Validation of a Multivariate Serum Profile for Epithelial Ovarian Cancer using a Prospective Multi-Site Collection

Partha Seshaiah[1], Greg P. Bertenshaw[1], Tzong-Hao Chen[1], Katharine J. Bergstrom[1], Jinghua Zhao[1], James P. Mapes[2], Laurie L. Stephen[2], Suraj Amonkar[1], Michael E. McCollum[3], Brigitte E. Miller[4], Lynda D. Roman[5], Beth Y. Karlan[6], Eva Chalas[7], Paul A. DiSilvestro[8], James F. Barter[9], James W. Orr, Jr.[10], Glenn E. Bigsby IV[11], Robert W. Holloway[11], Ronald D. Alvarez[12], Ping F. Yip[1], Brian C. Mansfield[1]*

[1] Correlogic Systems, Inc., Germantown, MD; [2] Rules-Based Medicine, Inc., Austin, TX; [3] The Harry and Jeanette Weinberg Cancer Institute at Franklin Square Hospital, Baltimore, MD; [4] North East Oncology Associates, Concord, NC; [5] University of Southern California, Norris Cancer Center, Los Angeles, CA and Women's and Children's Hospital, Los Angeles, CA; [6] Women's Cancer Research Institute at the Samuel Oschin Comprehensive Cancer Institute, Cedars-Sinai Medical Center, Los Angeles, CA; [7] SUNY at Stony Brook, NY, Stony Brook, NY; [8] Women and Infants Hospital of Rhode Island, Providence, RI; [9] Holy Cross Hospital, Silver Spring, MD; [10] Florida Gynecologic Oncology, Fort Meyers, FL; [11] Florida Hospital Cancer Institute, Orlando, FL; [12] University of Alabama at Birmingham, Birmingham, AL.

* Brian C. Mansfield, PhD, Correlogic Systems, Inc., 20271 Goldenrod Lane, Suite 2070 Germantown, MD 20876 USA; e-mail: bmansfield@correlogic.com

ABSTRACT

In previous studies we described the use of a retrospective collection of ovarian cancer and benign disease samples, in combination with a large set of multiplexed immunoassays and a multivariate pattern recognition algorithm, to develop an 11-biomarker classification profile that is predictive for the presence of epithelial ovarian cancer. In this study, customized, Luminex-based multiplexed immunoassay kits were GMP-manufactured and the classification profile was refined from 11 to 8 biomarkers (CA-125, epidermal growth factor receptor, CA 19-9, C-reactive protein, tenascin C, apolipoprotein AI, apolipoprotein CIII, and myoglobin). The customized kits and the 8-biomarker profile were then validated in a double-blinded manner using prospective samples collected from women scheduled for surgery, with a gynecologic oncologist, for suspicion of having ovarian cancer. The performance observed in model development held in validation, demonstrating 81.1% sensitivity (95% CI 72.6 – 87.9%) for invasive epithelial ovarian cancer and 85.4% specificity (95% CI 81.1 – 88.9%) for benign ovarian conditions. The specificity for normal healthy women was 95.6% (95% CI 83.6 – 99.2%). These results have encouraged us to undertake a second validation study arm, currently in progress, to examine the performance of the 8-biomarker profile on the population of women not under the surgical care of a gynecologic oncologist.

INTRODUCTION

Ovarian cancer is the most lethal gynecological cancer. Within the USA, there are approximately 21,880 new cases and 13,850 deaths each year.[1] More than 80% of these cases are detected as late stage disease with poor survival rates.[1,2] The best clinical outcome for patients with advanced disease depends upon complete surgical staging and optimal debulking.[3] These are complex surgeries that are best performed by gynecologic oncologists.[4,5] However, in the USA, more than 50% of women with ovarian cancer receive care from non-specialists and up to 80% of these individuals receive inadequate staging.[4,6] Since only 13-21% of women scheduled for surgery based on symptoms consistent with ovarian cancer, actually have ovarian cancer, and there are only ~ 1,000 of board certified gynecologic oncologists in the USA, a test that can both increase the number of cancer cases referred to a specialist, and increase the prevalence of cancers within that referred population, by minimizing the number of benign referrals, would be useful.

There are no highly accurate tests for the detection of ovarian cancer. For the assessment of cancer risk, an oncologist usually assesses multiple non-specific lines of evidence including family history, patient symptoms, a physical examination, a radiographic evaluation, and

laboratory findings. The common symptoms associated with the disease, such as pelvic and abdominal pain, urinary urgency, urinary frequency, abdominal bloating, and difficulty eating, do not differentiate well between cancerous and benign conditions.[7] A physical exam and radiographic evidence can help in the detection of a pelvic mass, however the commonly used imaging techniques – transvaginal sonography (TVS), positron-emission tomography (PET), magnetic resonance imaging (MRI), radio-immunoscintigraphy and computed tomography (CT) – tend to have a low specificity in the hands of many radiographers.[8] While some reports suggest ultrasound alone, or in combination with other prognostic variables, may be significantly more informative in the hands of an ovarian ultrasound expert,[9,10] many patients lack access to such specialized imaging services.

