

Ovarian malignancy risk stratification of the adnexal mass using a multivariate index assay[☆]

Robert E. Bristow^{a,*}, Alan Smith^b, Zhen Zhang^c, Daniel W. Chan^c, Gillian Crutcher^d, Eric T. Fung^d, Donald G. Munroe^e

^a University of California, Irvine-Medical Center, Orange, CA, USA

^b Applied Clinical Intelligence, Bala Cynwyd, PA, USA

^c The Johns Hopkins Medical Institutions, Baltimore, MD, USA

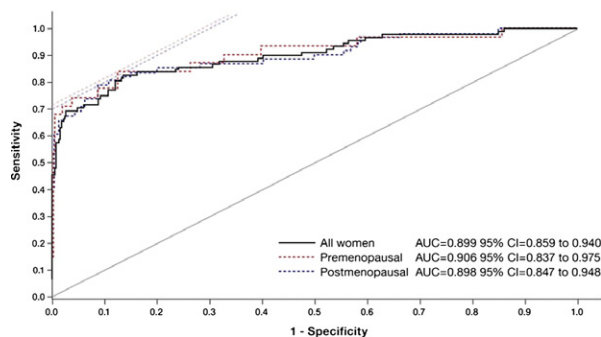
^d Vermillion, Inc., Mountain View, CA, USA

^e Vermillion, Inc., Austin, TX, USA

HIGHLIGHTS

- ▶ The multivariate index assay (MVA) was evaluated in an intended-use population of non-gynecologic oncology practices.
- ▶ The MVA demonstrated high sensitivity and NPV for ovarian malignancy.
- ▶ The MVA correctly identified 83.3% malignancies missed by clinical impression and 70.8% cases missed by CA125-II.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:

Received 25 September 2012

Accepted 15 November 2012

Available online 21 November 2012

Keywords:

Ovarian malignancy

Risk stratification

ABSTRACT

Objective. To validate the effectiveness of a multivariate index assay in identifying ovarian malignancy compared to clinical assessment and CA125-II, among women undergoing surgery for an adnexal mass after enrollment by non-gynecologic oncology providers.

Methods. A prospective, multi-institutional trial enrolled female patients scheduled to undergo surgery for an adnexal mass from 27 non-gynecologic oncology practices. Pre-operative serum samples and physician assessment of ovarian cancer risk were correlated with final surgical pathology.

Results. A total of 494 subjects were evaluable for multivariate index assay, CA125-II, and clinical impression. Overall, 92 patients (18.6%) had a pelvic malignancy. Primary ovarian cancer was diagnosed in 65 patients (13.2%), with 43.1% having FIGO stage I disease. For all ovarian malignancies, the sensitivity of the multivariate index assay was 95.7% (95%CI=89.3–98.3) when combined with clinical impression. The multivariate index assay correctly predicted ovarian malignancy in 91.4% (95%CI=77.6–97.0) of cases of early-stage disease, compared to 65.7% (95%CI=49.2–79.2) for CA125-II. The multivariate index assay correctly identified 83.3% malignancies missed by clinical impression and 70.8% cases missed by CA125-II. Multivariate index assay was superior in predicting the absence of an ovarian malignancy, with a negative predictive value of 98.1% (95%CI=95.2–99.2). Both clinical impression and CA125-II were more accurate at

[☆] This study was funded by Vermillion Inc., Austin, TX.

* Corresponding author at: Division of Gynecologic Oncology, Department of Obstetrics and Gynecology, University of California, Irvine School of Medicine, 101 The City Drive, Building 56, Room 260, Orange, CA 92868, USA. Fax: +1 714 4546 7754.

E-mail address: rbristow@uci.edu (R.E. Bristow).

identifying benign disease. The multivariate index assay correctly predicted benign pathology in 204 patients (50.7%, 95%CI=45.9–55.6) when combined with clinical impression.

Conclusion. The multivariate index assay demonstrated higher sensitivity and negative predictive value for ovarian malignancy compared to clinical impression and CA125-II in an intended-use population of non-gynecologic oncology practices.

© 2012 Elsevier Inc. All rights reserved.

Background

The American Cancer Society has estimated that 22,280 new cases of ovarian cancer will be diagnosed in the United States in 2012 [1]. With an estimated 15,500 women dying from disease in 2012, ovarian cancer accounts for more deaths than all other gynecologic cancers combined [1]. Despite clinical data supporting referral of all women with suspected ovarian cancer to a gynecologic oncologist, the proportion of ovarian cancer patients operated on by a gynecologic oncologist is less than 50% [2–10]. A major impediment to appropriate referral patterns is the challenge of identifying which subgroup of women with an adnexal mass is most likely to have ovarian cancer [10,11]. Particularly challenging subgroups are those with early-stage disease, where 50% of patients will have a normal CA125, and premenopausal women, where the prevalence of ovarian cancer is low.

Although numerous prediction models and referral guidelines have been tested in the preoperative evaluation of the adnexal mass, no single method has received widespread acceptance [12–14]. Recently, novel biomarker testing decision algorithms have been developed to aid in the preoperative evaluation process. These triage tools are not screening tests, which are designed to detect disease in asymptomatic patients. The multivariate index assay (OVA1®) is a multiple biomarker test that was cleared for use in clinical practice by the FDA in 2009 based on a high sensitivity and negative predictive value for identifying ovarian malignancy [15]. However, the pivotal trial reported by Ueland et al. in 2011 contained a mix of subjects that were enrolled by both gynecologists and gynecologic oncologists, with an associated prevalence of ovarian malignancy of 29%. Since the intended use of OVA1 is as a diagnostic triage aid in determining the need for gynecologic oncology referral, the objective of the current study was to validate the effectiveness of OVA1 in identifying ovarian malignancy, compared to clinical assessment and serum CA125-II, in a unique and prospectively acquired cohort of women undergoing surgery for an adnexal mass after enrollment by non-gynecologic oncology providers, independent of the original dataset reported by Ueland et al.

