



Article

An Improved Prediction Model for Ovarian Cancer Using Urinary Biomarkers and a Novel Validation Strategy

Shin-Wha Lee ¹, Ha-Young Lee ², Hyo Joo Bang ³, Hye-Jeong Song ⁴, Sek Won Kong ⁵ and Yong-Man Kim ^{1,*}

¹ Department of Obstetrics & Gynecology, University of Ulsan, ASAN Medical Center, Seoul 05505, Korea; swhlee@amc.seoul.kr

² ASAN Institute for Life Science, ASAN Medical Center, Seoul 05505, Korea; leehayoung@gmail.com

³ Ahngook Pharmaceutical Co., Ltd., Seoul 07445, Korea; jwnme@ahn-gook.com

⁴ Bio-IT Research Center, Hallym University, Chuncheon, Gangwon-do 24252, Korea; hjsong@hallym.ac.kr

⁵ Computational Health Informatics Program, Boston Children's Hospital, Boston, MA 02115, USA; sekwon.kong@childrens.harvard.edu

* Correspondence: ymkim@amc.seoul.kr; Tel.: +82-2-3010-3640

Received: 9 August 2019; Accepted: 29 September 2019; Published: 5 October 2019

Abstract: This study was designed to analyze urinary proteins associated with ovarian cancer (OC) and investigate the potential urinary biomarker panel to predict malignancy in women with pelvic masses. We analyzed 23 biomarkers in urine samples obtained from 295 patients with pelvic masses scheduled for surgery. The concentration of urinary biomarkers was quantitatively assessed by the xMAP bead-based multiplexed immunoassay. To identify the performance of each biomarker in predicting cancer over benign tumors, we used a repeated leave-group-out cross-validation strategy. The prediction models using multimarkers were evaluated to develop a urinary ovarian cancer panel. After the exclusion of 12 borderline tumors, the urinary concentration of 17 biomarkers exhibited significant differences between 158 OCs and 125 benign tumors. Human epididymis protein 4 (HE4), vascular cell adhesion molecule (VCAM), and transthyretin (TTR) were the top three biomarkers representing a higher concentration in OC. HE4 demonstrated the highest performance in all samples with OC (mean area under the receiver operating characteristic curve (AUC) 0.822, 95% CI: 0.772–0.869), whereas TTR showed the highest efficacy in early-stage OC (AUC 0.789, 95% CI: 0.714–0.856). Overall, HE4 was the most informative biomarker, followed by creatinine, carcinoembryonic antigen (CEA), neural cell adhesion molecule (NCAM), and TTR using the least absolute shrinkage and selection operator (LASSO) regression models. A multimarker panel consisting of HE4, creatinine, CEA, and TTR presented the best performance with 93.7% sensitivity (SN) at 70.6% specificity (SP) to predict OC over the benign tumor. This panel performed well regardless of disease status and demonstrated an improved performance by including menopausal status. In conclusion, the urinary biomarker panel with HE4, creatinine, CEA, and TTR provided promising efficacy in predicting OC over benign tumors in women with pelvic masses. It was also a non-invasive and easily available diagnostic tool.

Keywords: ovarian cancer; prediction model; urinary biomarker

1. Introduction

Ovarian cancer (OC) is the seventh most common malignancy among women worldwide, accounting for 3.6% of cancers in women [1]. It is the most lethal of all gynecological malignancies, with more than 70% of women presenting with the advanced-stage disease [2], and the primary treatment for advanced-stage disease involves both cytoreductive surgery and chemotherapy.

Despite aggressive treatment, long-term survival has not significantly improved in the past 30 years, with the five-year survival rate remaining between 30% and 45% [3]. Pelvic masses are often a diagnostic challenge for physicians, as discriminating malignant from benign masses as early intervention is crucial for prognosis. Patients with advanced-stage OC have a 20–40% five-year survival rate, in contrast to a >90% five-year survival rate for patients identified with and treated for stage-I OC [4]. The management of OC with complete cytoreductive surgery by a gynecologic oncologist enables a favorable prognosis [5]; therefore, the decision of an appropriate referral is important. Unfortunately, it has been reported that 65% of patients with malignant adnexal masses are inappropriately referred to gynecologic oncologists in the United States of America [6].

Cancer antigen 125 (CA-125) is the most widely used biomarker in OC; however, it lacks adequate diagnostic sensitivity (SN) and specificity (SP) to be used reliably as a screening test. Recently, in a randomized controlled trial, annual multimodal screening with serum CA-125 and transvaginal ultrasound sonography as a second-line test revealed limited efficacy when prevalent cases were excluded [7]. The Risk of Malignancy Index (RMI) was developed to assess the malignancy risk in women with adnexal masses, which incorporates a woman's CA-125 level, ultrasound morphology score, and menopausal status [8]. Two algorithms have been developed for the same reason. The Risk of Malignancy Algorithm (ROMA) is based on serum levels of human epididymis protein 4 (HE4) and CA-125 according to menopausal status, and OVA1 includes five serum biomarkers: CA-125, β -2 microglobulin, transferrin, transthyretin (TTR), and apolipoprotein [9–11]. Various studies have been published evaluating the effectiveness of RMI, ROMA, and OVA1 tests, as well as other strategies to distinguish the risk of malignancy in the setting of a pelvic mass. These are not screening tests for women in certain age groups, but could be used as decision-supporting tools for surgical intervention. However, to date, definitive conclusions have not been drawn regarding these tests due to a lack of prospective multi-institutional trials and cost-benefit analyses. Furthermore, the low compliance of patients for transvaginal ultrasound sonography and blood tests is another unresolved issue.

For these reasons, discovering biomarkers in urine can provide a non-invasive method for detecting OC, and enable the frequent testing of women who belong to high-risk groups. The protein profile in the urine is also less complex than that in the blood, and thus the measurement of urinary proteins may also enhance overall clinical performance. Several biomarkers have been identified in the urine of women with OC including eosinophil-derived neurotoxin, a fragment of osteopontin, mesothelin, and Bcl-2 [12–14]. The level of HE4 in the urine has been reported to have a relationship with primary platinum resistance that is not evident in serum levels [15]. However, there is a limit to the clinical applications of urinary proteins for patients with OC, and published studies have only been presented for a single biomarker with a simple statistical association. Here, we report an extended analysis of 23 protein biomarkers in the urine associated with OC and a panel of urinary biomarkers that demonstrate high performance in distinguishing OC from benign ovarian tumors in patients with several types of pelvic mass.

