-1, beta 2 microglobulin, and transferrin. The Risk of Ovarian Malignancy Algorithm (ROMA™) test includes the biomarkers human epididymis secretory protein 4 (HE4) and CA 125, along with menopausal status.

Regulatory Status

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

The OVA1™ test

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

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

The ROMA™ test

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

MEDICAL POLICY CRITERIA

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

1. Preoperative evaluation of adnexal masses to triage for malignancy
2. Screening for ovarian cancer
3. Selecting patients for surgery for an adnexal mass
4. Evaluation of patients with clinical or radiologic evidence of malignancy
5. Evaluation of patients with nonspecific signs or symptoms suggesting possible malignancy
6. Postoperative testing and monitoring to assess surgical outcome and/or to detect recurrent malignant disease following treatment

SCIENTIFIC EVIDENCE

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

- Technical feasibility of a device is typically assessed with 2 types of studies, those that compare test measurements with a gold standard and those that compare results taken with the same device on different occasions (test-retest). Demonstration of technical feasibility should include an assessment of its reproducibility and precision.

- Diagnostic performance is evaluated by the ability of a test to accurately diagnose a clinical condition in comparison with the gold standard. The sensitivity of a test is the ability to detect a disease when the condition is present (true-positive), while specificity indicates the ability to detect patients who are suspected of disease but who do not have the condition (true-negative). Evaluation of diagnostic performance, therefore, requires independent assessment by the 2 methods in a population of patients who are suspected of disease but who do not all have the disease.

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

Literature Appraisal

Technology Assessment

A BlueCross BlueShield Association Technology Evaluation Center (TEC) Assessment was completed in 2012 on "Multi-analyte testing for the evaluation of adnexal masses."[8] The Assessment included evaluation of both the OVA1 and ROMA tests and their impact on health outcomes. The single existing study assessing OVA1 was selected. Studies were selected showing the diagnostic characteristics of ROMA using prespecified cutoff values that assessed diagnostic performance for all types of malignancy, and that did not include healthy subjects as non-malignant control subjects. Since performance of the original trials prepared for submission to the U.S. Food and Drug Administration (FDA), neither test has had evaluation of performance independently confirmed by independent investigators. The following conclusions were made in the TEC Assessment:

Technical Feasibility

Evidence on the technical performance of these tests was evaluated by the U.S. Food and Drug Administration (FDA)[9] and is available through the FDA website. The final analysis indicates acceptable technical performance of OVA1 and ROMA for use in clinical care.[10]

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

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

Diagnostic Performance

Use of the ROMA and OVA1 proteomic tests in combination with clinical assessment appeared to produce very modest changes in diagnostic performance for identifying adnexal masses negative for ovarian cancer.

OVA1

Diagnostic performance of the OVA1 test was evaluated in a prospective, double-blind clinical study using 27 demographically mixed subject enrollment sites.[9] Patients underwent a complete clinical evaluation prior to surgical intervention, and only patients with planned surgical intervention were included in the study. The pre-surgical process for identifying patients for surgery and for establishing a preliminary diagnosis as benign or malignant were not specifically described but were noted to be "based on a variety of clinical assessments." The study did require at least one imaging test be performed within 12 weeks of surgery. Presumably, use of this somewhat non-standardized diagnostic methodology provides information on how the test works in conjunction with real-world decision making. The study enrolled a total of 743 patients with 146 subjects used in the training set and 516 in the testing set. Seventy-four patients were excluded because of missing information or samples. All patients had adnexal masses and were scheduled for surgery. The final prevalence of cancer in the population was 27%.

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

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

* Confidence intervals not provided.

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

In 2013, Bristow and colleagues reported on a prospective non-randomized study of 494 patients evaluated for multivariate index assay (OVA1), CA125-II, and clinical impression.[11] Patients were all scheduled to undergo surgery for an adnexal mass and all were recruited from non-gynecological oncology practices. Authors sought to assess the OVA1 test in determining the need for gynecological oncology referral by comparing OVA1 to clinical assessment and CA 125-II in identifying women with

ovarian cancer. Sensitivity, specificity, positive predictive value, negative predictive value, and likelihood ratios were calculated to estimate the performance of OVA1, CA125, clinical impression, and OVA1 with clinical impression. For ovarian malignancies, authors reported a sensitivity of 95.7% when combined with clinical impression. The negative predictive value was reported at 98.1%. However, both clinical impression and CA125-II were more accurate in identifying benign disease. Similar to the previous study, although sensitivity improved with the OVA1 test, specificity declined and was reported as 53.5% with OVA1 alone and 50.7% with OVA1 combined with clinical impression compared to 92.5% with clinical impression and 86.1% -94.5% for CA125-II in predicting disease.

