

Simultaneous Measurement of 92 Serum Protein Biomarkers for the Development of a Multiprotein Classifier for Ovarian Cancer Detection



Amy P.N. Skubitz^{1,2,3,4}, Kristin L.M. Boylan^{1,2}, Kate Geschwind^{1,2}, Qing Cao⁴, Timothy K. Starr^{3,4,5}, Melissa A. Geller^{3,4}, Joseph Celestino⁶, Robert C. Bast Jr.⁶, Karen H. Lu⁶, and Joseph S. Koopmeiners^{4,7}

Abstract

The best known ovarian cancer biomarker, CA125, is neither adequately sensitive nor specific for screening the general population. By using a combination of proteins for screening, it may be possible to increase the sensitivity and specificity over CA125 alone. In this study, we used Proseek Multiplex Oncology II plates to simultaneously measure the expression of 92 cancer-related proteins in serum using proximity extension assays. This technology combines the sensitivity of the PCR with the specificity of antibody-based detection methods, allowing multiplex biomarker detection and high-throughput quantification. We analyzed 1 μ L of sera from each of 61 women with ovarian cancer and compared the values obtained with those from 88 age-matched healthy women. Principle component analysis and unsupervised hier-

archical clustering separated the ovarian cancer patients from the healthy, with minimal misclassification. Data from the Proseek plates for CA125 levels exhibited a strong correlation with clinical values for CA125. We identified 52 proteins that differed significantly ($P < 0.006$) between ovarian cancer and healthy samples, several of which are novel serum biomarkers for ovarian cancer. In total, 40 proteins had an estimated area under the ROC curve of 0.70 or greater, suggesting their potential to serve as biomarkers for ovarian cancer. CA125 alone achieved a sensitivity of 93.4% at a specificity of 98%. By adding the Oncology II values for five proteins to CA125 in a multiprotein classifier, we increased the assay sensitivity to 98.4% at a specificity of 98%, thereby improving the sensitivity and specificity of CA125 alone.

Introduction

Ovarian cancer is the fifth leading cause of cancer deaths in women in the United States with over 16,000 women succumbing to the disease annually. Although early detection has the potential to increase patient survival, a test

with sufficient sensitivity and specificity for screening the general population has not yet been developed (1, 2).

Two ovarian cancer serum biomarkers, CA125 and HE4, are FDA approved and used for monitoring ovarian cancer recurrence. However, neither biomarker has adequate sensitivity or specificity, when used alone, to screen the general population for early stage disease. The risk of ovarian cancer algorithm has been used in screening trials to measure increases in CA125 levels over time to trigger transvaginal sonography in a small fraction of women screened (3). These trials showed improvement over using a single cutoff value for CA125 in women with an average risk of ovarian cancer, with excellent specificity and positive predictive value (4). Early stage disease was detected in the Normal Risk Ovarian Cancer Screening Study (NROSS; ref. 5), but only the United Kingdom Collaborative Trial of Ovarian Cancer Screening (UKCTOCS) trial was adequately powered to detect an improvement in survival with more than 200,000 participants (6). Although the trial failed to achieve statistical significance in the overall population, a 20% reduction in mortality ($P < 0.021$) was observed in a

¹Department of Laboratory Medicine and Pathology, University of Minnesota, Minneapolis, Minnesota. ²Ovarian Cancer Early Detection Program, University of Minnesota, Minneapolis, Minnesota. ³Department of Obstetrics, Gynecology, and Women's Health, University of Minnesota, Minneapolis, Minnesota. ⁴Masonic Cancer Center, University of Minnesota, Minneapolis, Minnesota. ⁵Department of Genetics, Cell Biology and Development, University of Minnesota, Minneapolis, Minnesota. ⁶University of Texas, MD Anderson Cancer Center, Houston, Texas. ⁷Division of Biostatistics, School of Public Health, University of Minnesota, Minneapolis, Minnesota.

Note: Supplementary data for this article are available at Cancer Prevention Research Online (<http://cancerprevres.aacrjournals.org/>).

Corresponding Author: Amy P.N. Skubitz, University of Minnesota, MMC 395, 420 Delaware Street, S.E., Minneapolis, MN 55455. Phone: 612-625-5920; Fax: 612-626-0665; E-mail: skubi002@umn.edu

doi: 10.1158/1940-6207.CAPR-18-0221

©2019 American Association for Cancer Research.

prespecified subset of women with prevalent nonprimary peritoneal disease. With wide confidence limits around this estimate, additional follow-up will be required, but it is already apparent that greater benefit might be attained with a more sensitive initial step in screening that included multiple biomarkers.