2. Results

2.1. Urinary Protein Biomarkers in Patients with Ovarian Cancer

Figure 1 shows the modeling process used to find the optimum biomarkers and the best combination. Twelve patients who were diagnosed with borderline tumors after surgery were excluded from the analysis. After exclusion, 283 urine samples were analyzed. The mean age of patients was different in the benign group and the cancer group (mean $\pm$ standard deviation: 40.6 $\pm$ 12.2 vs 53.8 $\pm$ 10.4, respectively). Among the 158 patients with OC, 70.3% exhibited papillary serous adenocarcinoma, 69.6% had advanced-stage cancer (International Federation of Gynecology and Obstetrics (FIGO) stage III and IV), and 58.2% exhibited postmenopausal status (Table 1). The urinary concentration of 17 biomarkers showed significant differences between OC and benign tumors (corrected $p < 0.05$). The concentrations of HE4, vascular cell adhesion molecule (VCAM), TTR, macrophage migration inhibitory factor (MIF), leptin, C-reactive protein (CRP),

platelet-derived growth factor (PDGF)-AA, Cyfra21-1, neural cell adhesion molecule (NCAM), prolactin, myeloperoxidase (MPO), Mesomark, CA19-9, and apolipoprotein A1 (ApoA1) were significantly higher in the urine of patients with OC than in those with benign tumors, whereas the concentrations of carcinoembryonic antigen (CEA), creatinine and plasminogen activator inhibitor-1 (PAI-1) were significantly lower in patients with OC than in those with benign tumors. HE4, VCAM, and TTR were the top three altered biomarkers in this analysis and, among them, HE4 exhibited a significant difference with a maximum p -value of 1.26×10^{-24} . CA-125 is a traditional serum tumor biomarker in OC, but it was not significant in our urine data (Table S1).

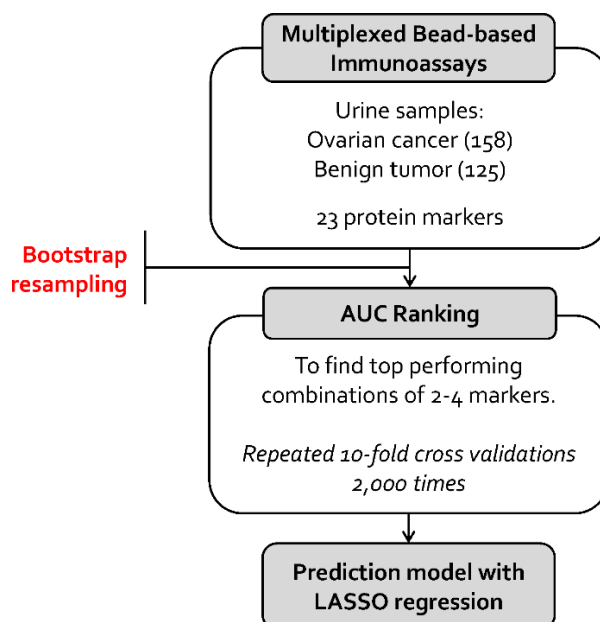


Figure 1. Process of the differential diagnosis modeling between ovarian cancer (OC) and benign tumor in this study. AUC, area under the receiver operating characteristic curve; LASSO, Least Absolute Shrinkage and Selection Operator.

Table 1. Clinical and demographic characteristics of the study patients.

	Number of Totals	Premenopausal	Postmenopausal
Total	295	175	119
Age (years), median and range	48 (20–82)		
Benign tumor	125 (42.4%)	99	25
Endometriosis	44	41	3
Teratoma	30 *	27	2
Mucinous cystadenoma	16	13	3
Serous cystadenoma	9	5	4
Inflammation ¹	5	1	4
Others ²	21	12	9
Borderline tumor	12 (4.1%)	10	2
Malignant tumor	158 (53.5%)	66	92
Serous adenocarcinoma	111	40	71
Mucinous adenocarcinoma	12	6	6
Endometrioid adenocarcinoma	15	11	4
Clear cell carcinoma	12	4	8
Other EOC ³	3	1	2
Granulosa cell tumor	2	2	-
Dysgerminoma	1	1	-
Other non-EOC ⁴	2	1	1
FIGO stage of malignancy			
I	36	23	13
II	12	5	7

III	91	32	59
IV	19	6	13

EOC, epithelial ovarian cancer; FIGO, the International Federation of Gynecology and Obstetrics. * The menopausal status of one patient was not stated; ¹ three tubo-ovarian abscesses/two chronic granulomatous inflammations; ² eight paratubal cysts, six hemorrhagic cysts, five fibromas, and two leiomyomas; ³ two transitional cell carcinomas and one mixed cell type (endometrioid and serous adenocarcinoma); ⁴ two mixed germ cell tumors.

Next, we tested whether the urine concentration of each protein marker could classify OC from benign tumors. We used a bootstrap resampling method to calculate the mean area under the receiver operating characteristic (AUC) from 2000 randomly chosen subsets while maintaining the OC-to-benign tumor ratios, and 95% confidence intervals (CIs) were estimated using DeLong's method. Table 2 shows the performance of each single biomarker in urine samples for predicting cancer over benign masses. HE4 exhibited the greatest AUC values in total samples of OC followed by VCAM, leptin, TTR, and prolactin. Interestingly, TTR showed the highest performance in early-stage cancers, whereas HE4 demonstrated the highest value in advanced-stage cancers.

Table 2. The performance of each single biomarker in the urine samples to predict cancer over benign tumors.

Markers	All Samples		Stage I and II		Stage III and IV	
	AUC	(95% CIs)	AUC	(95% CIs)	AUC	(95% CIs)
HE4	0.822	(0.772–0.869)	0.759	(0.659–0.852)	0.847	(0.791–0.898)
VCAM	0.776	(0.717–0.829)	0.744	(0.650–0.830)	0.788	(0.725–0.847)
Leptin	0.771	(0.701–0.837)	0.779	(0.657–0.889)	0.772	(0.689–0.848)
TTR	0.767	(0.706–0.824)	0.789	(0.714–0.856)	0.757	(0.693–0.819)
Prolactin	0.713	(0.624–0.793)	0.732	(0.573–0.870)	0.701	(0.600–0.794)
CRP	0.710	(0.644–0.772)	0.644	(0.541–0.743)	0.734	(0.662–0.800)
PDGF-AA	0.697	(0.632–0.758)	0.734	(0.644–0.820)	0.677	(0.607–0.747)
NCAM	0.678	(0.613–0.741)	0.672	(0.576–0.761)	0.678	(0.609–0.742)
Mesomark	0.670	(0.578–0.756)	0.648	(0.520–0.769)	0.680	(0.583–0.769)
MPO	0.668	(0.598–0.737)	0.640	(0.554–0.723)	0.684	(0.610–0.756)
Cyfra21-1	0.660	(0.591–0.726)	0.728	(0.630–0.821)	0.628	(0.549–0.701)
CEA	0.627	(0.558–0.692)	0.684	(0.583–0.778)	0.600	(0.524–0.676)
Creatinine	0.622	(0.554–0.687)	0.678	(0.585–0.770)	0.596	(0.518–0.668)
CA19-9	0.598	(0.529–0.666)	0.578	(0.470–0.678)	0.604	(0.528–0.677)
IL6	0.576	(0.450–0.701)	0.490	(0.323–0.657)	0.599	(0.459–0.730)
MIF	0.572	(0.500–0.642)	0.640	(0.531–0.743)	0.537	(0.456–0.618)
ApoAI	0.557	(0.485–0.623)	0.517	(0.422–0.612)	0.569	(0.493–0.644)
ApoCIII	0.524	(0.448–0.600)	0.580	(0.466–0.687)	0.562	(0.479–0.644)
PAI-1	0.523	(0.446–0.598)	0.514	(0.417–0.616)	0.533	(0.451–0.611)
CA125	0.523	(0.453–0.591)	0.486	(0.370–0.597)	0.533	(0.456–0.611)
OPN	0.521	(0.45–0.591)	0.540	(0.443–0.635)	0.514	(0.435–0.594)
IL8	0.488	(0.416–0.563)	0.541	(0.449–0.633)	0.531	(0.453–0.607)
CA15-3	0.486	(0.417–0.552)	0.540	(0.439–0.638)	0.536	(0.460–0.613)

