Received: 29 July 2022 | Revised: 23 January 2023 | Accepted: 20 February 2023

DOI: 10.1002/ijgo.14733

CLINICAL ARTICLE

Gynecology

A reflex testing protocol using two multivariate index assays improves the risk assessment for ovarian cancer in patients with an adnexal mass

Herbert A. Fritsche[1] | Rowan G. Bullock[2]

[1]Aspira Labs, Austin, Texas, USA

[2]Aspira Women's Health, Inc., Trumbull, Connecticut, USA

Correspondence
Rowan G. Bullock, Aspira Women's Health, Inc., 35 Nutmeg Dr. Suite 260, Trumbull, CT 06611, USA.
Email: rbullock@aspirawh.com

Abstract

Objectives: Patients with adnexal masses suspicious for malignancy benefit from referral to oncology specialists during presurgical assessment of the mass. OVA1 is a multivariate assay using a five-biomarker panel which offers high overall and early-stage sensitivity. However, OVA1 has a high false-positive rate for benign masses. Overa, a second-generation multivariate index assay was developed to reduce the false-positive rate. The aim of the present study was to use Overa as a reflex for OVA1 and increase specificity.

Methods: OVA1 cut-off scores were established to place patients into three categories: low, intermediate, and high cancer risk. Samples with intermediate-risk OVA1 scores were reflexed to the Overa and defined as high or low risk. This protocol was tested with 1035 prospectively collected serum samples and validated with an independent prospectively collected sample set ($N = 207$).

Results: Thirty-five per cent (359) of samples had intermediate OVA1 scores. Reflexing these to Overa eliminated 58% of the false-positives and improved the overall specificity from 50% to 72%. This finding was confirmed in the independent dataset, in which the specificity increased from 56% to 73%.

Conclusions: Reflexing samples with intermediate OVA1 scores significantly decreases the false-positive rate, thereby reducing unnecessary surgical referrals.

KEYWORDS
adnexal mass, biomarkers, CA125, cancer risk assessment, OVA1, ovarian cancer, Overa, ROMA

1 | INTRODUCTION

OVA1® (Aspira Women's Health Inc.) is a serum multivariate index assay routinely used in the clinical work-up of patients with an adnexal mass. When an adnexal mass is determined to be at high risk for cancer, guidelines suggest that the patient be referred for specialty surgical care to achieve the best patient outcome.[1] The use of a non-invasive biomarker-based test such as OVA1 increases confidence in the decision to refer the patient for surgery, particularly given that the accuracy of malignancy detection by common imaging modalities is limited.[2,3]

The OVA1 data input consists of the concentration of five serum proteins: apolipoprotein A-1 (APO-A1), transthyretin (TT), beta

2-microglobulin (B2M), transferrin (TRF), and Cancer Antigen 125-II (CA125). These inputs are combined by a proprietary support vector machine learning algorithm to generate a unitless score ranging from 1.0 to 10.0. Elevated risk for ovarian cancer is defined as:

- premenopausal patients: OVA1 score ≥5.0;
- postmenopausal patients: OVA1 score ≥4.4.

As ovarian cancer is the most deadly gynecologic cancer, and one of the most difficult to detect in early stages, OVA1 was developed to favor high sensitivity in order to miss as few malignancies as possible. Its cancer detection rate is on the order of 90% with a positive predictive value in the range of 30%–40%.[4,5] OVA1 is not meant to be used as a standalone test, and its sensitivity in conjunction with clinical assessment is 96%.[5] Consequently, most cancers, including early-stage cancers, are readily detected by OVA1. However, the patient referral rate is higher than desired, at approximately 50%, including many benign cases. While there is no negative impact of these unnecessary patient referrals on the quality of care and patient outcome, these false positives can result in the ineffective use of medical resources, and undue anxiety and stress for the patient.[6]