In the past, many groups have attempted to improve upon the use of CA125 alone as a serum biomarker for ovarian cancer by evaluating additional proteins individually and in combination with CA125 (1, 2). Although an ELISA can quantify the level of an individual protein, it loses its capacity to accurately measure more than 10 proteins at a time. The ability to screen dozens of biomarkers, simultaneously in one multiplex assay, has the potential to optimize biomarker discovery and facilitates analysis of multiprotein classifiers. In addition, multiplex screening could improve the sensitivity and specificity to the extent that the assay could reasonably be used to screen an asymptomatic general population for early stage cases. Several groups have used other technologies such as Lumines bead-based assay panels to identify subsets of biomarkers to detect ovarian cancer (2, 7, 8).

Technology developed by OLink Bioscience permits the simultaneous quantification of 92 cancer-related protein biomarkers based on the proximity extension assay (PEA; refs. 9–12). PEA is an innovative technology that combines the specificity of antibody-based detection methods (such as ELISA) with the sensitivity of the PCR, permitting multiplex biomarker detection and quantification with similar assay precision to other multiplex detection methods using only microliter quantities of patient sera.

Recently, we tested a small sample set (~20 cases each of healthy, benign ovarian, early stage serous ovarian cancer, and late-stage serous ovarian cancer) on the Proseek Multiplex Oncology Iv2 panel to determine its feasibility as a means to identify candidate biomarkers for early stage serous ovarian cancer (13). We demonstrated that the Proseek technology provides similar results to conventional clinical assays for CA125, and can identify new candidate biomarkers that improve the sensitivity and specificity of CA125 alone. From the Proseek Oncology Iv2 data, we developed a multiprotein classifier that combined 12 proteins. This multiprotein classifier improved the sensitivity (i.e., ability to detect positive cases) of CA125 alone from 93% to 95% when comparing sera from healthy women with early stage ovarian cancer patients. This improvement was calculated when we set the specificity, i.e., ability to correctly identify negative cases, at 95%. The Oncology Iv2 plate has been discontinued and is no longer commercially available.

We report herein the first use of the Proseek Oncology II plate to identify candidate biomarkers for high-grade serous ovarian cancer, because it is the most prevalent and deadly subtype of ovarian cancer. In this study, a larger cohort of serum samples from 61 women with advanced stage serous ovarian cancer and 88 healthy, age-matched

controls were tested on the updated Oncology II panel, which includes 92 proteins that are differentially expressed in a variety of cancers. Our goal was to identify proteins that could aid in the detection of ovarian cancer. Similar to the Oncology Iv2 panel, the Oncology II panel contains CA125 and HE4, as well as several other proteins that are known to be associated with ovarian cancer, such as ERBB2, ERBB3, ERBB4, VEGFR2, midkine (MK), kallikrein 11 (KLK11), folate receptor-alpha, IL6, and TGF-alpha. The Oncology II panel also contains new assays for ovarian cancer-associated proteins kallikrein 13 (KLK13), Nectin-4 (NECT4 or PVRL4), and mesothelin (MSLN).

Materials and Methods

Patient and control sera

Blood samples were obtained preoperatively from 61 patients with epithelial ovarian/fallopian tube or primary peritoneal cancer and from 88 healthy participants who had not developed cancer in the NROSS screening trial coordinated by the MD Anderson Ovarian SPORE (5). All participants provided written consent for use of samples in protocols approved by the MD Anderson's Institutional Review Board. All samples were processed by the same standard operating procedure, whereby immediately after the blood was drawn, the collection tubes were refrigerated at 4°C and then processed by centrifugation as soon as possible on the day of collection. Samples were then aliquoted and frozen at -80°C. Clinical, pathologic, and demographic information on subjects are presented in Supplementary Table S1. Preoperative samples from ovarian cancer patients and control samples from healthy individuals enrolled in the NROSS trial were assayed for CA125 on a Roche platform in the clinical laboratory of MD Anderson Cancer Center.