Methods

Consecutive patients who met inclusion criteria were prospectively enrolled at 27 sites throughout the United States, with Institutional Review Board approval from each site. All clinicians initially enrolling patients were from non-gynecologic oncology specialty practices, although patients may ultimately have had consultation with or undergone surgery by a gynecologic oncologist. Inclusion criteria were: females age ≥ 18 years, signed informed consent and agreeable to phlebotomy, documented pelvic mass planned for surgical intervention within 3 months of imaging. A pelvic mass was confirmed by imaging (computed tomography, ultrasonography, or magnetic resonance imaging) prior to enrollment. Exclusion criteria included a diagnosis of malignancy in the previous 5 years (except of non-melanoma skin cancers) or enrollment by a gynecologic oncologist. Menopause was defined as the absence of menses for ≥ 12 months, or age ≥ 50 . Demographic and clinic-pathologic information were collected on case report forms.

A preoperative blood sample (≤ 80 ml) was processed and serum frozen at the collection site. Biomarker measurements were performed according to the OVA1 Instructions for Use at the Division of Clinical

Chemistry, Department of Pathology, Johns Hopkins Medical Institutions. OVA1 is a multivariate biomarker assay that incorporates CA125-II, transferrin, transthyretin (prealbumin), apolipoprotein A1, and beta-2-microglobulin. The individual biomarker results were then transformed by the OvaCalc™ software using a proprietary multivariate algorithm, to generate an ovarian malignancy risk score as described previously [15]. The numeric result ranged from 0.0 to 10.0, and patient risk of malignancy was stratified using a cut-off specific to menopausal status:

- Premenopausal
 - Low probability of malignancy (OVA1 < 5.0).
 - High probability of malignancy (OVA1 ≥ 5.0).
- Postmenopausal
 - Low probability of malignancy (OVA1 < 4.4).
 - High probability of malignancy (OVA1 ≥ 4.4).

For CA 125-II measurement, the same biomarker value entered into the OvaCalc software was used for individual analysis, and compared with clinical cutoff values in accordance with published American College of Obstetricians and Gynecologists (ACOG) referral criteria ≥ 200 units/ml for premenopausal women or more than 35 units/ml for postmenopausal women [13]. The effect of substituting the modified ACOG criteria for premenopausal women ≥ 67 units/ml was also evaluated [16].

Clinicians were required to document the results of physical examination, family history, imaging, laboratory tests (including CA125, when available, but not OVA1), and formal pre-surgical prediction of malignancy. In cases where the formal prediction was done by a clinician other than the enrolling physician, the referral history and the specialty of the clinician who made the prediction were recorded, as was the specialty of the surgeon who ultimately operated on each patient. In order to reflect their routine clinical judgment and referral behavior, physicians were not asked to either follow any specific prediction algorithm or justify their prediction. Postoperative pathology diagnosis was recorded at each enrolling site and independently reviewed.

Case report forms, biomarker values and OVA1 scores were sent to Applied Clinical Intelligence for statistical analysis. Results were statistically stratified based on menopausal status, stage of malignancy, and surgical pathology. Clinical diagnostic performance criteria (sensitivity, specificity, positive predictive value, negative predictive value and likelihood ratios) were calculated for OVA1, CA125, and clinical impression alone or in combination with OVA1. The OVA1 Instructions for Use indicate that referral should use a Boolean “OR” between physician assessment and OVA1 risk stratification, so that either criterion will trigger patient evaluation by a gynecologic oncologist. Accordingly, the performance of combined OVA1 use with physician assessment was simulated by an “OR” function and the resulting combined diagnostic predictions were used for concordance analyses and ROC curve analysis. For example, if either clinical impression or OVA1 predicted malignancy, the result for OVA1 combined with clinical impression was a prediction of high probability of malignancy. 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. Concordances between OVA1

results of high or low probability of malignancy and CA125 sensitivity by pathological diagnosis, stage of disease, and menopausal status were assessed using McNemar's test. Ninety-five percent confidence intervals were constructed where appropriate, and those for the area under the receiver operating characteristic curve were calculated using a bootstrap procedure. Statistical analysis was performed with SAS 9.2 (SAS Institute Inc).

Results

Demographic characteristics

From August, 2010 to December, 2011 a total of 520 subjects were enrolled, of which 494 were evaluable for OVA1 and CA125-II, and clinical impression. Subjects were excluded from the final analysis for the following reasons: failed exclusion criteria (imaging outside of window prior to inclusion, surgery beyond 12 weeks, previous cancer <5 years, $n = 12$), primary contact gynecologic oncologist contact ($n = 6$), and no ovarian pathology ($n = 8$). The clinical characteristics of the subject population are shown in Table 1. All 494 subjects had a non-gynecologic oncologist as their primary contact (obstetrician/gynecologists, $n = 339$; family practice physicians, nurse practitioners, emergency department physicians, or medical oncologists, $n = 155$). The specialty of physicians making the clinical assessment was a non-gynecologic oncologist in 253 patients and a gynecologic oncologist in 241 patients. Overall, 402 patients (81.4%) had benign pathology, 92 patients (18.6%) had a pelvic malignancy. There were 65 primary ovarian cancers, with 53.8% of these patients having Stage I/II disease.

OVA1 performance

Receiver operating characteristic (ROC) curves for OVA1 combined with clinical impression are shown in Fig. 1A. The area under the curve (AUC) for all patients was 0.899 (95%CI = 0.859–0.940). ROC curves stratified by menopausal status show similar discriminative ability according to the specified cutoff values. The correlative clinical performance in the study population according to the pre- and post-menopausal cutoff values for OVA1 combined with clinical impression are shown in Figs. 1B and C.