Overa®, a second-generation multivariate index assay, was developed to improve the specificity of OVA1.[7] Overa was derived from the serum concentrations of APO-A1, TRF, CA125, Human Epididymis Protein 4 (HE4) and follicle-stimulating hormone (FSH). These inputs were processed by a support vector machine learning algorithm to combine the five inputs into a new risk score ranging from 1.0 to 10.0.[7–9] The sensitivity of Overa is equivalent to that of OVA1, but the specificity is greatly improved. While both OVA1 and Overa algorithms use five biomarker inputs, of which three (CA125, TRF, and APO-A1) are the same, Overa uses FSH and HE4 to replace the B2M and TT biomarkers that are used in the OVA1 algorithm. Both algorithms utilize a modified support vector machine (unified maximum separability algorithm [UMSA]-SVM) to define a linear solution for classification of cancer versus non-cancer. In the standard SVM model, there is a constant value which limits the maximum influence of any sample point on the final SVM solution. The UMSA modification to the standard SVM makes it so that constant values become individualized parameters for each data point by incorporating additional statistical information about the data point's position relative to the distribution of all the classes of samples in the training data. The rationale behind UMSA is that information about the overall data distribution can be used to qualify the "trustworthiness" of any of the training samples to become a support vector. The final UMSA-SVM solution, therefore, will rely on the weighted contributions of the support vectors and be less sensitive to labeling errors of a small percentage of the samples.

Here, we report the use of Overa as a reflex test to reduce the false-positive rate of OVA1. A two-step procedure was employed. First, OVA1 was used to determine the risk category as low, intermediate, or high. The low-risk category was defined as patients having OVA1 scores <4.4 (postmenopausal) and <5.0 (premenopausal). High-risk OVA1 scores >6.0 (postmenopausal) and >7.0 (for premenopausal) are used to identify patients with a high risk of malignancy. Samples with OVA1 scores falling within ranges of 4.4–6.0 for postmenopausal patients and 5.0–7.0 for premenopausal patients are classified as intermediate risk, and were reflexed to Overa. The process of categorizing patients into high-, intermediate-, and low-risk groups, followed by Overa testing of the intermediate risk group, is called OVA1plus.

TABLE 1 Patient characteristics of the combined data set (*N* = 1035).

	All evaluable subjects (*N* = 1035)		Premenopausal subjects (*N* = 545)		Postmenopausal subjects (*N* = 490)	
Age (years)						
N	1035		545		490	
Mean (SD)	49.7 (14.1)		39.9 (8.6)		60.6 (10.6)	
Median	48		42		60	
Range (min–max)	18–92		18–60		33–92	
Pathology diagnosis (*n* [%])						
Benign ovarian condition	804	77.7%	474	87.0%	330	67.4%
Epithelial primary ovarian malignancy	170	16.4%	46	8.4%	124	25.3%
Non-epithelial primary ovarian malignancy	15	1.5%	9	1.6%	6	1.2%
Low malignant potential (borderline) tumor	46	4.4%	16	2.9%	30	6.1%
Cancer stage for EOC and other primary (*N* [%])						
Stage I	66	35.7%	23	41.8%	43	33.1%
Stage II	25	13.5%	10	18.2%	15	11.5%
Stage III	82	44.3%	19	34.5%	63	48.5%
Stage IV	9	4.9%	2	3.6%	7	5.4%
Not given	3	1.6%	1	1.8%	2	1.5%

Abbreviation: EOC, endometrioid ovarian cancer.

TABLE 2 Performance of Ova1plus compared with OVA1, Overa, Cancer Antigen 125-II (CA125) and Risk of Malignancy Algorithm (ROMA) in the initial dataset ($N = 1035$).