ROMA

Diagnostic performance of the ROMA test was also evaluated in a prospective, blinded clinical trial using 13 demographically mixed subject enrollment sites with company sponsorship.[10] Patients all presented with an adnexal mass and were scheduled to undergo surgery. An Initial Cancer Risk Assessment (ICRA) was performed to determine the detection of benign versus malignant lesions before testing. The prevalence of cancer was 15%.

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

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

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

The ROMA test did not appear to improve the sensitivity of testing to a great extent.

Several studies have evaluated the diagnostic accuracy of the ROMA test compared to other tools for distinguishing between benign and malignant pelvic masses. Two prospective studies by Moore et al. described the development and utility of the ROMA test, comparing the CA 125 tumor marker to the HE4 marker.[12,13]

- One recent European study[14] demonstrated that in the hands of radiologists at a cancer institute, subjective assessment by ultrasound was superior to ROMA in discriminating benign from malignant adnexal masses. Two additional studies evaluating ROMA testing were performed and both questioned the value of HE4 in combination with the CA 125 marker in identifying ovarian cancers.[15,16]

- More recently, Karlsen and colleagues evaluated 1218 women presenting with pelvic masses using the ROMA test.[17] Prior to diagnosis, HE4 and CA125 levels were obtained, and ROMA and the Risk of Malignancy Index (RMI) (an index consisting of ultrasound findings, menopausal status and CA125 levels) were calculated. At a fixed sensitivity of 94.4%, the specificity of ROMA was 76.5% and the specificity of RMI was 81.5%. At a fixed specificity of

75%, the sensitivity of ROMA was 94.8% and the sensitivity of RMI was 96.0%. Accuracy of ROMA and RMI were not compared statistically, but appeared to be similar.

- In another study, Kaijser and colleagues evaluated 360 women with pelvic masses who were scheduled for surgery. The study compared the diagnostic accuracy of ROMA and an ultrasound-based prediction model (LR2) developed by the International Ovarian Tumor Analysis Study (IOTA). Histology was used as the reference standard. The overall performance of LR2 (94% sensitivity and 82% specificity) was significantly better than ROMA (84% sensitivity and 80% specificity).

There are a limited number of studies comparing diagnostic accuracy of the ROMA and OVA1 proteomic tests.[18] Studies have found that ROMA has similar or lower accuracy to other risk prediction measures that use components of the standard workup, such as the RMI and the LR2 measures. There are fewer published studies evaluating the diagnostic accuracy of OVA1. Further prospective studies are needed for both assays to understand their proper role in patient care.

Clinical Utility

No outcome studies have been performed using the OVA1or the ROMA test. It is not clear what impact either test would have on long-term health outcomes or referral patterns to specialty physicians. The use of proteomic testing to triage patients for malignancy may be only one of many factors in decision making about where treatment should be delivered.

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

Clinical Practice Guidelines and Position Statements

<u>The Society for Gynecologic Oncology (SGO)</u>

The SGO addresses the use of the OVA1 test in a bulletin issued in 2009.[19] This document includes the following statements:

- The OVA1 test…may be a useful tool in identifying women who should be referred to a gynecologic oncologist for their ovarian cancer surgery.
- This test has not been approved for use as an ovarian cancer screening tool, nor has it been proven to result in early detection or reduce the risk of death from this disease.
- Results from the OVA1 test should not be interpreted independently, nor be used in place of a physician's clinical assessment.
- Physicians are strongly encouraged to reference the SGO/ACOG Pelvic Mass Guidelines to determine an appropriate care plan for their patients.