Olink Proseek Oncology II multiplex assay

The reagents used in the Olink Proseek Multiplex 96-well plates are based on PEA technology in which 92 oligonucleotide-labeled antibody pairs bind to their respective protein targets in proximity to one another forming a PCR reporter sequence by DNA base-pairing which is subsequently detected and quantified using real-time PCR (9–11). This combination of antibody detection followed by PCR quantification permits the specific and sensitive analysis of 92 proteins in each well of a 96-well plate. The precision, reproducibility, and scalability of the PEA assay have been documented by the manufacturer (<http://www.olink.com>) and others (9–11). The protein names, gene names, and abbreviations for the 92 proteins that are included on the Proseek Oncology II plate are listed in Supplementary Table S2.

Sample processing

We randomized the patients' serum samples across four 96-well plates such that each sample was assayed once. On

each plate, we included two "Pooled Reference" controls, containing equal volumes of serum from each of the samples. The Proseek platform also includes three "Interplate controls" for data normalization between plates and three "Negative controls" to establish background levels. One microliter of serum from each of the samples was mixed with the Oncology II reagents following the manufacturer's protocol, then processed in combination with the Fluidigm BioMark HD high-throughput PCR instrument at the University of Minnesota Genomics Center as previously described (13). Data generated from the plates were analyzed, including normalization and linearization, per manufacturer's protocol. The assay reports relative quantification on a \log_2 scale, as Normalized Protein eXpression (NPX) values, which are cycle threshold (C_t) values normalized by the subtraction of values for the extension control. All assay characteristics including detection limits and measurements of assay performance and validation are available from the manufacturer's website (<http://www.olink.com/products/>).

ELISA

Serum samples were randomly selected from the cases and controls at the same ratio as was used for the Proseek plates (35 ovarian cancers and 53 healthy controls). ELISAs for HE4 (R&D Systems), EGF (Raybiotech and R&D Systems), and MSLN (R&D Systems and Biolegend) were performed following the manufacturer's instructions.

Data analysis and statistical analysis

Linear protein values were \log_2 -transformed and mean-centered to produce a data matrix of 149 patients by 92 proteins. Unsupervised clustering methods were applied to the data to identify clusters of proteins and visually evaluate their association with disease status. Unsupervised hierarchical clustering (uncentered correlation using centroid linkage) was completed using Cluster 3.0 (14) and visualized using Treeview (v1.1.6r4; ref. 15). Principle component analysis (PCA) was performed using the *prcomp* function in R (16). The protein folate receptor-gamma (FR-gamma/FOLR3) was not included in the PCA because it was expressed at high levels in a subset of serum samples that did not correlate with ovarian cancer status.

The protein values obtained from the Proseek plates were compared with the Roche platform or ELISA values using Pearson correlation coefficient. For consistency, all values were analyzed on the \log_2 scale.

Protein expression was compared between healthy and late-stage ovarian cancer patients using the two-sample *t* test assuming unequal variances. Proteins were considered significant if the *P* value for the two-sample *t* test was less than 0.006, which was chosen to obtain an FDR of 0.05 (17). The classification accuracy was evaluated using the ROC curve and was summarized by the area under the ROC curve (AUC) and the sensitivity corresponding to a

specificity of 95% [i.e., ROC(0.05)] or 98% [i.e., ROC(0.02)]. The ROC curve and its summaries were estimated using a nonparametric approach (18). Confidence intervals (CI) for AUC, ROC(0.05) and ROC(0.02), were calculated using the nonparametric bootstrap (19).

A multiprotein classifier for discriminating between sera from women with ovarian cancer and healthy women was developed using supervised machine-learning methods. Variable selection was completed using the LASSO. Proteins were sequentially added to the model by manipulating the LASSO penalty parameter (20). The classification accuracy of our predictive model was summarized by the ROC curve and its summaries [AUC, ROC(0.05), and ROC(0.02)]. Leave-one-out cross-validation was used to correct for the bias that results from validating our model on the same data that were used to build the model (21), and confidence intervals were calculated using the nonparametric bootstrap (19).

Results

Reproducibility of results on Oncology II plates

The Proseek multiplex assay uses a PEA that is capable of simultaneously quantifying expression levels of up to 92 proteins in a 1 μ L serum sample. We used the Proseek assay in conjunction with the Oncology II plate to measure the levels of 92 cancer-associated proteins (Supplementary Table S2) in serum samples taken from 61 patients with high-grade serous ovarian cancer and 88 age-matched healthy controls. Two "Pooled Reference" samples were included on each of the four plates, for a total of eight identical controls. The intraplate coefficient of variation (CV) was 8% before and after normalization across the four plates, with only 4 of the 92 proteins having a CV >15%. The interplate CV was 21% before normalization and 12% after normalization, with only 6 of the 92 proteins having a CV >20% after normalization.