A comparison of performance in predicting ovarian malignancy for clinical impression, OVA1, OVA1 plus clinical impression, and CA125-II is shown in Table 2. Clinical impression and the two pre-menopausal CA125 cutoff values had comparatively low sensitivity (<80%). In contrast, OVA1 demonstrated a sensitivity of 95.7% when combined with clinical impression. OVA1 demonstrated superior performance in predicting the absence of an ovarian malignancy, with a negative predictive value (NPV) of 98.1% when combined with clinical impression and 96.8% as an independent risk stratification tool. OVA1 correctly identified 83.3% (20/24) malignancies missed by clinical impression and 70.8% (17/24) cases missed by CA125-II. Both clinical impression and CA125-II were more accurate than OVA1 in correctly identifying benign disease. The individual OVA1® score distributions stratified by menopausal status and pathological diagnosis are shown in Fig. 2.

The comparative sensitivity for OVA1 and CA125-II for detecting malignancy, stratified according to histology, FIGO stage of disease, and menopausal status, is shown in Table 3. As a risk-stratification test, OVA1 had a higher sensitivity for detecting ovarian cancer than CA125-II at both pre-menopausal high-risk cutoff values. Using the specified OVA1 threshold values, the overall sensitivity of a positive OVA1 test result was 95.0% for epithelial ovarian cancer, 82.4% for ovarian tumors of low malignant potential, and 80.0% for non-epithelial ovarian cancer. This relationship was also observed across FIGO stage of disease and menopausal status. Notably, among the 277 pre-menopausal patients, CA125-II correctly identified just 45.5% (>200U/ml) and 72.7% (>67U/ml) of Stage I/II ovarian

Table 1
Demographics of all evaluable subjects.

	All evaluable subjects (N = 494)	Premenopausal women (N = 277)	Postmenopausal women (N = 217)
Age, years			
N	494	277	217
Mean (SD)	48.6 (14.15)	39.6 (8.95)	60.2 (10.74)
Median	48	41	60
Range	18 to 87	18 to 60	33 to 87
(min to max)			
Ethnicity/race, n (%)			
White	348 (70.4)	174 (62.8)	174 (80.2)
African-American	81 (16.4)	54 (19.5)	27 (12.4)
Hispanic or Latino	46 (9.3)	36 (13.0)	10 (4.6)
Asian	13 (2.6)	8 (2.9)	5 (2.3)
Native Hawaiian/ Pacific islander	1 (0.2)	1 (0.4)	0 (0.0)
Other	5 (1.0)	4 (1.4)	1 (0.5)
No. of pregnancies, n (%)			
None	80 (16.2)	56 (20.2)	24 (11.1)
1	87 (17.6)	53 (19.1)	34 (15.7)
2	131 (26.5)	70 (25.3)	61 (28.1)
3	94 (19.0)	50 (18.1)	44 (20.3)
4 or more	102 (20.6)	48 (17.3)	54 (24.9)
Physician's assessment, n (%)			
Malignant	98 (19.8)	39 (14.1)	59 (27.2)
Benign	396 (80.2)	238 (85.9)	158 (72.8)
Pathology diagnosis, n (%)			
Benign	402 (81.4)	246 (88.8)	156 (71.9)
Non-ovarian	4 (0.8)	1 (0.4)	3 (1.4)
primary malignancy with no involvement of the ovaries			
Non-ovarian	6 (1.2)	2 (0.7)	4 (1.8)
primary malignancy with involvement of the ovaries			
Ovarian low malignant potential (Borderline)	17 (3.4)	5 (1.8)	12 (5.5)
Primary malignant ovarian tumor	65 (13.2)	23 (8.3)	42 (19.4)
Epithelial ovarian cancer	60 (12.1)	18 (6.5)	42 (19.4)
Serous	24 (4.9)	8 (2.9)	16 (7.4)
Mucinous	9 (1.8)	1 (0.4)	8 (3.7)
Endometrioid	13 (2.6)	5 (1.8)	8 (3.7)
Clear cell	5 (1.0)	1 (0.4)	4 (1.8)
Carcinosarcoma	1 (0.2)	1 (0.4)	0 (0.0)
Mixed	1 (0.2)	0 (0.0)	1 (0.5)
Other	7 (1.4)	2 (0.7)	5 (2.3)
Non-Epithelial:	5 (1.0)	5 (1.8)	0 (0.0)
Other			
Tumor stage, n (% of all primary malignant ovarian tumor)			
Stage 1	28 (43.1)	9 (39.1)	19 (45.2)
Stage 2	7 (10.8)	2 (8.7)	5 (11.9)
Stage 3	25 (38.5)	10 (43.5)	15 (35.7)
Stage 4	5 (7.7)	2 (8.7)	3 (7.1)

malignancies. In contrast, the sensitivity of OVA1 for detecting ovarian malignancy in this same diagnostically challenging patient subset was 90.9%.

Discussion

The Society of Surgical Oncology provides guidelines for ovarian cancer surgery: "Surgeons undertaking operations for possible ovarian cancer should have both the necessary technical expertise and a thorough understanding of the management of the disease itself... and is best carried out in centers in which an experienced and coordinated multidisciplinary team is available" [17]. It has been extensively documented that initial surgery by a gynecologic oncologist is associated