AUC, area under the receiver operating characteristic curve; CIs, confidence intervals. AUC for each biomarker to predict cancer over benign tumors. The 95% confidence intervals were calculated using a bootstrap resampling method. We repeated resampling 2000 times with a replacement in proportion to the cancer-to-benign ratio of all samples. HE4, human epididymis protein 4; VCAM, vascular cell adhesion molecule; TTR, transthyretin; CEA, carcinoembryonic antigen; NCAM, neural cell adhesion molecule; CA-125, cancer antigen 125; CRP, C-reactive protein; PDGF, platelet-derived growth factor; MPO, myeloperoxidase; IL, interleukin; MIF, macrophage migration inhibitory factor; ApoAI, apolipoprotein A1; ApoCIII, apolipoprotein C3; PAI-1, plasminogen activator inhibitor-1; OPN, osteopontin.

2.2. Urinary Multimarker Panel Analysis in Patients with Ovarian Cancer

To build and compare the performance of multimarker prediction models, we used the Least Absolute Shrinkage and Selection Operator (LASSO) regression model to prioritize each marker. Using all biomarkers in the current study, we built LASSO models to predict OC over benign tumors with cross-validations as implemented in the R library glmnet. For each iteration of cross-validation, the regularization parameter lambda was determined. The magnitude of lambda for each marker was summarized from 2000 iterations of cross-validation to rank the importance of each biomarker. HE4 is the most influential biomarker in the differential diagnosis between OC and benign tumors, followed by creatinine, CEA, NCAM, and TTR (Figure 2). Interestingly, HE4 demonstrated not only the maximum t-score but also the greatest coefficient in the logistic regression model. However, the subsequent t-score biomarker VCAM revealed a small coefficient. Several biomarkers (e.g., NCAM) had a wide variation in the distribution coefficients. Notably, this coefficient result was closely related to the ROC curve of the next panel analysis.

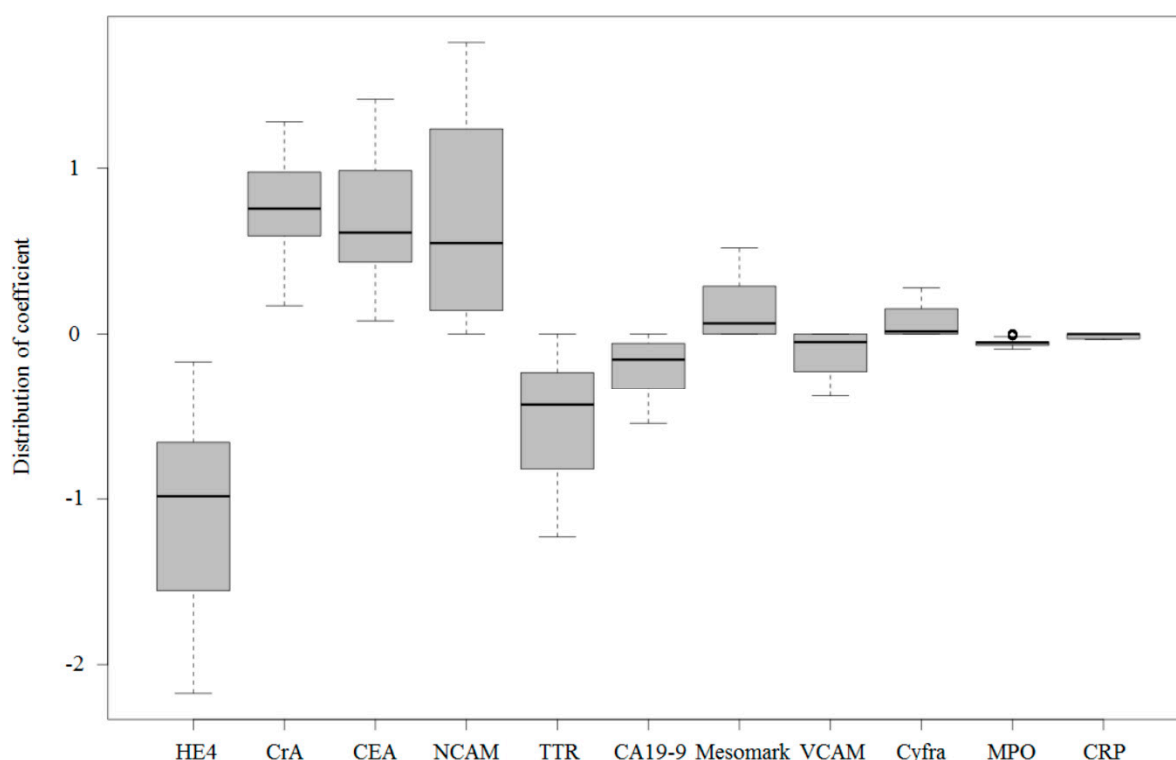


Figure 2. Distribution of the coefficient by the LASSO regression method. The y-axis is the distribution coefficient of the logistic regression model analyzed by the LASSO regression method. There are multiple values because the method was repeated 10-fold, with 2000 iterations of cross-validation. The mean value is in descending order from the left. As the absolute value is large, it will have a large coefficient from the model. In this sense, when the standard deviation of the measured value is changed to 1, it affects the results of the entire model. HE4 is the greatest contributor, followed by creatinine, CEA, NCAM, and TTR. There is large variation in the case of NCAM, and VCAM does not have a large contribution.

When the bootstrap resampling method was applied to the multivariate analysis, the top-performing two-biomarker combinations identified each included HE4 and creatinine. This combination outperformed the best individual biomarkers in terms of SN, SP, and AUC values. The top-performing three-biomarker panel included combinations of HE4, creatinine, and CEA. This panel provided notably elevated AUC values in comparison to the two-biomarker combinations and individual biomarkers, and it was modestly increased after multimarker combinations. Our final model included HE4, creatinine, CEA, and TTR; this panel provided 93.7% SN and a 78.7% positive predictive value (PPV) as well as 90.6% negative predictive value (NPV) for all-stage OC at 70.6% SP. When adding the menopausal status as an independent variable to the selected panel, it was confirmed that the AUC value increased, even if at very weak levels. (Figure 3 and Table 3).