Proseek values strongly correlate with CA125 and HE4 measurements

To confirm the validity of protein expression measurements by Proseek using PEA, we compared the CA125 and HE4 values obtained from the Proseek plate with the clinical values obtained from the patients' medical records of CA125 (Fig. 1A) and ELISA values for HE4 (Fig. 1B). The data from the Proseek plate exhibited a strong correlation with the protein values, with a Pearson correlation coefficient of 0.91 (95% CI, 0.88–0.94) for CA125 and a Pearson correlation coefficient of 0.86 (95% CI, 0.80–0.91) for HE4, indicating the PEA technology is comparable with clinically validated testing and ELISAs.

A protein signature distinguishes ovarian cancer from healthy patients

We performed PCA and plotted all 149 samples using the first three principle components (Fig. 1C; Supplementary

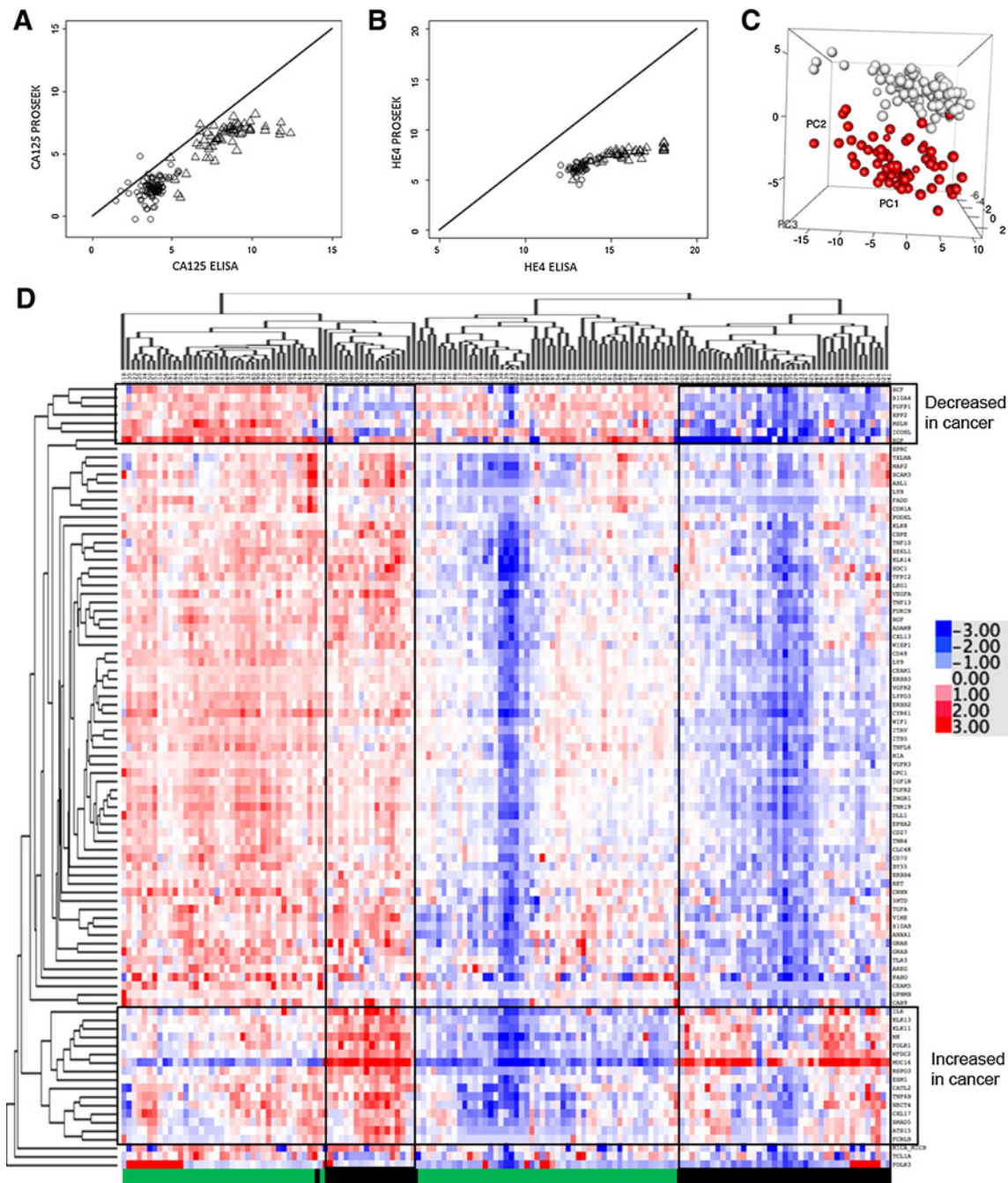


Figure 1. Oncology II protein results correlate with clinical and ELISA values of CA125 and HE, and cluster ovarian cancer samples from healthy controls. **A**, Comparison of the CA125 values obtained on the Roche platform versus the Proseek plate. Scatterplot comparison of the CA125 values for each of the serum samples from the 88 healthy women (circles) and the 61 late-stage serous ovarian cancer patients (triangles). Correlation coefficient of 0.91 (95% CI, 0.88–0.94). **B**, Comparison of the HE4 values obtained by ELISA versus the Proseek plate. Scatterplot comparison of the HE4 values for each of the serum samples from the 53 healthy women (circles) and the 35 late-stage serous ovarian cancer patients (triangles). Correlation coefficient of 0.86 (95% CI, 0.8–0.91). **C**, Principal component analysis of Proseek protein expression data. Principal component analysis plots were based on expression levels of 91 proteins measured in the sera of 149 patients using the Proseek Oncology II Multiplex plates. Gray circles represent the 88 healthy women, and red circles represent 61 late-stage high-grade serous ovarian cancer patients. A