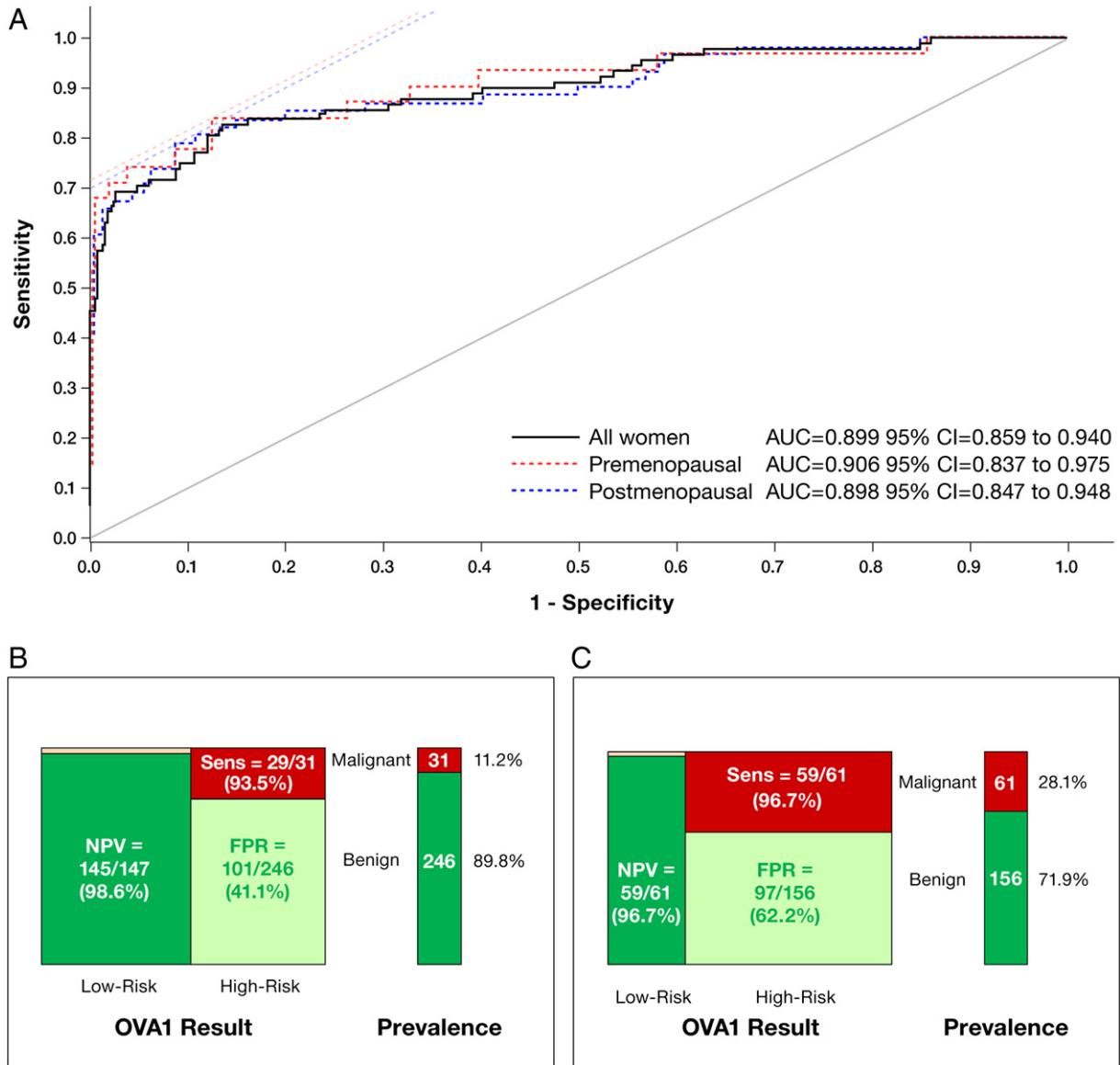


Fig. 1. OVA1 + clinical impression ROC curve analysis and matrix plot. A: ROC curve for OVA1 combined with clinical impression. B: Matrix plot OVA1 + clinical impression: pre-menopausal ($n = 277$); SENS = sensitivity, NPV = negative predictive value, FPR = false positive rate. C: Matrix plot OVA1 + clinical impression: post-menopausal ($n = 217$); SENS = sensitivity, NPV = negative predictive value, FPR = false positive rate.

Table 2
Test performance in predicting ovarian malignancy.

	Clinical assessment ($n = 494$)	OVA1 ($n = 494$)	Clinical Assessment + OVA1 (OR function) ($n = 494$)	CA125-II ^a ($n = 494$)	CA125-II ^b ($n = 494$)
Sensitivity	73.9	92.4	95.7	73.9	79.3
n/N	68/92	85/92	88/92	68/92	73/92
95% CI	64.1–81.8	85.1–96.3	89.3–98.3	64.1–81.8	70.0–86.4
Specificity	92.5	53.5	50.7	94.5	86.1
n/N	372/402	215/402	204/402	380/402	346/402
95% CI	89.5–94.7	48.6–58.3	45.9–55.6	91.9–96.4	82.3–89.1
PPV	69.4	31.3	30.8	75.6	56.6
n/N	68/98	85/272	88/286	68/90	73/129
95% CI	59.7–77.6	26.0–37.0	26.2–36.8	65.8–83.3	48.0–64.8
NPV	93.9	96.8	98.1	94.1	94.8
n/N	372/396	215/222	204/208	380/404	346/365
95% CI	91.1–95.9	93.6–98.5	95.2–99.2	91.3–96.0	92.0–96.6
Pre-test odds of malignancy	0.23:1	0.23:1	0.23:1	0.23:1	0.23:1
Post-test odds of malignancy with high risk test score	2.27:1	0.45:1	0.44:1	3.09:1	1.30:1
Pre-test odds of no malignancy	4.37:1	4.37:1	4.37:1	4.37:1	4.37:1
Post-test odds of no malignancy with low risk test score	15.50:1	30.71:1	51.00:1	15.83: 1	18.21:1

^a High risk cutoff: premenopausal subjects CA125 > 200 U/ml; postmenopausal subjects CA125-I > 35 U/ml.

^b High risk cutoff: premenopausal subjects CA125 > 67U/ml; postmenopausal subjects CA125-II > 35U/ml.