In May 2013, the Society for Gynecologic Oncology (SGO) issued the following consensus based statement on multiplex serum testing for women with pelvic masses:[20]

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

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

The American Congress of Obstetricians and Gynecologists (ACOG)[21]

ACOG addresses the use of the OVA1 test in their opinion statement on the role of the obstetrician-gynecologist in the early detection of epithelial ovarian cancer. This document makes the following statements:

- The OVA1 test appears to improve the predictability of ovarian cancer in women with pelvic masses
- This is not a screening test, but may be useful for evaluating women with a pelvic mass
- Clinical utility is not yet established

The ACOG Committee Opinion Statement regarding the use of the OVA1test is not based upon scientific evidence and does not recommend the use of the OVA1 test as part of their referral guidelines for women who present with an adnexal mass.

The National Cancer Institute (NCI)

The NCI included a discussion of proteomics in their publication on the Genetics of Breast and Ovarian Cancer.[22] The following statement is included in their discussion on proteomics:

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

National Comprehensive Cancer Network (NCCN)

The 2013 version of the NCCN guidelines[23] on ovarian cancer indicated that they do not recommend the use of the OVA1 or ROMA tests for determining the status of an undiagnosed pelvic mass or for determining who should proceed to surgery.

Summary

The OVA1 and ROMA tests have been analytically validated and clinical performance has been reported in prospective multi-center clinical trials. Changes in the observed sensitivity and negative predictive value of testing have been small and of uncertain diagnostic value. Studies on the diagnostic accuracy of these tests compared to other diagnostic tools have had mixed findings, but do not report that ROMA is superior to other risk prediction tools that use standard clinical information. No studies have been performed that directly evaluate the impact of test results on referral patterns or health outcomes. Underlying these issues is some uncertainty regarding the benefit of initial treatment by a gynecologic oncologist beyond the need for reoperation in some cases. Prospective and randomized trials with long-term follow-up that provide direct evidence on the clinical utility of proteomic tests in ovarian cancer risk assessment are needed. Therefore, all uses of these tests, including use as a screening tool for ovarian cancer, are considered investigational.

REFERENCES

1. Van Holsbeke, C, Van Belle, V, Leone, FP, et al. Prospective external validation of the 'ovarian crescent sign' as a single ultrasound parameter to distinguish between benign and malignant adnexal pathology. *Ultrasound Obstet Gynecol*. 2010 Jul;36(1):81-7. PMID: 20217895
2. du Bois, A, Rochon, J, Pfisterer, J, Hoskins, WJ. Variations in institutional infrastructure, physician specialization and experience, and outcome in ovarian cancer: a systematic review. *Gynecol Oncol*. 2009 Feb;112(2):422-36. PMID: 18990435
3. Vernooij, F, Heintz, P, Witteveen, E, van der Graaf, Y. The outcomes of ovarian cancer treatment are better when provided by gynecologic oncologists and in specialized hospitals: a systematic review. *Gynecol Oncol*. 2007 Jun;105(3):801-12. PMID: 17433422
4. Giede, KC, Kieser, K, Dodge, J, Rosen, B. Who should operate on patients with ovarian cancer? An evidence-based review. *Gynecol Oncol*. 2005 Nov;99(2):447-61. PMID: 16126262
5. Tingulstad, S, Skjeldestad, FE, Hagen, B. The effect of centralization of primary surgery on survival in ovarian cancer patients. *Obstet Gynecol*. 2003 Sep;102(3):499-505. PMID: 12962932
6. Grabowski, JP, Harter, P, Buhrmann, C, et al. Re-operation outcome in patients referred to a gynecologic oncology center with presumed ovarian cancer FIGO I-IIIA after sub-standard initial surgery. *Surgical oncology*. 2012 Mar;21(1):31-5. PMID: 20875732
7. Medical Devices: Ovarian adnexal mass assessment score test system; Labeling; Black box restrictions. 21 CFR Part 866, Federal Register 2011;76(251):82128-82123.
8. TEC Assessment 2012. "Multi-analyte Testing for the Evaluation of Adnexal Masses." BlueCross BlueShield Association Technology Evaluation Center, Vol. 27, Tab TBA.
9. U.S. Food and Drug Administration. 510(k) Substantial Equivalence Determination Decision Summary: OVA1™ Test (K081754). [cited 03/2011]; Available from: http://www.accessdata.fda.gov/cdrh_docs/reviews/K081754.pdf
10. U.S. Food and Drug Administration. 510(k) Substantial Equivalence Determination Decision Summary: ROMA™ Test (K103358). [cited 10/2012]; Available from: http://www.accessdata.fda.gov/cdrh_docs/reviews/K103358.pdf
11. Bristow, RE, Smith, A, Zhang, Z, et al. Ovarian malignancy risk stratification of the adnexal mass using a multivariate index assay. *Gynecol Oncol*. 2013 Feb;128(2):252-9. PMID: 23178277
12. 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*. 2008 Feb;108(2):402-8. PMID: 18061248