Menopausal status	Sensitivity (95% CI) (*n*/*N*)	Early stage (I and II) sensitivity (95% CI) (*n*/*N*)	Specificity (95% CI) (*n*/*N*)	PPV (95% CI) (*n*/*N*)	NPV (95% CI) (*n*/*N*)
OVA1					
All	92.21%	91.21%	49.63%	34.47%	95.68%
	(88.75–95.66) 213/231	(83.41–96.13) 83/91	(46.68–52.57) 399/804	(30.72–38.21) 213/618	(93.73–97.63) 399/417
Pre	88.73%	84.85%	58.86%	24.42%	97.21%
	(81.38–96.09) 63/71	(68.10–94.89) 28/33	(55.09–62.63) 279/474	(19.18–29.66) 63/258	(95.31–99.12) 279/287
Post	93.75%	94.83%	36.36%	41.67%	92.31%
	(90.00–97.50) 150/160	(85.62–98.92) 55/58	(31.94–40.78) 120/330	(36.57–46.80) 150/360	(87.73–96.89) 120/130
Overa					
All	90.91%	89.01%	66.17%	43.57%	96.20%
	(87.20–94.62) 210/231	(80.71–94.60) 81/91	(62.90–69.44) 532/804	(39.14–47.99) 210/482	(94.61–97.80) 532/553
Pre	91.55%	87.88%	71.10%	32.18%	98.25%
	(85.08–98.02) 65/71	(71.80–96.60) 29/33	(67.02–75.18) 337/474	(25.74–38.62) 65/202	(96.86–99.64) 337/343
Post	90.63%	89.66%	59.09%	51.79%	92.86%
	(86.11–95.14) 145/160	(78.83–96.11) 52/58	(53.79–94.40) 195/330	(45.93–57.64) 145/280	(89.37–96.34) 195/210
OVA1+					
All	87.88%	86.81%	72.01%	47.43%	95.39%
	(83.67–92.09) 203/231	(78.10–93.00) 79/91	(68.91–75.12) 579/804	(68.91–75.12) 203/428	(93.72–97.06) 579/607
Pre	87.32%	81.82%	78.48%	37.80%	97.64%
	(79.59–95.06) 92/71	(64.54–93.02) 27/33	(74.78–82.18) 372/474	(30.38–45.23) 62/164	(96.11–99.16) 372/381
Post	88.13%	89.66%	62.73%	53.41%	91.59%
	(83.11–93.14) 141/160	(78.83–96.11) 52/58	(57.51–67.94) 207/330	(47.39–59.43) 141/264	(87.98–95.21) 207/226
CA125*					
All	79.22%	74.73%	79.85%	53.04%	93.04%
	(73.99–84.45) 183/231	(64.53–83.25) 68/91	(77.08–82.62) 642/804	(47.78–58.31) 183/345	(91.15–94.94) 642/690
Pre	74.65%	69.70%	79.54%	35.33%	95.44%
	(64.53–84.77) 53/71	(51.29–84.41) 23/33	(75.90–83.17) 377/474	(27.68–42.98) 53/150	(93.39–97.50) 377/395
Post	81.25%	77.59%	80.30%	66.67%	89.83%
	(75.20–87.30) 130/160	(64.73–87.49) 45/58	(76.01–84.59) 265/330	(60.05–73.28) 130/195	(86.38–93.28) 265/295
ROMA					
All	74.03%	67.03%	93.16%	75.66%	92.58%
	(68.37–79.68) 171/231	(59.39–76.53) 61/91	(91.41–94.90) 749/804	(70.07–81.26) 171/226	(90.78–94.39) 749/809
Pre	66.20%	54.55%	95.57%	69.12%	94.97%
	(55.19–77.20) 47/71	(36.35–71.89) 18/33	(93.72–97.42) 453/474	(58.14–80.10) 47/68	(93.01–96.93) 453/477

(Continues)

TABLE 2 (Continued)

Menopausal status	Sensitivity (95% CI) (*n*/*N*)	Early stage (I and II) sensitivity (95% CI) (*n*/*N*)	Specificity (95% CI) (*n*/*N*)	PPV (95% CI) (*n*/*N*)	NPV (95% CI) (*n*/*N*)
Post	77.50% (71.03–83.97) 124/160	74.14% (60.96–84.74) 43/58	89.70% (86.42–92.98) 296/330	78.48% (72.07–84.89) 124/160	89.16% (85.81–92.50) 296/332

Note: Cut-offs used: premenopausal, 67 U/mL; postmenopausal, 35 U/mL.

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

2 | MATERIALS AND METHODS

Testing of OVA1plus included sera from benign patients, and patients with malignancies primary to the ovary, excluding any cancers of extraovarian origin. This analysis utilized biomarker data obtained from 1035 eligible sera that were prospectively collected in two earlier institutional review board (IRB)-approved clinical trials (WIRB protocol numbers OVA1-001-CO1 and OVA2-001-CO3).[4,5] An independent IRB-approved prospectively collected sample set, FHCRC#7788, consisting of 207 eligible patients was also analyzed to corroborate the observed sensitivity and specificity of the test. This sample cohort was designed to have a high cancer prevalence for the specific use of validating the analytical performance of newly proposed ovarian cancer risk assessment tests. All patients in both sample sets had pelvic masses, were scheduled for surgical intervention, and provided informed consent to participate in research. All cancer diagnoses were confirmed by histology.