rotating three-dimensional version of this figure is provided as Supplementary Movie S1. **D**, Unsupervised hierarchical clustering of Proseek protein expression data. Unsupervised hierarchical clustering was based on mean-centered \log_2 -transformed protein expression data of 92 Proseek Oncology II proteins (Correlation uncentered, average linkage) measured in the sera of 149 patients. Dark red indicates high levels of the protein relative to the average value, white indicates the average value, and dark blue indicates that the protein levels are below average (shown in the color bar on the right-hand side). Color bar at the bottom indicates patient classification: healthy (green) and late-stage high-grade serous ovarian cancer (black).

Table 1. Twelve proteins that were expressed at significantly higher levels in ovarian cancer versus healthy serum samples

Proseek protein abbreviation	Protein name	HUGO gene symbol	Log ₂ value on Proseek Oncology II Plate ^a		P value ^b
			Healthy patients	Ovarian cancer patients	
MUC16	CA-125 or Mucin-16	MUC16	2.4 (0.84)	6.48 (1.33)	<0.001
WFDC2	WAP four-disulfide core domain protein 2 (HE4)	WFDC2	6.54 (0.6)	7.82 (0.7)	<0.001
MK	Midkine	MDK	5.38 (1.16)	6.67 (1.12)	<0.001
KLK13	Kallikrein 13	KLK13	2.61 (0.66)	3.33 (0.78)	<0.001
ADAM.TS.15	A disintegrin and metalloproteinase with thrombospondin motifs 15	ADAMTS15	0.33 (0.59)	1.06 (0.92)	<0.001
IL6	Interleukin-6	IL6	1.78 (1.01)	3.01 (1.6)	<0.001
hK11 or KLK11	Kallikrein 11	KLK11	4.62 (0.65)	5.35 (0.95)	<0.001
RSPO3	R-Spondin-3	RSPO3	1.17 (0.71)	1.9 (1.02)	<0.001
FR-alpha	Folate receptor alpha	FOLR1	6.44 (0.79)	7.24 (1.16)	<0.001
TNFRSF6B	Tumor necrosis factor receptor superfamily member 6B	TNFRSF6B	3.02 (1.05)	3.73 (1.22)	<0.001
FCRLB	Fc-receptor-like B	FCRLB	0.05 (0.52)	0.39 (0.7)	0.002
PVRL4	Nectin-4	NECTIN4	5.09 (1.07)	5.61 (1.12)	0.005

^aMean log₂ value with SD in parenthesis.^bFDR < 0.05.

Movie S1). Visual analysis of the PCA plot clearly separated almost all of the ovarian cancer cases (red) from the healthy samples (gray).