Most laboratory findings focus on the elevation of CA-125 associated with ovarian cancer. However, despite widespread use, CA-125 is not FDA-approved for diagnosis. A critical drawback for CA-125 is a lack of specificity, being elevated in many normal/benign conditions and non-ovarian malignancies.[11-13] Approaches to improve the predictive value of CA-125 through serial measurements,[14,15] or in combination with additional markers,[16,17] have been reported, but remain under study.

Professional societies, such as ACOG/SGO have provided recommended guidelines for referral to a specialist[18-21] and at least one report has suggested a unique combination of symptoms, if fully documented for each patient, may be more informative than previously recognized.[22] Despite these advances, more than 80% of patients referred to a specialist will have a benign condition.[8,23,24]

In the absence of a single informative biomarker, attention has focused on the combination of multiple biomarkers, in conjunction with non-obvious algorithms, to create multivariate index assays. Progress on three multivariate biomarker algorithms has been reported recently. A six-biomarker panel, composed of leptin, prolactin, osteopontin, insulin-like growth factor II, macrophage inhibitory factor, and CA-125 was reported to have a sensitivity of 95.3% and a specificity of 99.4%.[25] However, these results were not validated using independent samples, and contained a notable population bias, with all cases representing women committed to surgery for ovarian exploration, and all controls representing apparently healthy women attending a regular gynecologic exam. A 2-biomarker assay (risk of malignancy algorithm; ROMA), based on CA-125 and HE4 has also been reported recently.[17] This assay uses separate algorithms for pre- and post-menopausal patients. For post-menopausal women, at a fixed specificity of 75.0% (95% CI 66.9 – 81.4%), the sensitivity was 92.3% (95% CI 85.9 – 96.4%). For pre-menopausal women, at a fixed specificity of 74.8% (95% CI 68.2 – 80.6%) the sensitivity was 76.5% (95% CI 58.8 – 89.3%). Unfortunately this study failed to use an independent validation sample set, the validation samples being used themselves to identify the ROMA cut-off values that yielded the "clinically acceptable" 75% specificity. Therefore the performance of this assay, with pre-established cut-off values, remains unknown. Most recently, OVA1, a 5-biomarker ovarian malignancy assay, combining CA-125, transthyretin, apolipoprotein A-I (ApoA-I), transferrin, and beta-2 microglobulin has been validated in a prospective, blinded clinical study submitted to the USA Food and Drug Administration (FDA).[26] The results have yet to be published. In the FDA regulatory documents,[26] the performance of the assay varies significantly depending on the source of the surgical patient population (specialist or non-specialist oncologist), the menopausal status of the patient, and whether the assay is combined with a pre-surgical impression or not. Across these scenarios, the sensitivity of test lies between 80.8 and 98.9%, while the specificity ranges between 24.4 and 56.8%. In general, sensitivity increased at the expense of specificity for post-menopausal women. A similar trend occurred when the biomarker results were combined with pre-surgical assessment. This biomarker assay is not for stand-alone use or screening, but for referral of a restricted subset of surgical patients, in the care of a non-specialist, whose pre-surgical assessment does not indicate malignancy. Since the low specificity yields a high false positive rate, a test with a higher specificity could improve specialist referral by lowering the over-referral of benign cases.

In previous work we described a biomarker discovery approach in which we profiled stage I invasive epithelial ovarian cancer (EOC) and ovarian benign conditions, using a highly multiplexed immunoassay discovery platform.[27,28] Our hypothesis was that by avoiding the use of late stage cancers in discovery, which have elevated CA-125 levels, we would identify biomarkers independent of CA-125. From a total of 204 biomarkers we identified a unique combination of 11 biomarkers, which appeared informative for the presence of all stages of invasive EOC in women scheduled for surgery. In preliminary testing using a multi-site retrospective collection of stage II-IV EOC and benign ovarian condition samples, collected from gynecologic oncology practices, this classification profile demonstrated a sensitivity of 91.3% (95% CI 84.2 – 95.5%) and a specificity of 88.5% (95% CI 81.4 – 93.2%). These results were very promising. Following this, we manufactured several lots of a custom-designed multiplexed immunoassay kit for the 11 biomarkers and used that to develop a refined version of this classifier. The improved classifier consists of a subset of 8 biomarkers (CA-125, CA 19-9, C-reactive protein (CRP), ApoA-I, apolipoprotein CIII (ApoC-III), epidermal growth factor receptor (EGF-R), myoglobin and tenascin C). We now report a completely

independent validation of this 8-biomarker profile using a prospectively collected set of samples from gynecologic oncologists practices. Using a double-blinded experimental protocol, we show that the validated performance is consistent with the results seen along the biomarker discovery pathway, exhibiting a sensitivity of 81.1% (95% CI 72.6 – 87.9%) and specificity of 85.4% (95% CI 81.1 – 88.9%), validating our biomarker discovery strategy. We also show that the classifier meets the generally accepted criterion of an informative test.