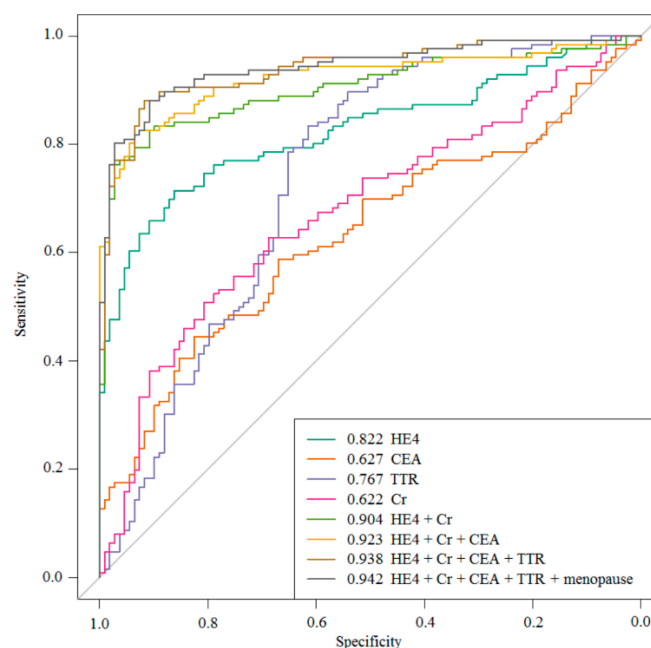


Figure 3. The ROC curve of the multimarker panels (i.e., HE4, creatinine, CEA, and TTR), and the results with the addition of menopausal status.

Table 3. The performance of multimarker models to predict cancer over benign tissues between early- and advanced-stage cancer (this analysis was performed using 109 benign and 131 cancer samples).

Markers	All Samples		Stage I and II		Stage III and IV	
	AUC	(95% CIs)	AUC	(95% CIs)	AUC	(95% CIs)
HE4	0.822	(0.772–0.869)	0.759	(0.659–0.852)	0.847	(0.791–0.898)
CEA	0.627	(0.558–0.692)	0.684	(0.583–0.778)	0.600	(0.524–0.676)
TTR	0.767	(0.706–0.824)	0.789	(0.714–0.856)	0.757	(0.693–0.819)
Creatinine (Cr)	0.622	(0.554–0.687)	0.678	(0.585–0.770)	0.596	(0.518–0.668)
HE4+Cr	0.904	(0.854–0.938)	0.825	(0.740–0.910)	0.926	(0.888–0.972)
HE4+Cr+CEA	0.923	(0.878–0.954)	0.883	(0.772–0.934)	0.943	(0.907–0.982)
HE4+Cr+CEA+TTR	0.938	(0.900–0.964)	0.932	(0.844–0.970)	0.946	(0.911–0.983)

AUC, area under the receiver operating characteristic curve; CIs, confidence intervals. AUC for each biomarker to predict cancer over benign tumors. The 95% confidence intervals were calculated using DeLong's method. We repeated resampling 2000 times with replacement in proportion to the cancer to benign ratio of all samples.

Each of the biomarker panels identified in this study was further evaluated for the disease status in patients with early-stage cancer and advanced-stage cancer. The four-biomarker panel with HE4, creatinine, CEA, and TTR was found to be the most useful for the classification of cancer compared with that of benign tumors regardless of disease status. In the early-stage group, in particular, the performance of multimarker models used to predict cancer over benign tumors improved remarkably compared with that in the advanced-stage group. When adding menopausal status as an independent variable to the selected panel, the AUC value increased regardless of the FIGO stage (AUC: 0.932 vs 0.933 in the early-stage group, 0.946 vs 0.951 in the advanced-stage group) (Figure 4).

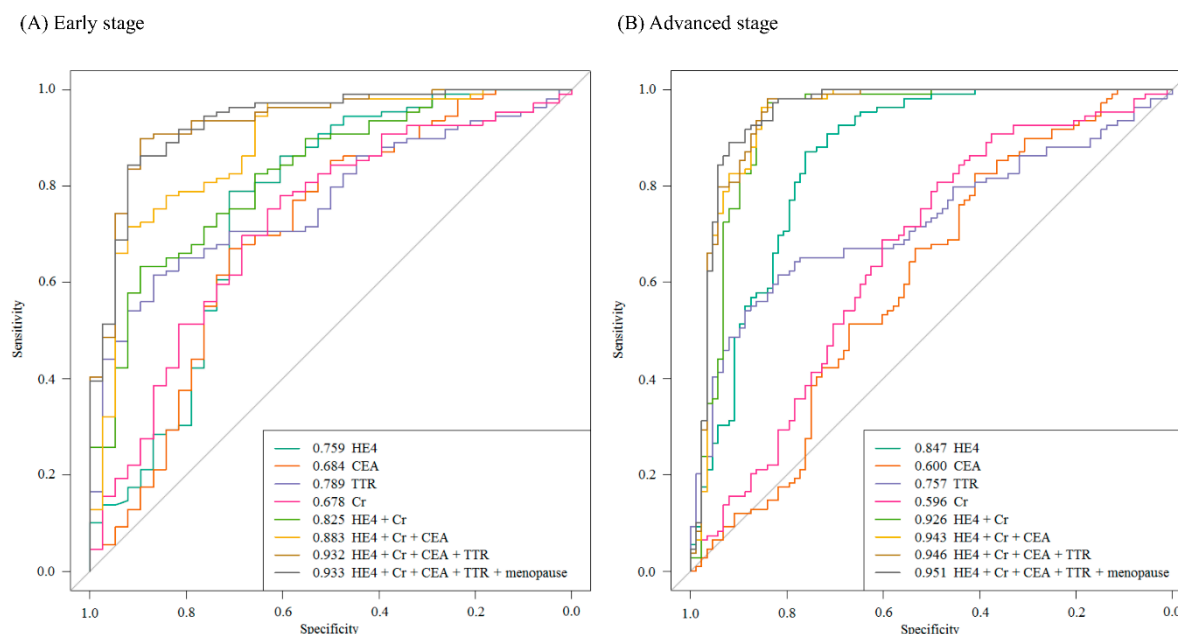


Figure 4. The ROC curve of multimarker panels (i.e., HE4, creatinine, CEA, and TTR) according to the FIGO stage, and the results from the addition of menopausal status.

2.3. Clinical Characteristics of Difficult Samples to Predict

We checked whether the samples that were not correctly predicted by our multimarker panel shared any clinical characteristics. Among 15 patients with OC representing a false-negative in our result, seven patients (46.7%) exhibited early-stage disease, including two patients with mucinous intraepithelial cancer, one with malignant granulosa cell tumor, and one with dysgerminoma. Among eight patients with advanced-stage disease, six patients revealed low-grade serous adenocarcinoma, and the remaining two patients had clear cell carcinoma and endometrioid adenocarcinoma. Our multimarker panel did not fail to identify the high-grade serous OC that accounts for most OCs, even in early-stage disease. The clinical characteristics of patients representing a false-negative in this study were summarized as nonepithelial OCs and early-stage non-serous OCs. In contrast, looking at two patients with a false-positive, we could confirm necrosis in the surgical specimens caused by the torsion of teratoma or ischemia of a hemorrhagic cyst. Therefore, rare ovarian malignancies and necrotic ischemic benign lesions should be considered if the results of biomarkers do not match the clinical judgment.