After determination of the OVA1 score, samples with OVA1 values in the intermediate risk range of 4.4–6.0 for premenopausal patients and in the range of 5.0–7.0 for postmenopausal patients were reflexed to Overa testing for final stratification of cancer risk. Pathology results obtained from surgical intervention was used to confirm all diagnoses. Sensitivity, specificity, and positive and negative predictive values were calculated for the entire cohort of 1035 patients.

Likelihood ratios and post-test probabilities were also calculated. The positive and negative likelihood ratios (LR+/−) were calculated from the performance of the reflex test on the total ($n = 1035$) subjects and post-test probabilities were calculated at assumed 5.0%, 7.5%, and 10.0% prevalence of cancer in patients with adnexal masses.[10]

The cut-off values used for CA125 were 67.0 U/mL for premenopausal patients, as suggested by Dearking et al.,[11] as the American College of Obstetricians and Gynecologists (ACOG) no longer recommends a specific numerical cut-off value for premenopausal patients.[12] A value of 35 U/mL was used for postmenopausal patients, as per ACOG recommendations.[12] Risk of Malignancy Algorithm (ROMA) (Fujirebio Diagnostics) values were calculated from the formula given by Moore et al.[13] All statistics were calculated using the DTComPair package of the R programming language (version 4.0.2).

3 | RESULTS

A clinical summary of the patients in the 1035-sample data set is shown in Table 1. The prevalence of cancer in this study group was 22% (231/1035) and included cancers of all epithelial cell types, non-epithelial cancers of the ovary, and borderline (low malignant potential) tumors of the ovary. Staging information is provided for the primary invasive tumors, but the borderline tumors in this analysis do not have staging due to a lack of quality data showing the necessity of surgical staging procedures for these low-risk tumors.[14]

The performance characteristics of OVA1plus are shown in Table 2 along with a comparison to the performance of OVA1, Overa, CA125 and ROMA. In the first step of testing with OVA1, 40% ($N = 417$) of the samples were stratified as low risk, 35% (359) were of intermediate risk, and 25% ($N = 259$) were high risk. Of the 359 samples that were classified in the OVA1 intermediate-risk group, Overa reflex testing reclassified 49.8% (180/359) as low risk. Overa reflex testing (OVA1plus) reduced the number of false positives in the OVA1 intermediate-risk group to a total of 179. The conversion of 180 patients from false positive to true negative improved the overall specificity of OVA1plus to 72%, a significant improvement over the 50% specificity of OVA1.

When assessing test performance by menopausal status, OVA1plus showed higher specificity in premenopausal patients (79%, 95% confidence interval [CI] 74.8–82.2) as compared with postmenopausal patients (63%, 95% CI 57.5–67.9), indicating a statistically significant difference in the sensitivity between menopausal groups, as has been shown for OVA1. For early-stage cancer detection, (International Federation of Gynecology and Obstetrics [FIGO] stages I and II, $n = 91$) the OVA1plus performance was statistically similar to the "all cancer" clinical sensitivity of 87%. Reflex testing with Overa eliminated 44.4% (180/405) of false-positive results reported by OVA1, albeit with a slight reduction of sensitivity compared with OVA1 (88% vs. 92%). OVA1plus misclassified 10 out of 213 cancers that were correctly identified by OVA1, resulting in a total of 28 false negatives, including the 18 that were missed by both tests (Table 2). Of these 28 false negatives, 50% ($n = 14$) were low-risk, borderline tumors; 32% ($n = 9$) were epithelial ovarian cancers, mixed across different subtypes (mucinous, serous, clear cell, endometrioid, and sarcoma); and 17.9% ($n = 5$) were non-epithelial ovarian cancers.

TABLE 3 Performance of OVA1plus compared with OVA1, Overa, Cancer Antigen 125-II (CA125) and Risk of Malignancy Algorithm (ROMA) in the independent dataset, FHCRC#7788.