13. 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*. 2011 Aug;118(2 Pt 1):280-8. PMID: 21775843
14. Van Gorp, T, Veldman, J, Van Calster, B, et al. Subjective assessment by ultrasound is superior to the risk of malignancy index (RMI) or the risk of ovarian malignancy algorithm (ROMA) in discriminating benign from malignant adnexal masses. *Eur J Cancer*. 2012 Jul;48(11):1649-56. PMID: 22226481
15. Partheen, K, Kristjansdottir, B, Sundfeldt, K. Evaluation of ovarian cancer biomarkers HE4 and CA-125 in women presenting with a suspicious cystic ovarian mass. *Journal of gynecologic oncology*. 2011 Dec;22(4):244-52. PMID: 22247801
16. Van Gorp, T, Cadron, I, Despierre, E, et al. HE4 and CA125 as a diagnostic test in ovarian cancer: prospective validation of the Risk of Ovarian Malignancy Algorithm. *Br J Cancer*. 2011 Mar 1;104(5):863-70. PMID: 21304524
17. Karlsen, MA, Sandhu, N, Hogdall, C, et al. Evaluation of HE4, CA125, risk of ovarian malignancy algorithm (ROMA) and risk of malignancy index (RMI) as diagnostic tools of epithelial ovarian cancer in patients with a pelvic mass. *Gynecol Oncol*. 2012 Nov;127(2):379-83. PMID: 22835718
18. Kaijser, J, Van Gorp, T, Van Hoorde, K, et al. A comparison between an ultrasound based prediction model (LR2) and the risk of ovarian malignancy algorithm (ROMA) to assess the risk of malignancy in women with an adnexal mass. *Gynecol Oncol*. 2013 May;129(2):377-83. PMID: 23360924
19. Society of Gynecologic Oncologists. Statement regarding OVA1. 2009. [cited 09/2012]; Available from: http://www.sgo.org/Media/Media_Archive/Media_Archive
20. Society of Gynecologic Oncologists. Multiplex Serum Testing for Women with Pelvic Mass (May 2013). [cited 12/03/2013]; Available from: https://www.sgo.org/newsroom/position-statements-2/multiplex-serum-testing-for-women-with-pelvic-mass/
21. Committee Opinion No. 477: the role of the obstetrician-gynecologist in the early detection of epithelial ovarian cancer. *Obstet Gynecol*. 2011 Mar;117(3):742-6. PMID: 21343791
22. National Cancer Institute. PDQ Genetics of Breast and Ovarian Cancer. 2012. [cited 09/2012]; Available from: http://www.cancer.gov/cancertopics/pdq/genetics/breast-and-ovarian/HealthProfessional/page4
23. National Comprehensive Cancer Network (NCCN). Clinical Practice Guidelines in Oncology™. Ovarian Cancer Including Fallopian Tube Cancer and Primary Peritoneal Cancer. v.2.2013 [cited 12/03/2013]; Available from: http://www.nccn.org/professionals/physician_gls/pdf/ovarian.pdf
24. BlueCross BlueShield Association Medical Policy Reference Manual "Proteomics-based testing for the Evaluation of Ovarian (Adnexal) Masses." Policy No. 2.04.62

CROSS REFERENCES

Analysis of Proteomic Patterns for Early Detection of Cancer, Laboratory, Policy No. 41

CODES	NUMBER	DESCRIPTION
CPT	81500	Oncology (ovarian), biochemical assays of two proteins (CA-125 and HE4), utilizing serum, with menopausal status, algorithm reported as a risk score (This code is for reporting the ROMA™ test)

CODES	NUMBER	DESCRIPTION
	81503	Oncology (ovarian), biochemical assays of five proteins (CA-125, apoliproprotein A1, beta-2 microglobulin, transferrin and pre-albumin), utilizing serum, algorithm reported as a risk score (This code is for reporting the OVA1™ test)
HCPCS	None	