MATERIALS AND METHODS

Prospective Patient Population

Sera from 1398 qualified individual participants were collected prospectively for the purpose of validating an ovarian cancer test. Of these, 260 had EOC, with 213 of the 260 being invasive and 47 being borderline-low malignant potential (LMP) cases. Twelve gynecologic oncology sites participated in the study, which was approved by the Institutional Review Board (IRB) of each participating site and by the Western IRB (Olympia, WA) for Correlogic Systems Inc. (Germantown, MD) under Study Number 1064143. The sites were Cedars-Sinai Medical Center; Wake Forest University, Winston-Salem; SUNY at Stony Brook; Florida Hospital Cancer Institute; Holy Cross Hospital; Florida Gynecologic Oncology; University of Southern California: Norris Cancer Center and Women's and Children's Hospital; Women and Infants Hospital of Rhode Island; University of Alabama; North Shore – Long Island Jewish Health System; and The Harry and Jeanette Weinberg Cancer Institute at Franklin Square Hospital. All sites provided both patients with EOC and benign ovarian conditions.

The study inclusion criteria were women, at least 18 years of age, symptomatic of ovarian cancer according to the National Comprehensive Cancer Network (NCCN) Ovarian Cancer Treatment Guidelines for Patients,[7] which includes women with or without a pelvic mass. Participants had to be scheduled for gynecologic surgery based on concern they had ovarian cancer, and post-surgical pathological evaluation of the ovaries and excised tissues was required. Exclusion criteria were women who did not meet the inclusion criteria, could not provide informed consent, were pregnant, or previously treated for ovarian cancer.

Prospective Serum Collection

Sera were collected, prior to intervention, using a well-defined blood collection protocol. Briefly, blood was drawn into red top Vacutainer tubes (Becton, Dickinson and Company; Franklin Lakes, NJ), clotted for at least 30 minutes at room temperature and centrifuged at 1,300g, at room temperature, to separate serum. Sera were then transferred to screw-top cryogenic vials and frozen (-80°C) within 2 hours of the draw. None of the samples used in the validation study were thawed or used in any manner prior to this validation study.

Nominally Healthy Individuals

Sera from healthy volunteers were collected by a commercial specimen bank (ProMedDx; Norton, MA) that was not involved in the gynecologic oncology collection. Sera were collected, using the same blood collection protocol, from 45 women (aged 19-62) and from 30 males (aged 18-57). A health review questionnaire was completed for each participant but ovarian pathology was not available for the women.

Blinding of Serum Samples and Clinical Data

The serum samples were blinded to both the analytical testing site, Rules-Based Medicine (RBM; Austin, TX), and the classification scoring site, Correlogic Systems, Inc. An independent, contract statistician experienced in clinical study design, identified the specific subset of samples to use in the study in a manner designed to ensure they were representative of the larger prospective collection. Two independent contract research organizations (CRO) worked with the statistician. CRO-1 performed all sample handling including masking, labeling, aliquoting, and shipping to the analytical test site. CRO-2 performed electronic encoding of the patient data and, following the classification of the masked samples, linked the results to the patient data for analysis by the statistician.

Study Design and Sample Size

The study was designed to demonstrate whether the 8-biomarker profile is informative for the detection of EOC in women, under the care of a gynecologic oncologist, with NCCN-defined symptoms of ovarian cancer, who are scheduled for surgery; and to estimate the sensitivity, specificity, positive predictive value (PPV), and negative predictive value (NPV), of the profile on this population.

A generally accepted criterion for an informative test is that the 95% lower confidence level (LCL) for the sum of sensitivity and specificity should be greater than 100%. Prior to the validation, and based on our model development results, we established a more stringent hypothesis that the point performances of sensitivity and specificity of the 8-biomarker assay should each exceed 80% and the exact one-sided LCL of the sum of sensitivity and specificity should be greater than 135%.

To achieve statistical significance, the hypothesis study design required a minimum of 103 patients with EOC, and 309 patients with benign ovarian disease. To accommodate any sample handling or assay errors, these numbers were increased an additional 10% to give a final sample size of 113 EOC patients and 335 control patients. To select

samples from the larger prospective collection, the statistician stratified the prospective collection by study site, to ensure a proportional representation of participating clinical sites, and from that selected the required number of samples randomly.