Menopausal status	Sensitivity (95% CI) (*n*/*N*)	Early stage (I and II) sensitivity (95% CI) (*n*/*N*)	Specificity (95% CI) (*n*/*N*)	PPV (95% CI) (*n*/*N*)	NPV (95% CI) (*n*/*N*)
OVA1					
All	89.16% (82.47–95.85) 74/83	94.74% (73.97–99.87) 18/19	55.65% (46.90–64.39) 69/124	57.36% (48.83–65.90) 74/129	88.46% (81.37–95.55) 69/78
Pre	78.95% (60.62–97.28) 15/19	66.67% (9.43–99.16) 2/3	59.57% (45.54–73.60) 28/47	44.12% (27.43–60.81) 15/34	87.50% (76.04–98.96) 28/32
Post	92.19% (85.61–98.76) 59/64	100.00% (79.41–100.00) 16/16	53.25% (42.10–64.39) 41/77	62.11% (42.10–64.39) 59/95	89.13% (80.14–98.13) 41/46
Overa					
All	87.95% (80.95–94.95) 73/83	89.47% (66.86–98.70) 17/19	66.13% (57.80–74.46) 82/124	63.48% (54.68–72.28) 73/115	89.13% (82.77–95.49) 82/92
Pre	78.95% (60.62–97.28) 15/19	66.67% (9.43–99.16) 2/3	72.34% (59.55–85.13) 34/47	53.57% (35.10–72.04) 15/28	89.47% (79.72–99.23) 34/38
Post	90.63% (83.48–97.77) 58/64	93.75% (69.77–99.84) 15/16	62.34% (51.52–73.16) 48/77	66.67% (56.76–76.57) 58/87	88.89% (80.51–97.27) 48/54
OVA1plus					
All	84.34% (76.52–92.16) 70/83	94.74% (73.97–99.87) 18/19	73.39% (65.61–81.17) 91/124	67.96% (58.95–76.97) 70/103	87.50% (81.14–93.86) 91/104
Pre	73.68% (53.88–93.48) 14/19	66.67% (9.43–99.16) 2/3	74.47% (62.00–86.93) 35/47	53.85% (34.68–73.01) 14/26	87.50% (77.25–97.75) 35/40
Post	87.50% (79.40–95.60) 56/64	100.00% (79.41–100.00) 16/16	72.27% (62.78–82.67) 56/77	72.27% (62.78–82.67) 56/77	87.50% (79.40–95.60) 56/64
CA125*					
All	74.70% (65.35–84.05) 62/83	78.95% (54.43–93.95) 15/19	80.65% (73.69–87.60) 100/124	72.09% (62.61–81.57) 62/86	82.64% (75.90–89.39) 100/121
Pre	63.16% (41.47–84.85) 12/19	33.33% (0.84–90.57) 1/3	76.60% (64.49–88.70) 36/47	52.17% (31.76–72.59) 12/23	83.72% (72.69–94.76) 36/43
Post	78.13% (68.00–88.25) 50/64	87.50% (61.65–98.45) 14/16	83.12% (74.75–91.48) 64/77	79.37% (69.37–89.36) 50/63	82.05% (73.53–90.57) 64/78
ROMA					
All	67.47% (57.39–77.55) 56/83	78.95% (54.43–93.95) 15/19	89.52% (84.12–94.91) 111/124	81.16% (71.93–90.39) 56/69	80.43% (73.82–87.05) 111/138
Pre	47.37% (24.92–69.82) 9/19	66.67% (9.43–99.16) 2/3	93.62% (86.63–100.00) 44/47	75.00% (50.50–99.50) 9/12	81.48% (71.12–91.84) 44/54

(Continues)

TABLE 3 (Continued)

Menopausal status	Sensitivity (95% CI) (*n*/*N*)	Early stage (I and II) sensitivity (95% CI) (*n*/*N*)	Specificity (95% CI) (*n*/*N*)	PPV (95% CI) (*n*/*N*)	NPV (95% CI) (*n*/*N*)
Post	73.44%	81.25%	87.01%	82.46%	79.76%
	(62.62–84.26) 47/64	(54.35–95.95) 13/16	(79.50–94.52) 67/77	(72.58–92.33) 47/57	(71.17–88.35) 67/84

Note: Cut-offs used: premenopausal, 67 U/mL; postmenopausal, 35 U/mL.

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

TABLE 4 Analysis of likelihood ratios for OVA1plus.