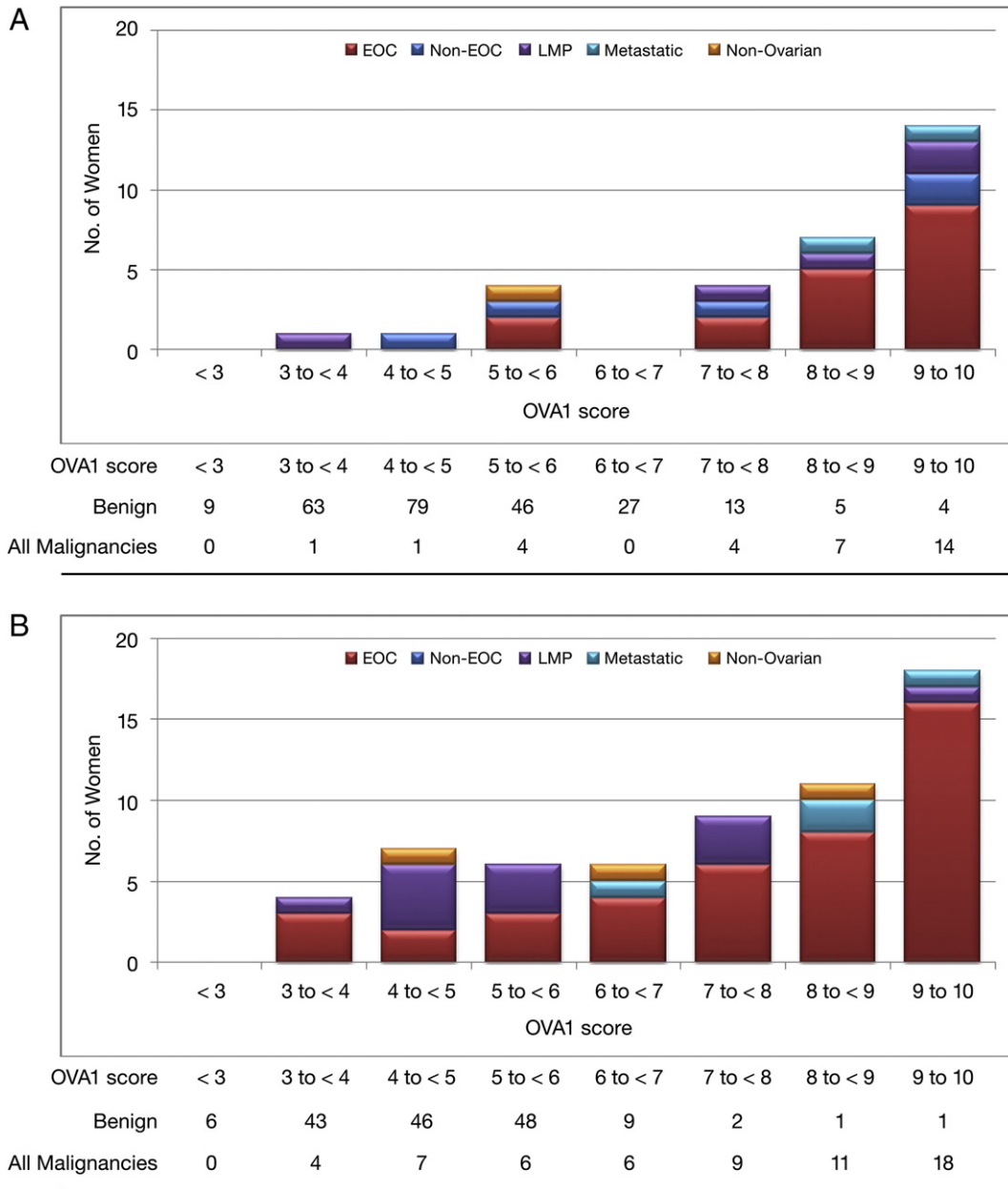


Fig. 2. OVA1 score distributions. A: pre-menopausal women ($n = 277$; 246 benign and 31 malignant cases). B: post-menopausal women ($n = 217$, 156 benign and 61 malignant cases).

with a higher likelihood of complete staging, optimal cytoreductive surgery, and superior overall survival [2,4,5,18–21]. Unfortunately, contemporary population-based studies have shown that the proportion of women with ovarian cancer that undergo surgery by these subspecialty surgeons is less than 50% [2,7–10]. Clearly, the accurate triage of women with an adnexal mass remains a clinical challenge for obstetrician/gynecologists and other providers of women's healthcare.

Ideally, all women with ovarian cancer would undergo surgery by a gynecologic oncologist. Two systematic approaches to the organization of care could accomplish this optimized healthcare delivery policy: 1) universal referral of all patients with an adnexal mass to a gynecologic oncologist or 2) selective referral based on a diagnostic triage strategy with high sensitivity for ovarian cancer [22,23]. While the universal referral strategy has been successfully implemented in single-payer healthcare systems, the complexity of the current multi-tiered healthcare coverage and reimbursement system in the United States makes adoption of such a model impractical [3]. Geography is also an impediment

to universal gynecologic oncology referral for patients with suspected ovarian cancer. Data from the Center for Disease Control and Prevention indicate that almost half of the incident ovarian cancer cases in the United States occur in a county without a gynecologic oncologist [24]. As a result, selective referral to specialty providers is necessary but requires striking a balance between the negative consequences of inappropriate under-referral and unnecessary over-referral.

Multiple decision-support algorithms are available to assist in the triage of women with an adnexal mass; however, none has received widespread acceptance [12]. The ACOG and Society of Gynecologic Oncology guidelines for the referral of women with a pelvic mass function well for detecting advanced-stage ovarian cancer but perform poorly in identifying patients with early-stage disease [13,14,16,25]. The Risk of Malignancy Index (RMI) is a risk stratification tool incorporating menopausal status, ultrasound findings, and CA125 but has a sensitivity of less than 80% [12,26]. Both the ACOG/SGO guidelines and the RMI are highly dependent on the CA125 level, although CA125 is

Table 3
OVA1® and CA125-II performance by stage and histological type.