Immunoassays

We previously reported an 11-biomarker profile for epithelial ovarian cancer,[27] for which we have now constructed custom Luminex-based, multiplexed immunoassay panels. The custom kits, designed for an automated, high-throughput use, are composed of two multiplex panels. Panel A, an 8-plex composed of assays for CA-125, CA 19-9, EGF-R, tenascin C, myoglobin, interleukin-6, interleukin-18, and macrophage inflammatory protein-1-alpha (MIP-1a), is assayed at a 1:5 dilution. Panel B, a 4-plex composed of assays for CRP, ApoA-I, ApoC-III, and fibrinogen (a quality check for serum) is assayed at a 1:5,000 dilution. Four independent assay lots of these kits were manufactured under good manufacturing practices (GMP). Kits are available commercially, for research use only, from Correlogic Systems, Inc.

Multiplexed immunoassays were performed in 96-well plates at room temperature. Sera were diluted in phosphate-buffered saline (PBS) containing 4% bovine serum albumin (BSA), pH 7.4 and assayed in a buffer containing 1% BSA, 0.05% ProClin 300, and a proprietary mix of antibodies and domestic animal proteins to minimize non-specific interactions. Each 96 well assay plate contained: 16 wells for standards, consisting of 2 sets of 8 different dilutions of pooled standards, used to create 8-point calibration curves; 6 wells of pooled quality controls, consisting of 2 sets of low, medium and high levels for each biomarker; and up to 72 wells for patient samples. The validation study was performed over 2 days using a combination of 3 manufacturing lots of the immunoassay kits, 5 different robots, 11 different Luminex instruments and 2 operators.

The 8-Biomarker Profile

During the discovery and development phase,[27,28] sera were analyzed using RBM's Human MAP v1.6, a set of multiplexed assays that measured 204 serum molecules. Prior to validation the 11 assays identified in the previous work, plus fibrinogen, were reformatted into a custom multiplex kit described above. Since multiplex assay performance characteristics change when assays are reformatted, the model development analysis previously described was repeated, using the same model development samples, same modeling strategy and same KDE-VS modeling algorithm.[27] The only difference this time was that the samples were assayed with the custom manufactured 12-biomarker multiplex kit rather than the multiple panels that make up Human MAP v1.6. Using these new data, the 11-biomarker classification model was refined to a single model that used only 8 of the previous 11 biomarkers (CA-125, CA 19-9, CRP, ApoA-I, ApoC-III, EGF-R, myoglobin and tenascin C), with no statistically significant change in performance. Importantly, none of the samples evaluated in the blinded validation study were used in this refinement.

The final classification profile or "model", consists of a series of fixed mathematical rules, that interpret the relative concentration of the 8 biomarkers in a patient's serum. The rules are encoded into a software package that automates the analysis[27] and provides a simple binary outcome for each patient that reflects either a low or high risk of having ovarian cancer.

RESULTS AND DISCUSSION

Patient Characteristics

Women participating in the prospective serum collection were scheduled for surgery, based on a concern that they had ovarian cancer. We selected a subset of these women for the double-blinded validation study as described in Materials and Methods. A comparison of the key characteristics demonstrate that the validation subset fairly represents the entire prospective collection (Table 1).

In our study population the incidence of invasive EOC was 15.2%, the incidence of borderline – low malignant potential (LMP) tumors within the EOC was 18.1%. The incidence of any malignancy was 31.3%.

In the overall population the mean age of the EOC patients was 61.1 ±11.7 years while women with benign conditions had a mean age of 53.2 ± 13.5 years. A pelvic mass was present in 1293 out of 1335 (96.9%) of patients who reported pelvic mass information. Women with EOC predominantly had a pelvic mass (211 out of 213) while there were slightly more benign cases without a mass (30 out of 1185). Of women presenting with EOC, 72.8% were self reported as post-menopausal compared to 50.8% of women with benign conditions. The dominant subtype and stage of EOC was serous and stage III disease, respectively, with distant disease (stage III/IV) representing 63.8% of all pathology confirmed EOC and 70.5% of all staged EOC.

Validation Results

During analysis, two samples were lost due to a pipette malfunction. The remaining samples were successfully assayed and classified. Upon completion of the study, unmasking identified both lost samples as stage III cancers leaving 111 invasive EOC and 335 benign samples. The test had a sensitivity of 81.1% (95% CI 72.6 – 87.9%) and specificity of 85.4% (95% CI 81.1 – 88.9%; Table 2). The exact one-sided LCL of the sum of the sensitivity and specificity for the validation sample set was 160% yielding a normal z statistic of 7.5 ($p < 0.0001$).