Pre-test probability (assumed prevalence)	Menopausal status	*N*	Results			
			LR+	Post-test probability of malignancy given a high-risk OVA1plus result	LR−	Post-test probability of a benign mass given a low-risk OVA1plus result
5.0%	All	1035	3.14	14.18%	0.17	99.11%
	Pre	545	4.06	17.61%	0.16	99.16%
	Post	490	2.36	11.05%	0.19	99.01%
7.5%	All	1035	3.14	20.29%	0.17	98.64%
	Pre	545	4.06	24.77%	0.16	98.72%
	Post	490	2.36	16.06%	0.19	98.48%
10.0%	All	1035	3.14	25.86%	0.17	98.15%
	Pre	545	4.06	31.09%	0.16	98.25%
	Post	490	2.36	20.77%	0.19	97.93%

In the 1035-sample set, Overa gave comparable sensitivity to OVA1, but demonstrated improved specificity over OVA1 (66% vs. 50%). However, the best specificity resulted from the OVA1plus reflex testing protocol (72%). Additionally, as shown in Table 2, both CA125 and ROMA tests showed unacceptably lower cancer detection rates of 79% and 74%, respectively, but did show higher specificity than OVA1plus.

A second independent sample set (FHCRC#7788) of 207 patients with a cancer prevalence of 40% also showed significant reduction in false positives when intermediate risk samples were reflexed to Overa (i.e. OVA1plus; Table 3).

The specificity was improved from 55.6% to 73.4%. The number of cancers that were correctly identified by OVA1 but misclassified by OVA1plus in this high-cancer-prevalence dataset was only 4 out of 74. Due to the exceptionally high prevalence of malignancy in this dataset, the positive predictive value (PPV) is high (68%), but the negative predictive value (NPV) was only slightly affected (88%). Lastly, in this dataset, CA125 and ROMA show low cancer detection rates (75% and 68%, respectively), but with high specificity.

The impact of reflex testing on likelihood ratios (LRs) and post-test probabilities are displayed in Table 4. This analysis was performed to demonstrate a more realistic utility of this test, as the datasets used in this analysis were taken from enriched populations with an approximately 20% malignancy prevalence, and, as such, the positive and negative predictive values demonstrated are not likely to reflect those observed in a general patient population, wherein the prevalence of cancer is more likely to be in the range of 5–10%.

The calculated LR+ values all exceeded at least twice the odds as predicted from prevalence. The LR− values demonstrate the improved specificity of OVA1plus with the overall post-test probabilities of benign disease being over 98% at all prevalences, given a negative risk result.

4 | DISCUSSION

There is a large body of literature showing that patients with a high risk of ovarian malignancies benefit from referral to a gynecologic oncologist for specialty surgical care, where they are more likely to receive appropriate staging and more complete debulking and cytoreduction, resulting in improved survival.[15] However, there are challenges associated with determining which patients should be referred and when. Gynecologic oncology is a specialized field with a small number of practitioners in the USA and even fewer outside of large metropolitan centers.[16] Access to these specialists may be limited, and travel time and cost are factors, as are the anxiety and psychological toll associated with referral to a cancer center for treatment. Therefore, an ideal risk assessment tool for ovarian cancer would combine a high sensitivity with a high specificity in order to detect most cancers, and minimize false positives that result in inappropriate referrals for patients who actually have benign masses.

A test with higher specificity helps to determine which patients may be safely retained by the general gynecologist for surgical intervention. Improved risk assessment reduces the burden on specialists, avoids the wasteful use of medical resources, and ultimately benefits the patient.

While appropriate management of low risk masses is important, ovarian cancer is the deadliest gynecologic cancer and the fifth leading cause of death in women.[17] Its symptoms are non-specific and early detection is particularly challenging, as many early-stage ovarian malignancies may be entirely asymptomatic.[18] As such, approximately 75% of patients are not diagnosed until the disease has progressed to an advanced stage, at which point the prognosis is poor, with a 5-year survival rate as low as 17%.[18] The potential consequences of missing a malignant mass are severe. Therefore, an ideal risk assessment tool for ovarian cancer cannot neglect sensitivity either. A balance between sensitivity and specificity is needed for cancer risk assessment.

There is currently no recommended screening or diagnostic test for ovarian cancer. Several options exist for risk assessment tools as part of a clinical work-up. Imaging, typically transvaginal ultrasound, and CA125 are commonly used by practitioners in the US. There are limitations for both of these modalities. The accuracy of imaging is limited and may depend on the experience and expertise of the ultrasonographer.[2,3] There are systems in use to improve the accuracy of ultrasound alone, such as IOTA simple rules, which are included in ACOG guidelines, which predict malignancy based on a number of imaging features. Their accuracy is high, but one drawback is that they classify approximately one-third of tumors as indeterminate, which can be problematic for clinicians.[19,20] CA125 is excellent for detecting certain histological subtypes of epithelial ovarian malignancies, particularly at advanced stages, but some subtypes and early-stage cancers do not cause elevated CA125.[21] Combining CA125 with HE4 in ROMA shows mixed results in improving detection over CA125 alone.[22–24] Our results in Tables 2 and 3 corroborate the low sensitivities of CA125 and ROMA for cancer detection.