3. Discussion

A multimarker panel appears to serve an essential role in the differential diagnosis of OC in women who have pelvic masses [16–20]. In particular, many investigators are currently performing biomarker research using urine, as urine samples can be acquired truly non-invasively (as opposed to blood samples) [15,21–24]. The goal of this study was to determine a panel of biomarkers that exhibit high SN and SP in order to distinguish OC from benign ovarian tumors in patients with the heterogeneous characteristics of tumors at different growth stages using bead-based xMAP multiplexing technology and truly non-invasive urine samples. Firstly, in agreement with published evidence [15,22], HE4 in urine was the most important biomarker for ovarian neoplasm. However, although HE4 was a strong biomarker for predicting the overall stages of OC (particularly early stage), TTR provided better performance than HE4. This is a good reason to develop a multimarker panel for the early detection and differential diagnosis of OC.

After performing a multimarker multiplexed immunoassay, we investigated (in line with reported evidence [16]) that four-biomarker panels seem to offer superior performance compared with two- and three-biomarker panels. Through repeating the 10-fold cross-validation 2000 times,

we identified a combination of the top four urinary protein biomarkers (i.e., HE4, creatinine, CEA, and TTR) that provided 94% SN, 79% PPV, and 91% NPV for all-stage OC with 71% SP. SP in the final analysis was relatively low, which is why the objective of this study was a biomarker for differential diagnosis and not for screening. For the screening test, a high SN should be achieved for distinguishing women with early-stage OC from healthy individuals with at least 98% SP. Due to the low prevalence of OC (accounting for 3.6% of cancers in women worldwide) [1], the required minimum SP of the serum test should be 98% when combined with transvaginal sonography (TVS) as a second-line test [25,26], and accurate assessment of such high SP require at least 1000 healthy participants. However, the present results are comparable to other differential diagnostic techniques using blood samples: (1) the urinary protein panel (our results) offered 93% SN at 75% SP, (2) the RMI showed 85–96% SN at 75% SP [27,28], (3) the ROMA provided 94% SN at 75% SP [27–29], and (4) the OVA1 test revealed 92–97% SN at a low SP of approximately 54% [29,30] (Table 4).

Table 4. Specificity and sensitivity results of various screening strategies in the setting of a pelvic mass.

Algorithm or Assay		N	SN (%)	SP (%)	PPV	NPV	AUC
<i>Serum Multimarker Panels</i>							
Moore, 2010 [27]	RMI	457	84.6	75.0	-	-	0.870
	ROMA	457	94.3	75.0	-	-	0.953
Karlsen, 2012 [28]	RMI	1218	96.0	75.0	-	-	0.958
	ROMA	1218	94.8	75.0	-	-	0.954
Bristow, 2013 [30]	OVA1	494	92.4	53.5	31.3	96.8	-
Grenache, 2015 [29]	ROMA	146	83.9	83.5	57.8	95.1	-
	OVA1	146	96.8	54.8	36.6	98.4	-
<i>Urinary Biomarker</i>							
Hellstrom, 2010 [22]	HE4	135 ¹	88.6	94.4			
Liao, 2015 [15]	HE4	279 ²	52.2	95.0			
<i>Urinary Multimarker Panel in this Study</i>							
HE4+Cr+CEA+TTR		283	81.0	95.3	95.3	81.3	0.938
			92.9	75.2	81.3	90.1	

SN, sensitivity; SP, specificity; PPV, positive predictive value; NPV, negative predictive value; AUC, area under the receiver operating characteristic curve; ¹ including 79 patients with ovarian cancer; ² including 92 patients with ovarian cancer; RMI, Risk of Malignancy Index; ROMA, Risk of Malignancy Algorithm.

Although the performance of the differential diagnosis of OC was similar to that of other diagnostic panels approved by the US Food and Drug Administration (FDA), our urinary panel had distinct advantages as a diagnostic test. This was because urine is an excellent alternative liquid biopsy for biomarker research, as it is available in larger quantities compared to blood and through a less invasive process, allowing for repeated tests. The identification of urinary biomarkers could provide a more convenient and accurate test for diagnosing malignant tumors. The urinary proteome is a direct product of renal filtration and consists of soluble, low molecular weight peptides that are highly appropriate for proteomic analyses that may represent disease-specific cleavages. Renal filtration also results in a less complex matrix compared to the serum and contains fewer factors known to interfere with biomarker assays, which can be an advantage in a proteomic analysis [12,13,24].

There were some limitations to this study, including the small sample size and 23 selected protein biomarkers. In the current study, the biggest limitation was the absence of an independent validation set. However, we improved the significance of our data by performing a resampling method and multivariate statistical analyses, which were applied to an in vitro diagnostic multivariate index assay (IVDMIA) of proteomic biomarkers. The most important method was to create random sampling and maintain independence for the logistic regression analysis by a bootstrap resampling method. Bootstrapping is the practice of estimating properties of an estimator (e.g., variance) by measuring properties when sampling from an approximate distribution, and can

refer to any test that relies on random sampling [31–33]. Among algorithms for comparing the areas under two or more correlated ROC curves, DeLong's algorithm is the most widely used due to its simplicity of implementation in practice [17,34]. In this way, we were able to maximize the independence of the hold-out testing set of samples [17]. The second method was the LASSO regression. Variable selection in the regression analysis is a key procedure when we have a large collection of possible covariates for the efficient prediction of a response variable. LASSO not only helps to improve prediction accuracy when dealing with multicollinear data, but also has several outstanding properties such as interpretability and numerical stability [35]. The last limitation was the certainty of creatinine as a biomarker. Urinary creatinine concentration is changeable based on age, race/ethnicity, body mass index, and time of day. These factors were not controlled in this study. However, other biomarkers from urine change following creatinine concentration. Therefore, when using multiple biomarkers, the interaction among biomarkers is more important rather than factors associated with the creatinine level.

4. Materials and Methods

4.1. Patient Populations and Urine Sample Collections

This study was an observational cohort study with a population of 295 patients diagnosed with ovarian tumors and scheduled for surgery. We obtained their urine samples prospectively before the pathological diagnosis, and excluded 12 cases with borderline tumors after enrollment. Other inclusion criteria were that patients be over the age of 18 years, agree to urine collection before surgery, and provide written informed consent. Exclusion criteria were patients who underwent surgical or open biopsy for any reason within the past 28 days, patients with clinically significant medical problems, and patients who were diagnosed with other cancers within the past 5 years. Urine samples were obtained before the operation as the first-morning urine after more than 8 hours of fasting, and were frozen at -70 or -80 °C within 1 hour of collection until testing. Each urine sample was annotated with information regarding age, menopausal status, histology, and the International Federation of Gynecology and Obstetrics (FIGO) stage (in epithelial ovarian cancer) (Table 1). A collection of samples at the ASAN Medical Center was performed according to strict protocols approved by the Institutional Review Board of the ASAN Medical Center (AMC IRB 2012-0067, approved on 10 February 2012).