Table 1. Comparison of the Prospective Collection to the Blinded Validation Subset

	Prospective Collection	Validation Study
Total patients n (%)	1398	448 *
Invasive EOC	213	113 *
FIGO Stage I	35 (16.4)	17 (15.0)
II	22 (10.3)	10 (8.8)
III	125 (58.7)	70 * (61.9)
IV	11 (5.2)	6 (5.3)
Stage unknown	20 (9.4)	10 (8.8)
Benign ovarian condition	930	335
Other gynecologic malignancy	224 †	0
Pathology not definitive	31	0
Age – Mean (±SD)		
Invasive EOC	61.1 (±11.7) ‡	61.4 (±11.5)
Benign condition	53.2 (±13.5) §	53.0 (±14.5)
Pelvic Mass n (%)		
Present	1293 (96.9)	439 (98.0)
Absent	42 (3.1)	9 (2.0)
No information	63 (N/A)	0 (N/A) ||

* Includes two stage III EOC samples which were not assayed due to a technical problem. † 89 of these samples were assayed following completion of the validation study, see Table 5. ‡ Age not reported for one individual, calculation based on 212 individuals. § Age not reported for six individuals, calculation based on 920 individuals. || Pelvic mass information was required. Abbreviations: EOC, epithelial ovarian cancer; FIGO, International Federation of Gynecology and Obstetrics; SD, standard deviation.

These findings met the criteria established prior to the study, independently validating the performance of the 8-biomarker classification profile, meeting our pre-validation definition of an informative test, and meeting the generally accepted criteria of a clinically useful test.[29]

There was no statistical difference in either the sensitivity or specificity of the 8-biomarker profile when patients with or without a pelvic mass were compared (Table 2). The one EOC and eight benign samples from patients without masses were all classified correctly. While the performance for symptomatic women lacking a pelvic mass is promising, the low number of these patients does not provide sufficient statistical support to validate the profile as informative within this group.

PPV and NPV

The 25% prevalence of EOC in the subset of prospective samples used in the validation study was a product of the statistical design of the hypothesis study. In the entire prospective collection, representative of the studied population, the prevalence was 15.2%. Extrapolation of the measured sensitivity and specificity of this population to the entire collection yields a PPV of 49.5% and NPV of 96.2%.

Stage and Subtype of Ovarian Cancer

The sensitivity for late stage disease (>90%) was greater than for early stage disease (Table 3). In our previous model development studies we estimated stage II - IV sensitivity of 91.3% (95% CI 84.2 – 95.5%).[27] In this prospective validation we observed a stage II - IV sensitivity of 88.1% (95% CI 78.8 – 93.8%), a point estimate lower than the initial point estimate, but still within the 95% confidence limits and consistent with the differences expected between testing estimates and validation. For specificity, the model development estimate of 88.5% (95% CI 81.4 – 93.2%) was close to the observed validation specificity of 85.4% (95% CI 81.1 – 88.9%).

Table 2. Sensitivity, Specificity and Informative Value of the 8-Biomarker Profile

	Epithelial Ovarian Cancer			Benign Conditions		
Population	**SN**	**95% CI**	**n/N**	**SP**	**95% CI**	**n/N**
All	81.1%	72.6 – 87.9	90/111	85.4%	81.1 – 88.9	286/335
Pelvic Mass						
With	80.9%	72.3 – 87.8	89/110	85.0%	80.7 – 88.7	278/327
Without	100.0%	2.5 – 100.0	1/1	100.0%	63.1 – 100.0	8/8
Menopause						
Post	80.9%	70.9 – 88.2	72/89	89.2%	83.8 – 93.3	165/185
Not post	89.5%	65.5 – 98.2	17/19	79.7%	72.2 – 86.0	114/143
Unknown	33.3%	1.8 – 87.5	1/3	100.0%	59.0 – 100.0	7/7

Abbreviations: SN, sensitivity; SP, specificity; 95% CI, 95% two-sided exact binomial confidence interval; n, number of samples classified correctly; N, number of samples assayed.

The retrospective sample set used in model development was deliberately enriched for stage I samples.[27] Since the majority of these samples were used to select the biomarkers and develop the classification profile, it had left few stage I samples for performance testing. In those earlier development studies, we used bootstrap analysis to estimate the profile had a stage I sensitivity of 83.4% ±12.4% with the caveat that this strategy limited the statistical confidence of this extrapolation. We could not confirm this performance with this study, and indeed this highlights the need for caution in interpreting bootstrap estimates in biomarker studies. In our prospective study population, stage I disease represented only 16.4% of the staged invasive EOC and the small number of stage I cancers prevented statistically significant conclusions about early stage performance. While the apparent performance on early stage disease was lower than anticipated, the surgical value of specialist referral may be much greater for the more complex later stage disease, where the profile scored well.[4,5] Sensitivity for the two major subtypes of EOC, serous (85.5%) and endometrioid (90.0%) was statistically similar (Table 4). There were too few clear cell, mucinous, undifferentiated, transitional cell and Brenner samples to make a judgment on these subtypes of EOC.