	OVA1			CA 125-II ^a			CA 125-II ^b		
	n/N	Sensitivity	95% CI	n/N	Sensitivity	95% CI	n/N	Sensitivity	95% CI
Histologic subtype									
Epithelial ovarian cancer	57/ 60	95.0 ^{c,d}	86.2–98.3	49/60	81.7	70.1–89.4	52/60	86.7	75.8–93.1
Non-epithelial cancer	4/ 5	80.0	37.6–96.4	2/5	40.0	11.8–76.9	3/5	60.0	23.1–88.2
Low malignant potential	14/17	82.4 ^{c,d}	59.0–93.8	9/17	52.9	31.0–73.8	10/17	58.8	36.0–78.4
Metastatic	6/6	100	61.0–100	6/6	100	61.0–100	6/6	100	61.0–100
Non-ovarian cancer	4/4	100	51.0–100	2/4	50.0	15.0–85.0	2/4	50.0	15.0–85.0
Stage									
Stage I	25/28	89.3 ^{c,d}	72.8–96.3	18/28	64.3	45.8–79.3	20/28	71.4	52.9–84.7
Stage II	7/7	100	64.6–100	5/7	71.4	35.9–91.8	6/7	85.7	48.7–97.4
Stage III	24/25	96.0	80.5–99.3	23/25	92.0	75.0–97.8	24/25	96.0	80.5–99.3
Stage IV	5/5	100	56.6–100	5/5	100	56.6–100	5/5	100	56.6–100
Early Stage (Stage I & II)	32/35	91.4 ^{c,d}	77.6–97.0	23/35	65.7	49.2–79.2	26/35	74.3	57.9–85.8
Late Stage (Stage III & IV)	29/30	96.7	83.3–99.4	28/30	93.3	78.7–98.2	29/30	96.7	83.3–99.4
Premenopausal									
Early Stage (Stage I and II)	10/11	90.9 ^c	62.3–98.4	5/11	45.5	21.3–72.0	8/11	72.7	43.4–90.3
Late Stage (Stage III and IV)	12/12	100	75.8–100	11/12	91.7	64.6–98.5	12/12	100	75.8–100
Postmenopausal									
Early Stage (Stage I & II)	22/24	91.7 ^{c,d}	74.2–97.7	18/24	75.0	55.1–88.0	18/24	75.0	55.1–88.0
Late Stage (Stage III & IV)	17/18	94.4	74.2–99.0	17/18	94.4	74.2–99.0	17/18	94.4	74.2–99.0

a. High risk cut-off: premenopausal subjects CA125 > 200U/mL; postmenopausal subjects CA125 > 35U/mL.

b. High risk cut-off: premenopausal subjects CA125 > 67U/mL; postmenopausal subjects CA125 > 35U/mL.

c. Significant difference in sensitivity over CA125II^a (from McNemar's test $p < 0.05$).

d. Significant difference in sensitivity over CA125II^b (from McNemar's test $p < 0.05$).

not indicated for ovarian cancer preoperative assessment, has low sensitivity for early-stage ovarian cancer, and suffers from low specificity due to non-specific expression by other epithelial tissues [12,27].

As a result of the limitations of the above methods, novel serum biomarkers have been combined into multi-marker testing decision algorithms in attempts to improve the preoperative evaluation women with an adnexal mass and facilitate appropriate triage. The Risk of Ovarian Malignancy Algorithm (ROMA), which incorporates HE4, CA125, and menopausal status into a logistic regression model, has demonstrated sensitivity for ovarian malignancy in pre-menopausal women of 53.3–72.7% and a specificity of 74.2–87.9% [28–30]. In postmenopausal women, the reported sensitivity has been 82.5–90.8%, with specificity ranging from 66.3%–84.6% [28–30]. In 2011, Ueland et al published the pivotal trial for OVA1, which included 524 patients and demonstrated sensitivity for ovarian malignancy of 95.7% and negative predictive value of 95% when combined with clinical impression [15]. In this study, 47% of patients were enrolled by gynecologic oncologists, and the overall prevalence of ovarian malignancy was 29%. The current study was initiated with the intended use of OVA1 in mind, namely as a diagnostic triage aid to determine the need for gynecologic oncology referral. The objective was to validate the effectiveness of OVA1 in identifying ovarian malignancy, compared to clinical assessment and serum CA125-II, among women undergoing surgery for an adnexal mass after initial enrollment from non-gynecologic oncology providers.

In the intended-use population of the current dataset, OVA1 was more sensitive in detecting ovarian malignancy than clinical impression and CA125-II. When combined with clinical impression, the sensitivity for OVA1 (95.7%) was virtually identical to that observed by Ueland et al., validating its usefulness as a preoperative cancer referral test [15]. OVA1 had higher sensitivity than CA125-II in the especially challenging subsets of pre-menopausal patients (93.5%) and those with early-stage disease (91.4%). OVA1 correctly identified 83% of malignancies missed by clinical impression and 71% malignancies missed by CA125-II. As expected, the high sensitivity requirement for OVA1 as a cancer referral test was necessarily associated with a lower specificity and positive predictive value relative to clinical impression alone and CA125-II level. When combined with clinical impression, OVA1 correctly identified 204/402 patients with benign pathology, for a specificity of 50.7%.

Conversely, a low-risk OVA1 test result correctly predicted the absence of malignancy in 204/208 cases, translating into a negative predictive value of 98.1%.

The primary strength of the current study is that it is a large, prospective, multi-institutional patient cohort representing the intended-use population for the biomarker panel under investigation. In addition, biomarker testing was independently performed and validated, and all ovarian tumor types were included in the statistical analysis of test performance.

A potential limitation of the current study is that inclusion criteria required that all subjects were to undergo a planned surgery for an adnexal mass. As such, the prevalence of pelvic malignancy (18.6%) is higher than would be expected in a population of patients for whom surgery is not planned. An inverse relationship between positive predictive value (decrease) and negative predictive value (increase) would be expected in a patient population with a lower prevalence if cancer. In the current dataset, the prevalence of malignancy was lower in the pre-menopausal patient subset (11.2%) compared to the postmenopausal group (28.1%) although the sensitivity of OVA1 was similar in both groups, demonstrating consistency of performance despite varying cancer incidence.