Similar results were obtained when we subjected the protein expression profiles to unsupervised hierarchical clustering (Fig. 1D). Interestingly, the first major split between samples (left vs. right) was not based on cancer versus healthy, but instead was based on generally high protein expression (red) of all proteins compared with generally low expression (blue) of all proteins. However, each of these two major groups was split into two subgroups that correlated almost perfectly with cancer (black bar at bottom) versus healthy patients (green bar at bottom). Visual inspection of the heat map identified proteins that were elevated in the sera for the majority of women with ovarian cancer, as depicted in red (bracketed). Importantly, CA125 (MUC16) and HE4 (WFDC2) are present in this region of the heat map, as well as other proteins that are reportedly elevated in the sera of women with ovarian cancer, such as IL6 (22), KLK13 (23), KLK11 (24), MK (25), and NECT4 (or PVRL4; ref. 26). A second cluster in the heat map contained proteins that were lower in the sera for the majority of women with ovarian cancer compared with the healthy controls, depicted in blue (bracketed). These proteins included stem cell factor/KIT ligand (SCF/KTLIG), S100 calcium binding protein A4 (S10A4 or S100A4), fibroblast growth factor binding protein 1 (FGFBP1 or FGFP1), X-prolyl aminopeptidase 2 (XPNPEP2 or XPP2), MSLN, inducible T-cell costimulator ligand (ICOSL or ICOSLG), and EGF. These results indicate it is possible to define a protein signature that can separate ovarian cancer samples from healthy samples.

Identification of potential novel biomarkers elevated in ovarian cancer sera

The 92 proteins measured on the Oncology II plate were originally selected because of their known roles

in multiple cancers, and were not selected to be ovarian cancer specific. To identify which of the 92 proteins could serve as potential biomarkers of ovarian cancer, we performed differential expression analysis and identified 12 proteins significantly elevated in ovarian cancer samples compared with healthy samples ($P < 0.006$ with an FDR < 0.05; Table 1). As anticipated, CA125 and HE4 were present at low levels in healthy controls, whereas higher levels were found in the late-stage serous ovarian cancer cases (Fig. 2A). Another six proteins have previously been reported to be elevated in the sera of ovarian cancer patients, and these were confirmed by our data. These proteins are folate receptor-alpha (FR-alpha; ref. 27), IL6 (22, 28), KLK13 (23), KLK11 (24), MK (25), and NECT4 (26). The remaining four proteins (ADAMTS.15, FCRLB, RSPO3, and TNFRSF6B) have not been previously shown to be elevated in the sera of ovarian cancer patients, and therefore, represent potential novel serum biomarkers of ovarian cancer. A complete listing of the expression values for all 92 proteins is provided in Supplementary Table S3.

To determine if there is a correlation between gene expression in tumor tissue compared with protein levels in sera, we mined The Cancer Genome Atlas (TCGA) data via cBioPortal for RNASeq expression levels of the 12 upregulated serum biomarkers. Almost half of TCGA ovarian cancer tissue samples (48%; 141/295) had elevated levels of expression of at least one of the 12 genes, indicating that in some cases, the elevated protein expression in the sera may correlate with increased gene expression in the tumor (Supplementary Fig. S1).

Biomarker specificity and sensitivity

To assess the sensitivity (i.e., the probability that an ovarian cancer serum sample will be identified correctly) and specificity (i.e., the probability that a healthy serum