Other Analyses

While the double-blinded validation study was limited to studying invasive EOC, there is significant interest in LMP cancers. We decided to exclude LMP cancers for the double-blinded validation study based on several considerations. In the 2001 International Classification of Diseases publication (ICD-0-3), the World Health Organization changed the classification of LMP tumors, from malignant to non-malignant, suggesting our study should assign these as benign conditions, contributing to estimates of specificity. On the other hand, while most LMP tumors are non-aggressive and respond well to surgery alone, with 10-year survival rates in excess of 95%,[30] a small proportion (~5%) are aggressive, which would suggest they should be grouped with EOC and contribute to estimates of sensitivity. Finally, there is an increasing concern in the literature that detection of LMP represents over-diagnosis.[31,32] Given the lack of clarity on the classification, we chose to report performance on LMP separately to the invasive EOC samples. To this end, the LMP samples were classified in a separate single-blinded study. No preoperative biomarker or pathognomonic ultrasound features have been described for diagnosing LMP.[33] Of the 42 LMP samples tested, 23 (54.8%) were classified as positive, a near neutral outcome (Table 5).

The prospective collection contained other forms of gynecologic cancers, beyond EOC, and a selection of these, randomized along with sera from a small number of apparently healthy male and female donors, were also analyzed in blinded manner. In most instances the number of each disease was low, introducing wide confidence intervals into the performance estimates (Table 5). Although the confidence intervals are wide, the results are interesting. The sensitivity for mixed cancers in which ovarian cancer is present with another cancer was 100%.

Table 3. Sensitivity of the 8-Biomarker Profile by Epithelial Ovarian Cancer FIGO Stage

FIGO Stage	SN	95% CI	n/N
I	47.1%	23.0 – 72.2	8/17
II	60.0%	26.2 – 87.8	6/10
III	91.2%	81.8 – 96.7	62/68
IV	100.0%	54.1 – 100.0	6/6
Unknown	80.0%	44.4 – 97.5	8/10

Abbreviations: FIGO, International Federation of Gynecology and Obstetrics; SN, sensitivity; n, number of samples classified as ovarian cancer; N, number of samples assayed.

The fallopian and peritoneal cancers, which are very similar in origin, treatment and outcome to EOC, were detected with similar sensitivity (~81%) to EOC. The malignant mixed Mullerian tumors (MMMT) were classified as a malignancy in 85.7% of cases, a promising result given CA-125 is not useful in monitoring most MMMT. Finally, CA-125 sensitivity for endometrial/uterine cancers is low making the 53.9% detection with the 8-biomarker profile interesting. Together these results suggest that the 8-biomarker profile is equally informative for other ovarian-associated malignancies. As anticipated, 100% of male specimens were classified as having a profile inconsistent with the presence of ovarian cancer. Only 2 of 45 apparently healthy women tested had profiles consistent with ovarian cancer (specificity = 95.6%; 95% CI 83.6 – 99.2%), a significantly higher specificity than was observed for women with benign ovarian conditions (85.4%). This finding underscores the necessity of validating classification profiles on case and control samples matching the correct intended use population. Despite the apparently higher specificity on nominally healthy women, the 8-biomarker profile specificity still reflects an unacceptably high false positive rate for a screening application within the general population.[34]

CONCLUSION

In conclusion, using a prospective collection from women assessed as at risk of having ovarian cancer, under the care of a gynecologic oncologist, we have validated an 8-biomarker profile built on a retrospective collection of stage I invasive EOC and demonstrated it meets the general requirements of an informative test. We are now focusing on validating an improvement of the KDE-VS algorithm that allows us to adjust the relative sensitivity and specificity of the profile, with the gain in one performance attribute offset by an acceptable decrease in the other performance attribute. At the same time we are undertaking a second prospective collection, this time focusing on women with symptoms of ovarian cancer currently under the care of a non-specialist, to determine the validated performance of our assay on the intended use population – those women who would benefit by referral to a specialist oncologist.