4.2. Multiplexed Urinary Biomarker Analysis

The xMAP bead-based technology (Luminex Corp., Austin, TX) facilitates the multiplexed analysis of multiple analytes in a single sample. The 23 bead-based xMAP immunoassays for the differential diagnosis of OC used in this study are listed in Table 5. This list of biomarkers was compiled based on a literature review of current proteins of interest within all fields of circulating biomarkers related to OC research, based on suitable immunoassay availability. The concentration of the biomarkers in urine was calculated by a microbead-based antibody multiplexed immunoassay using Luminex, and the multiplexed immunoassay kit including the cancer biomarkers specifically reacting to OC was used. Bead-based immunoassays targeting 23 protein biomarkers were developed according to strict quality control standards by the University of Pittsburgh Cancer Institute (UPCI) Luminex Core Facility [36] and were performed according to manufacturers' protocols as previously described [16]. As an exception, all urine samples were tested undiluted. Complete biomarker testing was performed immediately upon thawing without testing protein concentration or any other pretest manipulation. The total protein content of each urine sample was measured using the Bio-Plex suspension array system (Bio-Rad Life Science Research, Hercules, CA). Biomarker expression levels were expressed as the median fluorescent intensities generated by analyzing 100 microbeads for each analyte in each sample. The concentrations of the analytes were then quantitated from the median fluorescence intensities using standard curves generated by the Bio-Rad five-parameter curve fitting to the series of known concentrations of each analyte [37].

Table 5. Complete list of multiplexed biomarkers.

Inflammatory Mediators	IL-6 ¹ , IL-8 ¹ , MPO ¹ , MIF ¹ , OPN ¹
Tumor-associated antigens	CA19-9 ¹ , CA15-3 ¹ , CA-125 ¹ , HE4 ¹ , CEA ¹
Adhesion molecules	VCAM ¹ , NCAM ¹
Adipokines	Leptin ¹
Apolipoproteins	ApoAI ¹ , ApoCIII ¹
Apoptotic proteins	Cyfra21-1 ¹
Growth/angiogenic factors	PDGF-AA ¹
Carrier proteins	TTR ²
Proteases/inhibitors	PAI-1 ¹
Hormones	Prolactin ¹
Others	CRP ¹ , Mesomark ³ , Creatinine ⁴

¹ Merck Millipore corp., Darmstadt, Germany; ² Abnova corp., Taipei, Taiwan; ³ Fujirebio Diagnostics, Inc., Göteborg, Sweden; ⁴ Abcam plc., Cambridge, UK.

4.3. Statistical Analysis

The development of statistical models for distinguishing patients with OC from benign tumors was continued until one panel and one model of combining the candidate biomarkers in the panel was selected. The likelihood of OC for each sample was compared to the post-operative confirmatory diagnosis, to evaluate the accuracy of the model. For each biomarker, we used a subsampling strategy to rank markers according to the mean area under the receiver operating characteristic curve (AUC) from 2000 random subsamplings for 80% of the dataset, while maintaining the OC and benign tumor ratio throughout. Using top-performing markers, we constructed multimarker prediction models. To determine the best biomarker combination, 10-fold cross-validation was repeated 2000 times and the top biomarker combinations were selected using the average AUC. By repeating the cross-validation 2000 times, the deviation between the total set of combinations and the chosen subset could be decreased. The combination selected consisted of 2–4 biomarkers out of 23, and the score threshold for the logistic regression was determined using the F-score. Panels were also evaluated based on the sensitivity (SN) at predetermined specificity (SP) levels between 70% and 95%. These SP criteria were selected from reference data on diagnostic products (ROMA and OVA1) that the US FDA had approved for the discrimination of OC from patients with a pelvic mass. Due to the concern that discrepancies in menopausal status may contribute to biomarker panel performance, the final algorithm was applied to the biomarker results in a separate analysis that included menopausal status as an independent variable.

We used a nested cross-validation technique to minimize over-fitting to the dataset. Briefly, 80% of samples were randomly selected to create a training set (i.e., outer loop). In a training set, we used repeated cross-validations to maximize the performance of multimarker models (i.e., inner loop), and the overall performance of a panel was evaluated with the remaining 20% of samples. The OC-to-benign tumor ratio was maintained throughout the resampling procedure [38]. AUC was used as a performance measure, and we used DeLong's method to determine the statistical difference between two AUC values [17,31–34]. Figure 1 shows the modeling process to find the optimum biomarkers and classify. Comparisons between different sample sources were performed using a one-way analysis of variance (ANOVA) with Tukey's multiple comparison tests. An initial minimum level of significance of $p < 0.05$ was utilized. All statistical analyses were performed using the statistical language R.

5. Conclusions

In conclusion, the performance of the HE4, creatinine, CEA, and TTR panel for the differential diagnosis of OC in women with pelvic masses appears promising, with high levels of SN and acceptable SP. Our results suggest that urinary biomarkers provide diagnostic properties exceeding those reported for serum biomarkers by applying not only accuracy but also non-invasiveness. Assuming a patient has a pelvic mass and does not need immediate surgery (but is at high risk of

OC, including a history of breast cancer, familial history of breast/ovarian cancer, or BRCA mutation), she would require continuous check-up. Moreover, a non-invasive accurate test is needed for such a patient. Further evaluation of this panel and other candidate urinary biomarkers (particularly in early-stage OC) would, through prospective studies, expand the clinical utility of urinary multimarker panels.

Supplementary Materials: The following is available online at www.mdpi.com/1422-0067/20/19/4938/s1. Table S1. The estimated value of each single marker in urine samples. (DOCX)

Author Contributions: Conceived and designed the experiments: Y.M.K.; Performed the experiments: S.W.L., H.Y.L., and H.J.B.; Analyzed the data: S.W.L., Y.M.K., H.J.S., and S.W.K.; Analyzed the statistics: S.W.K.; Wrote the paper: S.W.L. and Y.M.K.

Funding: The research was supported by the Research and Business Development Program through the Ministry of Knowledge Economy, Science and Technology (N0000425). This study was supported by a grant from the National R&D Program for Cancer Control, Ministry for Health, Welfare and Family affairs, Republic of Korea (0920010) and a grant of the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health and Welfare, Republic of Korea. (HI06C0868).

Acknowledgments: The authors would like to thank Enago (<http://www.enago.co.kr>) for the English language review.

Conflicts of Interest: The authors declare no conflict of interest.

Abbreviations

OC	ovarian cancer
EOC	epithelial ovarian cancer
SN	sensitivity
SP	specificity
RMI	Risk of Malignancy Index
ROMA	Risk of Malignancy Algorithm
FIGO	the International Federation of Gynecology and Obstetrics
HE4	human epididymis protein 4
VCAM	vascular cell adhesion molecule
TTR	transthyretin
CEA	carcinoembryonic antigen
NCAM	neural cell adhesion molecule
CA-125	cancer antigen 125
CRP	C-reactive protein
PDGF	platelet-derived growth factor
MPO	myeloperoxidase
IL	interleukin
MIF	macrophage migration inhibitory factor
ApoA1	apolipoprotein A1
ApoCIII	apolipoprotein C3
PAI-1	plasminogen activator inhibitor-1
OPN	osteopontin
LASSO	Least Absolute Shrinkage and Selection Operator
PPV	positive predictive value
NPV	negative predictive value
AUC	area under the receiver operating characteristic curve
CI	confidence intervals
IVDMIA	in vitro diagnostic multivariate index assay

References

1. Ferlay, J.; Soerjomataram, I.; Dikshit, R.; Eser, S.; Mathers, C.; Rebelo, M.; Parkin, D.M.; Forman, D.; Bray, F. Cancer incidence and mortality worldwide: sources, methods and major patterns in GLOBOCAN 2012. *Int. J. Cancer* **2015**, *136*, E359–386. doi:10.1002/ijc.29210. PubMed PMID: 25220842.