A second potential limitation of the current study is an inability to interrogate the process of physician assessment in reaching the clinical impression. Physician assessment is a confounding variable in the adnexal mass clinical management algorithm because the contributing factors are unknown. Although OVA1 is FDA-cleared for use with clinical impression, the independent performance of the test yields a risk stratification score that can be reliably controlled and offers a more standardized practice algorithm than clinical impression, which varies between studies and from clinician to clinician. As a risk stratification test, OVA1 detected 92.4% of all ovarian malignancies, 91.4% (32/35) of early-stage cancers, and 93.5% (29/31) of malignancies in pre-menopausal patients. In a classic “needle-in-a-haystack” scenario, OVA1 correctly identified 10/11 (90.9%) Stage I/II cancers among the group of 277 pre-menopausal patients while maintaining a specificity of 61.4% in the current dataset (Table 4). The current data combined with the report from Ueland et al. represent over 1000 prospectively studied patients and demonstrate highly reproducible performance as a “first-time-right” ovarian cancer surgery referral test overall, by

Table 4
Comparison of OVA1® Performance Characteristics from Original Pivotal Trial (Ueland et al.) [15] and Present Study.

	All evaluable subjects		Premenopausal		Postmenopausal	
	Ueland et al.	Present study	Ueland et al.	Present study	Ueland et al.	Present study
Evaluable subjects, N	516	494	235	277	281	217
Prevalence, %	31.2	18.6	19.1	11.2	41.3	28.1
Sensitivity, %	92.5	92.4	86.7	93.5	94.8	91.8
95% CI	87.4–95.7	85.1–96.3	73.8–93.7	79.3–98.2	89.2–97.6	82.2–96.4
Specificity, %	42.8	53.5	51.6	61.4	32.7	41.0
95% CI	37.8–48.0	48.6–58.3	44.5–58.6	55.2–67.2	26.0–40.2	33.6–48.9
PPV, %	42.3	31.3	29.8	23.4	49.8	37.8
95% CI	37.3–47.5	26.0–37.0	22.6–38.1	16.8–31.6	43.2–56.3	30.4–45.9
NPV, %	92.7	96.8	94.2	98.7	90.0	92.8
95% CI	87.6–95.8	93.6–98.5	88.0–97.3	95.4–99.6	79.9–95.3	84.1–96.9
Ovarian cancer, N	105	65	28	23	77	42
Stage I: Sensitivity, %	90.3	89.3	66.7	88.9	100	89.5
95% CI	75.1–96.7	72.8–96.3	35.4–87.9	56.5–98.0	85.1–100	68.6–97.1
Stage II: Sensitivity, %	100	100	100	100	100	100
95% CI	82.4–100	64.6–100	67.6–100	34.2–100	72.2–100	56.6–100
Stage III: Sensitivity, %	100	96.0	100	100	100	93.3
95% CI	93.0–100	80.5–99.3	72.2–100	72.2–100	91.4–100	70.2–98.8
Stage IV: Sensitivity, %	100	100	–	100	100	100
95% CI	43.9–100	56.6–100	–	34.2–100	43.9–100	43.9–100

Note: Prevalence includes all malignancies: primary ovarian malignancies, LMP, metastatic to ovaries and non-ovarian malignancies.

menopausal status, and by stage of disease (Table 4) [15]. Combining the results of these two independent and prospectively conducted studies, OVA1 correctly identified 78/84 patients (92.8%) with Stage I/II disease as an independent risk stratification tool.

Finally, the current study was not designed to address the impact on referral patterns of the lower level of specificity with OVA1, which could lead to referral of a higher proportion of patients with non-malignant tumors and has been raised as a possible concern by general obstetrician/gynecologists. When OVA1 was combined with clinical impression, the overall specificity was 50.7%, improved from 34.6% in the Ueland study. Considering the high sensitivity and negative predictive value, utilization of OVA1 as a risk stratification instrument would be associated with retention of over 50% of patients with benign adnexal masses, ensure that more than 96% of retained surgical patients would not have ovarian cancer, and provide appropriate referral to a gynecologic oncologist for over 95% of patients with ovarian cancer.

Conflict of interest statement

Dr. Robert E. Bristow was the principal investigator for the OVA1 trial and has been a member of Vermillion Inc's speaker's bureau since November 2011. He has not received honoraria from Vermillion Inc.

Donald G. Munroe, Ph.D. is an employee of Vermillion Inc., which funded and sponsored this study. His position is Chief Scientific Officer and VP Research & Development for Vermillion Inc., and Dr. Munroe oversaw the Clinical Study from October 11, 2011 to the present.

Dr. Eric Fung was an employee of Vermillion when the work was conducted. Dr. Fung owns stock in Vermillion.

Gillian Crutcher is Senior Director of Clinical, Regulatory, Quality Systems, and Manufacturing at Vermillion, Inc. and owns stock in Vermillion.

Zhen Zhang, PhD is a faculty of the Johns Hopkins University School of Medicine. He is co-inventor of patents associated with the OVA1 product and is entitled to royalty payments from the sale of OVA1 test through a license agreement between Vermillion Inc. and Johns Hopkins University. His work on OVA1 has been funded through sponsored research agreements between Vermillion Inc. and Johns Hopkins University.

Daniel W. Chan, PhD serves on the Advisory Board at Vermillion Inc.

Alan Smith is Vice-President, Biometrics at Applied Clinical Intelligence and is a consultant for Vermillion Inc.

Author Contribution Statement

Study design: Bristow, Smith, Zhang, Chan, Crutcher, Fung, Munroe.

Data collection and analysis: Bristow, Smith, Zhang, Chan, Crutcher, Fung, Munroe.

Manuscript preparation: Bristow, Smith, Munroe.

Final approval of manuscript: Bristow, Smith, Zhang, Chan, Crutcher, Fung, Munroe.

Acknowledgments

Dr. Robert E. Bristow was supported in part by the Queen of Hearts Foundation.