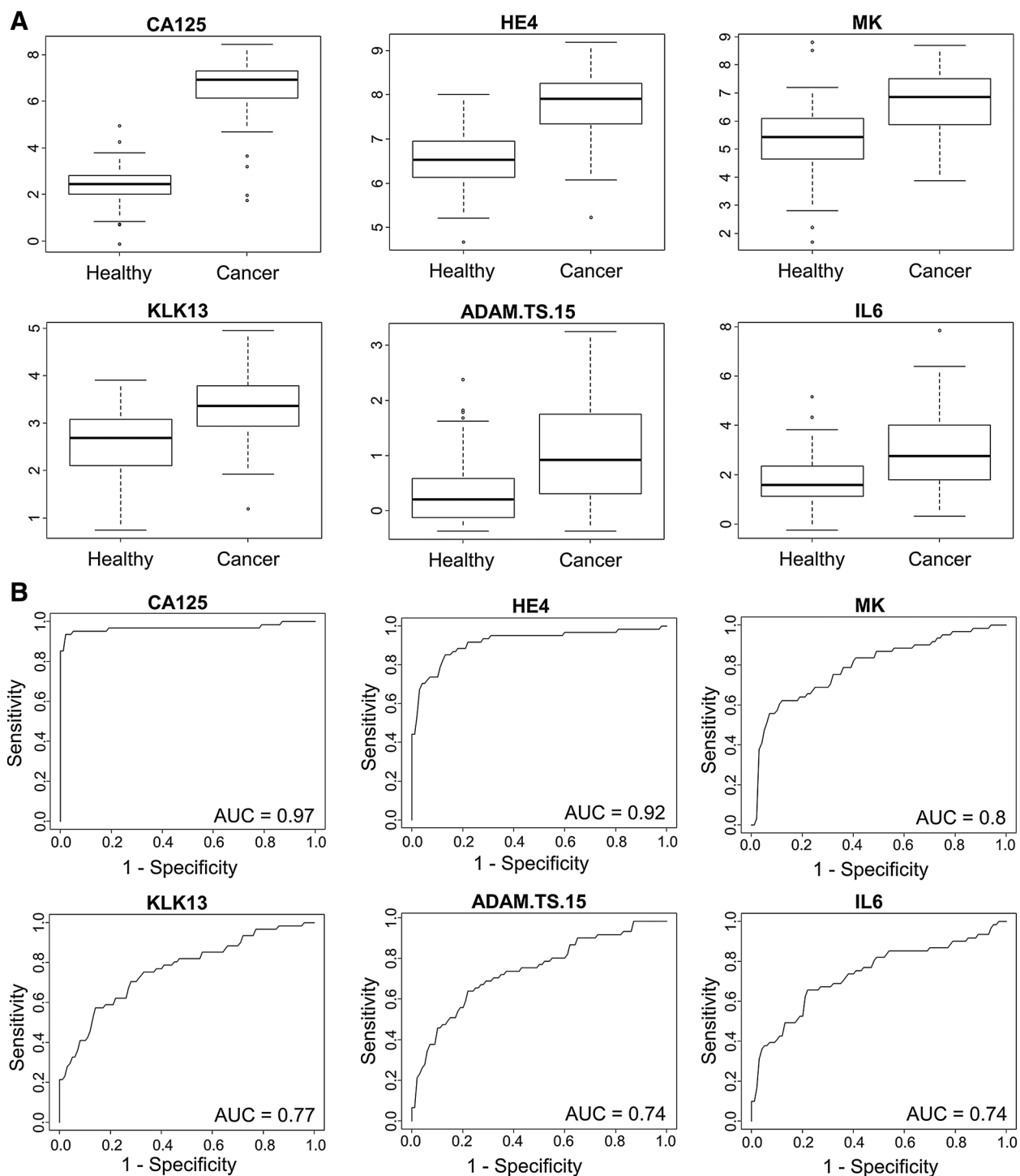


Figure 2.

Six Oncology II proteins expressed at elevated levels in the sera of ovarian cancer patients compared with healthy controls. **A**, Quantile boxplots showing expression of the top six overexpressed proteins in ovarian cancer serum (Cancer) compared with controls (Healthy). Outliers are defined as any value higher or lower than 1.5 multiplied by the interquartile range. A complete list of the mean values for all 92 proteins is provided in Supplementary Table S3. **B**, ROC curves for six proteins that showed increased Oncology II values in ovarian cancer serum compared with healthy controls. ROC curves were graphed for the same six proteins with the highest AUC values from **A**. Data for the 40 proteins with AUC values of 0.70 or greater are shown in Table 2. A complete list of the AUC values for all 92 proteins is provided in Supplementary Table S4.

Table 2. AUC, sensitivity, and specificity of proteins when comparing sera from ovarian cancer patients versus healthy women