Table 4. Sensitivity of the 8-Biomarker Profile by EOC Subtype

	Multiple Subtypes			Single Subtypes		
Subtype‡	SN	95% CI	n/N	SN	95% CI	n/N
Serous	81.4%	71.6 – 89.0	70/86	85.5%	75.0 – 92.8	59/69
Endometrioid	85.7%	63.7 – 97.0	18/21	90.0%	55.5 – 99.8	9/10
Clear Cell	78.6%	49.2 – 95.3	11/14	80.0%	28.4 – 99.5	4/5
Mucinous	50.0%	11.8 – 88.2	3/6	66.7%	9.4 – 99.2	2/3
Borderline	0.0%	0.0 – 84.2	0/2	-	-	-
Undifferentiated	100.0%	2.5 – 100.0	1/1	100.0%	2.5 – 100.0	1/1
Transitional Cell	50.0%	1.3 – 98.7	1/2	100.0%	2.5 – 100.0	1/1
Brenner	66.7%	9.4 – 99.2	2/3	100.0%	15.8 – 100.0	2/2
Mixed Subtypes	-	-	-	60.0%	36.1 – 80.9	12/20

‡ More than one subtype of epithelial ovarian cancer was reported for 20 individuals. The column "Multiple" reports the classification for all reported tumors either as single subtypes or mixed subtypes. The column "Single" reports the classification of tumors assigned only a single subtype. The 20 mixed subtypes are grouped together and their pathologies are as follows: serous/endometrioid (7), serous/clear cell (4), serous/borderline (2), serous/mucinous (1), serous/endometrioid/clear cell (2), serous/endometrioid/clear cell/mucinous (1), clear cell/mucinous (1), clear cell/endometrioid (1), transitional cell/Brenner (1). Abbreviations: SN, sensitivity; SP, specificity; n, number of samples classified correctly; N, total number of samples being classified; 95% CI, 95% two-sided exact binomial confidence interval.

Table 5. Classification of Other Gynecologic Cancers by the 8-Biomarker Profile

Cancer	Classified as Ovarian Cancer	95% CI	n/N
Cervical	100.0%	19.8 – 100.0	2/2
Endometrial/uterine	53.8%	26.1 – 79.6	7/13
Fallopian	80.0%	29.9 – 98.9	4/5
Peritoneal	81.8%	47.8 – 96.8	9/11
Extra-ovarian	100.0%	5.5 – 100.0	1/1
Ovarian germ cell	100.0%	5.5 – 100.0	1/1
Ovarian stromal cell	100.0%	5.5 – 100.0	1/1
Ovarian transitional cell	100.0%	5.5 – 100.0	1/1
Ovarian borderline (Low Malignant Potential)	54.8%	38.8 – 69.8	23/42
Ovarian malignant mixed Mullerian tumor	85.7%	42.0 – 99.2	6/7
Mixed Cancers	100.0%	46.3 – 100.0	5/5

Abbreviations: n, number of samples classified as ovarian cancer; N, number of samples assayed; 95% CI, 95% two-sided exact binomial confidence interval.

ACKNOWLEDGEMENTS

We thank John L. Lovecchio, MD; North Shore – Long Island Jewish Health System, Manhasset, NY.