References

- [1] Siegel R, Naishadham D, Jemal A. Cancer statistics 2012. *CA Cancer J Clin* 2012;62:10–29.
- [2] 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.
- [3] Tingulstad S, Skjeldestad FE, Hagen B. The effect of centralization of primary surgery on survival in ovarian cancer patients. *Obstet Gynecol* 2003;102:499–505.
- [4] Eisenkop SM, Spirtos NM, Montag TW, Nalick RH, Wang HJ. The impact of subspecialty training on the management of advanced ovarian cancer. *Gynecol Oncol* 1992;47:203–9.
- [5] Engelen MJA, Kos HE, Willemse PHB, Aalders JG, de Vries EGE, Schaapveld M, et al. Surgery by consultant gynecologic oncologists improves survival in patients with ovarian carcinoma. *Cancer* 2006;106:589–98.
- [6] Vernooj 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;105:801–12.
- [7] Olaitan A, Weeks J, Mocroft A, Smith J, Howe K, Murdoch J. The surgical management of women with ovarian cancer in the south west of England. *Br J Cancer* 2001;85:1824–30.
- [8] Elit L, Bondy SJ, Paszat L, Przybysz R, Levine M. Outcomes in surgery for ovarian cancer. *Gynecol Oncol* 2002;87:260–7.
- [9] Vernooj F, Heintz APM, Witteveen PO, van der Heiden-van der Loo M, Coebergh JW, van der Graaf Y. Specialized care and survival of ovarian cancer patients in the Netherlands: nationwide cohort study. *J Natl Cancer Inst* 2008;100:399–406.
- [10] Earle CC, Schrag D, Neville BA, Yabroff R, Topor M, Fahey A, et al. Effect of surgeon speciality on processes of care and outcomes for ovarian cancer patients. *J Natl Cancer Inst* 2006;98:173–80.
- [11] Boyd LR, Novetsky AP, Curtin JP. Ovarian cancer care for the underserved: are surgical patterns of care different in a public hospital setting? *Cancer* 2011;117:777–83.
- [12] Dodge JE, Covens AL, Lacchetti C, Elit LM, Le T, Devries-Aboud M, et al. Preoperative identification of a suspicious adnexal mass: a systematic review and meta-analysis. *Gynecol Oncol* 2012;126:157–66.
- [13] The role of the generalist obstetrician gynecologist in the early detection of ovarian cancer. ACOG Committee Opinion Number 280. American College of Obstetricians and Gynecologists. *Obstet Gynecol* 2002;100:1413–6.
- [14] The role of the obstetrician-gynecologist in the early detection of epithelial ovarian cancer. ACOG Committee Opinion Number 477. American College of Obstetricians and Gynecologists. *Obstet Gynecol* 2011;117:742–6.

- [15] Ueland FR, DeSimone CP, Seamon LG, Miller RA, Goodrich S, Podzielinski I, et al. Effectiveness of a multivariate index assay in the preoperative assessment of ovarian tumors. *Obstet Gynecol* 2011;117:1289–97.
- [16] Dearing 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;110:841–8.
- [17] Hoskins W, Rice L, Rubin S. Ovarian cancer surgical practice guidelines. *Oncology* 1997;11:896–904.
- [18] Grossi M, Quinn MA, Thursfield VJ, Francis PA, Rome RM, Planner RS, et al. Ovarian cancer: patterns of care in Victoria during 1993–1995. *Med J Aust* 2002;177:11–21.
- [19] Cress RD, Bauer K, O'Malley CD. CR ovarian patterns of care study. *Gynecol Oncol* 2011;121:94–9.
- [20] Goff BA, Matthews BJ, Wynn M, Muntz HG, Lishner DM, Baldwin LM. Ovarian cancer: patterns of surgical care across the United States. *Gynecol Oncol* 2006;103:383–90.
- [21] Giede KC, Kieser K, Dodge J, Rosen B. Who should operate on patients with ovarian cancer? An evidence-based review. *Gynecol Oncol* 2005;99:447–61.
- [22] Kim KH, Zsebk GN, Straughn JM, Landen CN. Management of complex pelvic masses using a multivariate index assay: a decision analysis. *Gynecol Oncol* 2012, <http://dx.doi.org/10.1016/j.ygyno.2012.05.027>.
- [23] Geomini PM, Kruitwagen RF, Bremer GL, Massuger L, Mol BW. Should we centralize care for the patient suspected of having an ovarian malignancy? *Gynecol Oncol* 2011;122:95–9.
- [24] Stewart SL, Rim SH, Ricjards TB. Gynecologic oncologists and ovarian cancer treatment: avenues for improved survival. *J Womens Health* 2011;20:1257–60.
- [25] Im SS, Gordon AN, Buttin BM, Leath CA, Gostout BS, Shah C, et al. Validation of referral guidelines for women with pelvic masses. *Obstet Gynecol* 2005;105:35–41.
- [26] Geomini P, Kruitwagen R, Bremer GL, Cnossen J, Mol BWJ. The accuracy of risk scores in predicting ovarian malignancy. *Obstet Gynecol* 2009;113:384–94.
- [27] Bast Jr RC, Xu FJ, Yu YH, Barnhill S, Zhang Z, Mills GB. CA-125: the past and the future. *Int J Biol Markers* 1998;13:179–87.
- [28] Moore RG, Miller C, Disilvestro P, Landrum LM, Gajewski W, Ball JJ, et al. Evaluation of the diagnostic accuracy of the risk of ovarian malignancy algorithm in women with a pelvic mass. *Obstet Gynecol* 2011;118:280–8.
- [29] Montagnana M, Danese E, Ruzzenente O, Bresciani V, Nuzzo T, Gelati M, et al. The ROMA (risk of ovarian malignancy algorithm) for estimating the risk of epithelial ovarian cancer in women presenting with pelvic masses: is it really useful? *Clin Chem Lab Med* 2011;49:521–5.
- [30] Van Gorp T, Cadron I, Despierre E, Daemen A, Leunen K, Amant F, et al. HE4 and CA125 as a diagnostic test in ovarian cancer: prospective validation of the risk of malignancy algorithm. *Br J Cancer* 2011;104:863–70.