OVA1plus, the reflex protocol that utilizes the OVA1 and Overa algorithms, produces a risk assessment score that gives high sensitivity and improved specificity compared with either OVA1 or Overa used alone. The non-invasive serum biomarker reflex test has high sensitivity for early-stage cancer detection and its improved specificity over OVA1 used alone leads to fewer false positives, resulting in fewer unnecessary referrals, which is beneficial for patients, practitioners, and the healthcare system at large. The OVA1plus risk result is a binary classification, eliminating the confusion caused by indeterminate results. In our data set, 35% of the OVA1 false positives were determined to be indeterminate. When Overa is employed as an automatic reflex test, the indeterminate category is eliminated.

OVA1plus has the same intended use as OVA1, in that it is a risk assessment tool used as part of the clinical work-up of a confirmed adnexal mass for which surgery is planned, in order to help guide referral decisions. The use of Overa as a reflex tool for samples that fall within the OVA1 within intermediate range significantly improves the accuracy of cancer risk assessment. As such, OVA1plus is a valuable tool when utilized as part of the clinical work-up of an adnexal mass, helping to guide the decision to refer a patient to specialty care for surgical intervention, or to determine that the patient can be safely treated without referral. A limitation of this report is the retrospective nature of the data. Studies are under way to investigate the clinical performance and utility of OVA1plus in a population that reflects a 'real-world' demographic of patients.

AUTHOR CONTRIBUTIONS

HAF and RGB contributed significantly to drafting, editing, formal analysis, and final approval of the manuscript.

CONFLICT OF INTEREST STATEMENT

Both authors were employed or contracted by Aspira Women's Health Inc. or its subsidiary, Aspira Labs, at the time of contribution. Aspira Women's Health, Inc. provided funding for this study. No grant funding was used in the referenced trials or the preparation of this manuscript.

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

REFERENCES

1. Carney ME, Lancaster JM, Ford C, Tsodikov A, Wiggins CL. A population-based study of patterns of care for ovarian cancer: who is seen by a gynecologic oncologist and who is not? *Gynecol Oncol*. 2002;84:36-42.
2. Iyer VR, Lee SI. MRI, CT, and PET/CT for ovarian cancer detection and adnexal lesion characterization. *AJR Am J Roentgenol*. 2010;194(2):311-321. doi:10.2214/AJR.09.3522
3. Valentin L, Ameye L, Jurkovic D, et al. Which extrauterine pelvic masses are difficult to correctly classify as benign or malignant on the basis of ultrasound findings and is there a way of making a correct diagnosis? *Ultrasound Obstet Gynecol*. 2006;27(4):438-444. doi:10.1002/uog.2707
4. Ueland FR, Desimone CP, Seamon LG, et al. Effectiveness of a multivariate index assay in the preoperative assessment of ovarian tumors. *Obstet Gynecol*. 2011;117:1289-1297.
5. Bristow RE, Smith A, Zhang Z, et al. Ovarian malignancy risk stratification of the adnexal mass using a multivariate index assay. *Gynecol Oncol*. 2013;128:252-259.
6. Khimani F, Curley B, Almubarak M. Survey of patients referred to a university cancer Center for Benign Hematology: quality measures and patient understanding. *J Oncol Pract*. 2015;11(1):26-29. doi:10.1200/JOP.2014.001543
7. Coleman RL, Herzog TJ, Chan DW, et al. Validation of a second-generation multivariate index assay for malignancy risk of adnexal masses. *Am J Obstet Gynecol*. 2016;215:e1-e11.
8. Zhang Z, Bast RC, Yu Y, et al. Three biomarkers identified from serum proteomic analysis for the detection of early stage ovarian cancer. *Cancer Res*. 2004;64:5882-5890.
9. Zhang Z, Chan DW. The road from discovery to clinical diagnostics: lessons learned from the first FDA-cleared in vitro diagnostic multivariate index assay of proteomic biomarkers. *Cancer Epidemiol Biomark Prev*. 2010;19:2995-2999.