REFERENCES

1. American Cancer Society. Cancer Facts & Figures 2010. Atlanta: American Cancer Society; 2010.
2. Horner, M.J., *et al.* (eds.). *SEER Cancer Statistics Review, 1975-2006*, (National Cancer Institute, Bethesda, MD, 2009).
3. Wakabayashi, M.T., Lin, P.S. & Hakim, A.A. The Role of Cyto-reductive/De-bulking Surgery in Ovarian Cancer CME. *J Natl Compr Canc Netw* **6**, 803-811 (2008).
4. Earle, C.C., *et al.* Effect of surgeon specialty on processes of care and outcomes for ovarian cancer patients. *J Natl Cancer Inst* **98**, 172-180 (2006).
5. Paulsen, T., Kjaerheim, K., Kaern, J., Tretli, S. & Trope, C. Improved short-term survival for advanced ovarian, tubal, and peritoneal cancer patients operated at teaching hospitals. *Int J Gynecol Cancer* **16 Suppl 1**, 11-17 (2006).
6. Carney, M.E., Lancaster, J.M., Ford, C., Tsodikov, A. & Wiggins, C.L. A population-based study of patterns of care for ovarian cancer: who is seen by a gynecologic oncologist and who is not? *Gynecol Oncol* **84**, 36-42 (2002).
7. NCCN. NCCN Clinical Practice Guidelines in Oncology - Ovarian Cancer. Version I (2008).
8. Myers, E.R., *et al.* Management of Adnexal Mass. Vol. Publication No. 06-E004 (Agency for Healthcare Research and Quality, 2006).
9. Timmerman, D., *et al.* Logistic regression model to distinguish between the benign and malignant adnexal mass before surgery: a multicenter study by the International Ovarian Tumor Analysis Group. *J Clin Oncol* **23**, 8794-8801 (2005).
10. Van Calster, B., *et al.* Discrimination Between Benign and Malignant Adnexal Masses by Specialist Ultrasound Examination Versus Serum CA-125. *J Natl Cancer Inst* **99**, 1706-1714 (2007).
11. Buamah, P. Benign conditions associated with raised serum CA-125 concentration. *J Surg Oncol* **75**, 264-265 (2000).
12. Kalantri, Y., *et al.* Role of cancer antigen-125 from pleural & ascitic fluid samples in non malignant conditions. *Indian J Med Res* **125**, 25-30 (2007).
13. Miralles, C., *et al.* Cancer antigen 125 associated with multiple benign and malignant pathologies. *Ann Surg Oncol* **10**, 150-154 (2003).
14. Greene, M.H., *et al.* A Prospective Study of Risk-Reducing Salpingooophorectomy and Longitudinal CA-125 Screening among Women at Increased Genetic Risk of Ovarian Cancer: Design and Baseline Characteristics: A Gynecologic Oncology Group Study. *Cancer Epidemiol Biomarkers Prev* **17**, 594-604 (2008).
15. Skates, S.J., *et al.* Calculation of the risk of ovarian cancer from serial CA-125 values for preclinical detection in postmenopausal women. *J Clin Oncol* **21**, 206s-210s (2003).
16. Skates, S.J., *et al.* Preoperative sensitivity and specificity for early-stage ovarian cancer when combining cancer antigen CA-125II, CA 15-3, CA 72-4, and macrophage colony-stimulating factor using mixtures of multivariate normal distributions. *J Clin Oncol* **22**, 4059-4066 (2004).
17. Moore, R.G., *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* **112**, 40-46 (2009).
18. Gostout, B.S. & Brewer, M.A. Guidelines for referral of the patient with an adnexal mass. *Clin Obstet Gynecol* **49**, 448-458 (2006).
19. Im, S.S., *et al.* Validation of referral guidelines for women with pelvic masses. *Obstet Gynecol* **105**, 35-41 (2005).
20. American College of Obstetricians and Gynecologists. The role of the generalist obstetrician–gynecologist in the early detection of ovarian cancer. *Obstet Gynecol* **100**, 1413-1416 (2002).
21. Dearking, A.C., *et al.* How Relevant Are ACOG and SGO Guidelines for Referral of Adnexal Mass? *Obstet Gynecol* **110**, 841-848 (2007).

22. Andersen, M.R., *et al.* Combining a symptoms index with CA 125 to improve detection of ovarian cancer. *Cancer* **113**, 484-489 (2008).
23. Guerriero, S., *et al.* Trans-vaginal Ultra-sonography in the Diagnosis of Extra-uterine Pelvic Diseases. *Expert Rev of Obstet Gynecol* **3**, 731-752 (2008).
24. Timmerman, D., *et al.* Inclusion of CA-125 does not improve mathematical models developed to distinguish between benign and malignant adnexal tumors. *J Clin Oncol* **25**, 4194-4200 (2007).
25. Visintin, I., *et al.* Diagnostic Markers for Early Detection of Ovarian Cancer. *Clin Cancer Res* **14**, 1065-1072 (2008).
26. FDA/Vermillion. Substantial Equivalence Decision k081754 - OVA1. (2009).
27. Amonkar, S.D., *et al.* Development and preliminary evaluation of a multivariate index assay for ovarian cancer. *PLoS ONE* **4**, e4599 (2009).
28. Bertenshaw, G.P., *et al.* Multianalyte profiling of serum antigens and autoimmune and infectious disease molecules to identify biomarkers dysregulated in epithelial ovarian cancer. *Cancer Epidemiol Biomarkers Prev* **17**, 2872-2881 (2008).
29. Wians Jr, F.H. Clinical Laboratory Tests: Which, Why, and What Do The Results Mean? *LabMedicine* **40**, 105-113 (2009).
30. Sherman, M.E., *et al.* Survival among women with borderline ovarian tumors and ovarian carcinoma: a population-based analysis. *Cancer* **100**, 1045-1052 (2004).
31. Menon, U., *et al.* Sensitivity and specificity of multimodal and ultrasound screening for ovarian cancer, and stage distribution of detected cancers: results of the prevalence screen of the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS). *Lancet Oncol* **10**, 327-340 (2009).
32. van den Akker, P.A., *et al.* Evaluation of the risk of malignancy index in daily clinical management of adnexal masses. *Gynecol Oncol* **116**, 384-388 (2010).
33. Timor-Tritsch, I.E. Adnexal Mass. in *Ultrasound in Gynecology* (eds. Timor-Tritsch, I.E. & Goldstein, S.R.) 100-116 (Churchill Livingstone, 2007).
34. Havrilesky, L.J., Sanders, G.D., Kulasingam, S. & Myers, E.R. Reducing ovarian cancer mortality through screening: Is it possible, and can we afford it? *Gynecol Oncol* **111**, 179-187 (2008).