2. Cohen, J.G.; White, M.; Cruz, A.; Farias-Eisner, R. In 2014, can we do better than CA125 in the early detection of ovarian cancer? *World J. Biol. Chem.* **2014**, *5*, 286–300. doi:10.4331/wjbc.v5.i3.286. PubMed PMID: 25225597.
3. Siegel, R.L.; Miller, K.D.; Jemal, A. Cancer statistics, 2015. *CA Cancer J. Clin.* **2015**, *65*, 5–29. doi:10.3322/caac.21254. PubMed PMID: 25559415.
4. Howlader, N. N.A.; Krapcho, M.; Garshell, J.; Neyman, N., A.S.; Kosary, C.L.; Yu, M.; Ruhl, J.; Tatalovich, Z.; Cho, H., M.A.; *et al.* SEER Cancer Statistics Review, 1975–2010. Available online: https://seer.cancer.gov/archive/csr/1975_2010/ (accessed on 1 December 2017).
5. American College of Obstetricians and Gynecologists Committee on Gynecologic Practice. Committee Opinion No. 477: the role of the obstetrician-gynecologist in the early detection of epithelial ovarian cancer. *Obstet. Gynecol.* **2011**, *117*, 742–746. doi:10.1097/AOG.0b013e31821477db. PubMed PMID: 21343791.
6. Chan, J.K.; Kapp, D.S.; Shin, J.Y.; Husain, A.; Teng, N.N.; Berek, J.S.; Osann, K.; Leiserowitz, G.S.; Cress, R.D.; O'Malley, C. Influence of the gynecologic oncologist on the survival of ovarian cancer patients. *Obstet. Gynecol.* **2007**, *109*, 1342–1350. doi:10.1097/01.Aog.0000265207.27755.28. PubMed PMID: 17540806.
7. Jacobs, I.J.; Menon, U.; Ryan, A.; Gentry-Maharaj, A.; Burnell, M.; Kalsi, J.K.; Amso, N.N.; Apostolidou, S.; Benjamin, E.; Cruickshank, D.; *et al.* Ovarian cancer screening and mortality in the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS): a randomised controlled trial. *Lancet.* **2016**, *387*, 945–956. doi:10.1016/s0140-6736(15)01224-6. PubMed PMID: 26707054.
8. Jacobs, I.; Oram, D.; Fairbanks, J.; Turner, J.; Frost, C.; Grudzinskas, J.G. A risk of malignancy index incorporating CA 125, ultrasound and menopausal status for the accurate preoperative diagnosis of ovarian cancer. *Br. J. Obstet. Gynaecol.* **1990**, *97*, 922–929. doi:10.1111/j.1471-0528.1990.tb02448.x. PubMed PMID: 2223684.
9. Moore, R.G.; McMeekin, D.S.; Brown, A.K.; DiSilvestro, P.; Miller, M.C.; Allard, W.J.; Gajewski, W.; Kurman, R.; Bast R.C., Jr.; Skates, S.J. 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. doi:10.1016/j.ygyno.2008.08.031. PubMed PMID: 18851871.
10. Zhang, Z.; Bast R.C., Jr.; Yu, Y.; Li, J.; Sokoll, L.J.; Rai, A.J.; Rosenzweig, J.M.; Cameron, B.; Wang, Y.Y.; Meng, X.Y. Three biomarkers identified from serum proteomic analysis for the detection of early stage ovarian cancer. *Cancer Res.* **2004**, *64*, 5882–5890. doi:10.1158/0008-5472.Can-04-0746. PubMed PMID: 15313933.
11. Ueland, F.R.; Desimone, C.P.; Seamon, L.G.; Miller, R.A.; Goodrich, S.; Podzielinski, I.; Sokoll, L.; Smith, A.; van Nagell, J.R., Jr.; Zhang, Z. Effectiveness of a multivariate index assay in the preoperative assessment of ovarian tumors. *Obstet. Gynecol.* **2011**, *117*, 1289–1297. doi:10.1097/AOG.0b013e31821b5118. PubMed PMID: 21606739.
12. Ye, B.; Skates, S.; Mok, S.C.; Horick, N.K.; Rosenberg, H.F.; Vitonis, A.; Edwards, D.; Sluss, P.; Han, W.K.; Berkowitz, R.S.; *et al.* Proteomic-based discovery and characterization of glycosylated eosinophil-derived neurotoxin and COOH-terminal osteopontin fragments for ovarian cancer in urine. *Clin. Cancer Res.* **2006**, *12*, 432–441. doi:10.1158/1078-0432.Ccr-05-0461. PubMed PMID: 16428483.
13. Badgwell, D.; Lu, Z.; Cole, L.; Fritsche, H.; Atkinson, E.N.; Somers, E.; Allard, J.; Moore, R.G.; Lu, K.H.; Bast, R.C., Jr. Urinary mesothelin provides greater sensitivity for early stage ovarian cancer than serum mesothelin, urinary hCG free beta subunit and urinary hCG beta core fragment. *Gynecol. Oncol.* **2007**, *106*, 490–497. doi:10.1016/j.ygyno.2007.04.022. PubMed PMID: 17532030.
14. Anderson, N.S.; Bermudez, Y.; Badgwell, D.; Chen, R.; Nicosia, S.V.; Bast, R.C., Jr.; Kruk, P.A. Urinary levels of Bcl-2 are elevated in ovarian cancer patients. *Gynecol. Oncol.* **2009**, *112*, 60–67. doi:10.1016/j.ygyno.2008.09.037. PubMed PMID: 19007973.
15. Liao, J.B.; Yip, Y.Y.; Swisher, E.M.; Agnew, K.; Hellstrom, K.E.; Hellstrom, I. Detection of the HE4 protein in urine as a biomarker for ovarian neoplasms: Clinical correlates. *Gynecol. Oncol.* **2015**, *137*, 430–435. doi:10.1016/j.ygyno.2015.03.044. PubMed PMID: 25866324.
16. Yurkovetsky, Z.; Skates, S.; Lomakin, A.; Nolen, B.; Pulsipher, T.; Modugno, F.; Marks, J.; Godwin, A.; Gorelik, E.; Jacobs, I.; *et al.* Development of a multimarker assay for early detection of ovarian cancer. *J. Clin. Oncol.* **2010**, *28*, 2159–2166. doi:10.1200/jco.2008.19.2484. PubMed PMID: 20368574.
17. Amonkar, S.D.; Bertenshaw, G.P.; Chen, T.H.; Bergstrom, K.J.; Zhao, J.; Seshiaiah, P.; Yip, P.; Mansfield, B.C. Development and preliminary evaluation of a multivariate index assay for ovarian cancer. *PLoS One* **2009**, *4*, e4599. doi:10.1371/journal.pone.0004599. PubMed PMID: 19240799.