10. National Institutes of Health Consensus Development Conference Statement. Ovarian cancer: screening, treatment, and follow-up. *Gynecol Oncol*. 1994;55:S4-S14.
11. Dearking AC, Aletti GD, McGree ME, Weaver AL, Sommerfield MK, Cliby WA. How relevant are ACOG and SGO guidelines for referral of adnexal mass? *Obstet Gynecol*. 2007;110(4):841-848. doi:10.1097/01.AOG.0000267198.25223.bc
12. American College of Obstetricians and Gynecologists' Committee on Practice Bulletins—Gynecology. Practice bulletin No. 174: evaluation and management of adnexal masses. *Obstet Gynecol*. 2016;128(5):e210-e226. doi:10.1097/AOG.0000000000001768
13. Moore RG, McMeekin DS, Brown AK, et al. A novel multiple marker bioassay utilizing HE4 and CA125 for the prediction of ovarian cancer in patients with a pelvic mass. *Gynecol Oncol*. 2009;112:40-46.
14. Shazly SA, Laughlin-Tommaso SK, Dowdy SC, Famuyide AO. Staging for low malignant potential ovarian tumors: a global perspective. *Am J Obstet Gynecol*. 2016;215(2):153-168.e2. doi:10.1016/j.ajog.2016.04.035
15. Chan JK, Kapp DS, Shin JY, et al. Influence of the gynecologic oncologist on the survival of ovarian cancer patients. *Obstet Gynecol*. 2007;109:1342-1350.
16. Ricci S, Tergasc AI, Long K, et al. Geographic disparities in the distribution of the U.S. gynecologic oncology workforce: a Society of Gynecologic Oncology study gynecologic. *Oncol Rep*. 2017;22:100-104.
17. Torre LA, Trabert B, DeSantis CE, et al. Ovarian cancer statistics, 2018. *CA Cancer J Clin*. 2018;68(4):284-296. doi:10.3322/caac.21456
18. Doubeni CA, Doubeni AR, Myers AE. Diagnosis and management of ovarian cancer. *Am Fam Physician*. 2016;93(11):937-944.
19. Kaijser J, Sayasneh A, Van Hoorde K, et al. Presurgical diagnosis of pelvic tumours using mathematical models and scoring systems: a systematic review and meta-analysis. *Hum Reprod Update*. 2014;20(3):449-462. doi:10.1093/humupd/dmt059
20. Patel-Lippmann KK, Sadowski EA, Robbins JB, et al. Comparison of international ovarian tumor analysis simple rules to Society of Radiologists in ultrasound guidelines for detection of malignancy in pelvic cysts. *Am J Roentgenol*. 2020;214(3):694-700. doi:10.2214/ajr.18.20630
21. Choi JH, Sohn GS, Chay DB, Cho HB, Kim JH. Preoperative serum levels of cancer antigen 125 and carcinoembryonic antigen ratio can improve differentiation between mucinous ovarian carcinoma and other epithelial ovarian carcinomas. *Obstet Gynecol Sci*. 2018;61(3):344-351. doi:10.5468/ogs.2018.61.3.344
22. Nolen BM, Lokshin AE. Biomarker testing for ovarian cancer: clinical utility of multiplex assays. *Mol Diagn Ther*. 2013;17(3):139-146. doi:10.1007/s40291-013-0027-6
23. Montagnana M, Danese E, Ruzzenente O, et al. The ROMA (risk of ovarian malignancy algorithm) for estimating the risk of epithelial ovarian cancer in women presenting with pelvic mass: is it really useful? *Clin Chem Lab Med*. 2011;49(3):521-525. doi:10.1515/CCLM.2011.075
24. Zheng H, Tie Y, Wang X, Yang Y, Wei X, Zhao X. Assessment of the diagnostic value of using serum CA125 and GI-RADS system in the evaluation of adnexal masses. *Medicine*. 2019;98:1-6.

How to cite this article: Fritsche HA, Bullock RG. A reflex testing protocol using two multivariate index assays improves the risk assessment for ovarian cancer in patients with an adnexal mass. *Int J Gynecol Obstet*. 2023;00:1-8. doi:10.1002/ijgo.14733