18. Kim, K.; Visintin, I.; Alvero, A.B.; Mor, G. Development and validation of a protein-based signature for the detection of ovarian cancer. *Clin. Lab. Med.* **2009**, *29*, 47–55. doi:10.1016/j.cll.2009.02.001. PubMed PMID: 19389550.
19. Kim, Y.W.; Bae, S.M.; Lim, H.; Kim, Y.J.; Ahn, W.S. Development of multiplexed bead-based immunoassays for the detection of early stage ovarian cancer using a combination of serum biomarkers. *PLoS ONE* **2012**, *7*, e44960. doi:10.1371/journal.pone.0044960. PubMed PMID: 22970327.
20. Nolen, B.M.; Lokshin, A.E. Biomarker testing for ovarian cancer: clinical utility of multiplex assays. *Mol. Diagn. Ther.* **2013**, *17*, 1391–46. doi:10.1007/s40291-013-0027-6. PubMed PMID: 23552992.
21. Chambers, A.F.; Vanderhyden, B.C. Ovarian cancer biomarkers in urine. *Clin. Cancer Res.* **2006**, *12*, 323–327. doi:10.1158/1078-0432.Ccr-05-2563. PubMed PMID: 16428467.
22. Hellstrom, I.; Heagerty, P.J.; Swisher, E.M.; Liu, P.; Jaffar, J.; Agnew, K.; Hellstrom, K.E. Detection of the HE4 protein in urine as a biomarker for ovarian neoplasms. *Cancer Lett.* **2010**, *296*, 43–48. doi:10.1016/j.canlet.2010.03.013. PubMed PMID: 20381233.
23. Nolen, B.M.; Orlichenko, L.S.; Marrangoni, A.; Velikokhatnaya, L.; Prosser, D.; Grizzle, W.E.; Ho, K.; Jenkins, F.J.; Bovbjerg, D.H.; Lokshin, A.E. An extensive targeted proteomic analysis of disease-related protein biomarkers in urine from healthy donors. *PLoS ONE* **2013**, *8*, e63368. doi:10.1371/journal.pone.0063368. PubMed PMID: 23723977.
24. Nolen, B.M.; Lomakin, A.; Marrangoni, A.; Velikokhatnaya, L.; Prosser, D.; Lokshin, A.E. Urinary protein biomarkers in the early detection of lung cancer. *Cancer Prev. Res (Phila)*. **2015**, *8*, 111–119. doi:10.1158/1940-6207.Capr-14-0210. PubMed PMID: 25416410.
25. Menon, U.; Jacobs, I.J. Ovarian cancer screening in the general population. *Curr. Opin. Obstet. Gynecol.* **2001**, *13*, 61–64. PubMed PMID: 11176234.
26. Jacobs, I.J.; Menon, U. Progress and challenges in screening for early detection of ovarian cancer. *Mol. Cell. Proteomics* **2004**, *3*, 355–366. doi:10.1074/mcp.R400006-MCP200. PubMed PMID: 14764655.
27. Moore, R.G.; Jabre-Raughley, M.; Brown, A.K.; Robison, K.M.; Miller, M.C.; Allard, W.J.; Kurman, R.J.; Bast, R.C.; Skates, S.J. Comparison of a novel multiple marker assay vs the Risk of Malignancy Index for the prediction of epithelial ovarian cancer in patients with a pelvic mass. *Am. J. Obstet. Gynecol.* **2010**, *203*, 228.e1–6. doi:10.1016/j.ajog.2010.03.043. PubMed PMID: 20471625.
28. Karlsen, M.A.; Sandhu, N.; Hogdall, C.; Christensen, I.J.; Nedergaard, L.; Lundvall, L.; Engelholm, S.A.; Pedersen, A.T.; Hartwell, D.; Lydolph, M.; et al. Evaluation of HE4, CA125, risk of ovarian malignancy algorithm (ROMA) and risk of malignancy index (RMI) as diagnostic tools of epithelial ovarian cancer in patients with a pelvic mass. *Gynecol. Oncol.* **2012**, *127*, 379–383. doi:10.1016/j.ygyno.2012.07.106. PubMed PMID: 22835718.
29. Grenache, D.G.; Heichman, K.A.; Werner, T.L.; Vucetic, Z. Clinical performance of two multi-marker blood tests for predicting malignancy in women with an adnexal mass. *Clin. Chim. Acta* **2015**, *438*, 358–363. doi:10.1016/j.cca.2014.09.028. PubMed PMID: 25283731.
30. Bristow, R.E.; Smith, A.; Zhang, Z.; Chan, D.W.; Crutcher, G.; Fung, E.T.; Munroe, D.G. Ovarian malignancy risk stratification of the adnexal mass using a multivariate index assay. *Gynecol. Oncol.* **2013**, *128*, 252–259. doi:10.1016/j.ygyno.2012.11.022. PubMed PMID: 23178277.
31. Efron, B.; Tibshirani, R. An introduction to the bootstrap. Chapman and Hall, London. **1993**.
32. Efron, B. Second Thoughts on the Bootstrap. *Statist. Sci.* **2003**, *18*, 135–140.
33. Henderson, A.R. The bootstrap: a technique for data-driven statistics. Using computer-intensive analyses to explore experimental data. *Clin. Chim. Acta* **2005**, *359*, 1–26. doi:10.1016/j.cccn.2005.04.002. PubMed PMID: 15936746.
34. DeLong, E.R.; DeLong, D.M.; Clarke-Pearson, D.L. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. *Biometrics* **1988**, *44*, 837–845. PubMed PMID: 3203132.
35. Musoro, J.Z.; Zwinderman, A.H.; Puhan, M.A.; ter Riet, G.; Geskus, R.B. Validation of prediction models based on lasso regression with multiply imputed data. *BMC Med. Res. Methodol.* **2014**, *14*, 116. doi:10.1186/1471-2288-14-116. PubMed PMID: 25323009.
36. UPCI Luminex Core Facility. Available online: <http://www.upci.upmc.edu/luminex/index.cfm> (accessed on 1 December 2017).

37. Nolen, B.M.; Marks, J.R.; Ta'san, S.; Rand, A.; Luong, T.M.; Wang, Y.; Blackwell, K.; Lokshin, A.E. Serum biomarker profiles and response to neoadjuvant chemotherapy for locally advanced breast cancer. *Breast Cancer Res.* **2008**, *10*, R45. doi:10.1186/bcr2096. PubMed PMID: 18474099.
38. Kong, S.W.; Collins, C.D.; Shimizu-Motohashi, Y.; Holm, I.A.; Campbell, M.G.; Lee, I.H.; Brewster, S.J.; Hanson, E.; Harris, H.K.; Lowe, K.R.; et al. Characteristics and predictive value of blood transcriptome signature in males with autism spectrum disorders. *PLoS One* **2012**, *7*, e49475. doi:10.1371/journal.pone.0049475. PubMed PMID: 23227143.



© 2019 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).