Protein	AUC (95% CI)	Rank	Sensitivity at 95% specificity (95% CI)	Rank	Sensitivity at 98% specificity (95% CI)	Rank
CA125/MUC.16	0.97 (0.93-1)	1	0.95 (0.88-1)	2	0.93 (0.8-0.98)	1
FGF.BP1	0.97 (0.93-0.99)	2	0.97 (0.13-1)	1	0.15 (0.03-0.98)	17
S100A4	0.92 (0.87-0.96)	3	0.69 (0.25-0.87)	4	0.16 (0.05-0.79)	14
HE4/WFDC2	0.92 (0.87-0.96)	4	0.7 (0.52-0.84)	3	0.54 (0.36-0.79)	2
ITGAV	0.9 (0.84-0.94)	5	0.43 (0.07-0.84)	9	0.07 (0-0.67)	24
ICOSLG	0.88 (0.81-0.94)	6	0.51 (0.34-0.77)	5	0.41 (0.2-0.61)	3
SCF	0.84 (0.77-0.9)	7	0.11 (0.02-0.56)	55	0.03 (0-0.38)	40
GPNMB	0.82 (0.74-0.89)	8	0.44 (0.07-0.67)	8	0.1 (0.03-0.56)	21
MSLN	0.81 (0.74-0.88)	9	0.38 (0.2-0.63)	11	0.26 (0.15-0.43)	4
EGF	0.81 (0.74-0.87)	10	0.48 (0.16-0.61)	6	0.18 (0.1-0.51)	13
ITGB5	0.8 (0.73-0.87)	11	0.23 (0.01-0.66)	30	0.02 (0-0.35)	52
FADD	0.8 (0.73-0.87)	12	0.33 (0-0.52)	16	0 (0-0.36)	65
CD207	0.8 (0.73-0.87)	13	0.28 (0.05-0.61)	20	0.07 (0.02-0.38)	24
MK	0.8 (0.73-0.87)	14	0.48 (0.05-0.66)	6	0.03 (0-0.52)	40
FASLG	0.78 (0.71-0.86)	15	0.23 (0.02-0.58)	30	0.03 (0-0.29)	40
CEACAM5	0.78 (0.7-0.85)	16	0 (0-0.52)	88	0 (0-0)	65
XPNPEP2	0.78 (0.7-0.85)	17	0.38 (0.18-0.51)	11	0.21 (0-0.43)	9
KLK13	0.77 (0.68-0.84)	18	0.33 (0.2-0.49)	16	0.23 (0.14-0.4)	8
IFN.gamma.R1	0.76 (0.68-0.83)	19	0.3 (0.02-0.5)	19	0.03 (0-0.34)	40
CD48	0.76 (0.68-0.84)	20	0.08 (0.02-0.51)	67	0.05 (0-0.19)	35
MIA	0.75 (0.67-0.83)	21	0.16 (0-0.43)	44	0.02 (0-0.22)	52
LY9	0.74 (0.67-0.82)	22	0.2 (0-0.52)	34	0.02 (0-0.36)	52
ADAM.TS.15	0.74 (0.67-0.83)	23	0.28 (0.15-0.51)	20	0.21 (0.03-0.36)	9
IL6	0.74 (0.65-0.83)	24	0.38 (0.15-0.51)	11	0.16 (0.05-0.44)	14
TGFR.2	0.74 (0.65-0.82)	25	0.34 (0.02-0.6)	15	0.05 (0-0.45)	35
GZMB	0.73 (0.65-0.82)	26	0.25 (0-0.57)	24	0 (0-0.36)	65
hK11	0.73 (0.65-0.82)	27	0.39 (0.21-0.57)	10	0.26 (0.05-0.44)	4
LYPD3	0.73 (0.64-0.81)	28	0.18 (0-0.54)	38	0.02 (0-0.22)	52
GPC1	0.73 (0.65-0.8)	29	0.25 (0.03-0.38)	24	0.07 (0.02-0.3)	24
TLR3	0.73 (0.65-0.81)	29	0.1 (0.03-0.34)	63	0.07 (0.02-0.13)	24
VEGFR.2	0.72 (0.64-0.8)	31	0.13 (0-0.41)	49	0.02 (0-0.18)	52
WIF.1	0.72 (0.63-0.81)	32	0.25 (0.03-0.54)	24	0.07 (0-0.3)	24
ERBB2	0.72 (0.62-0.8)	33	0.16 (0.02-0.41)	44	0.03 (0-0.25)	40
CRNN	0.72 (0.63-0.8)	34	0.08 (0.02-0.21)	67	0.05 (0-0.11)	35
CEACAM1	0.72 (0.62-0.8)	35	0.13 (0-0.54)	49	0 (0-0.2)	65
FR-alpha	0.71 (0.62-0.8)	36	0.26 (0.15-0.51)	22	0.21 (0.11-0.34)	9
CD70	0.71 (0.63-0.79)	37	0.08 (0-0.34)	67	0.02 (0-0.11)	52
RSPO3	0.71 (0.62-0.8)	38	0.38 (0.21-0.52)	11	0.25 (0.02-0.41)	6
RET	0.71 (0.62-0.79)	39	0.26 (0-0.41)	22	0 (0-0.3)	65
TNFRSF19	0.7 (0.62-0.79)	40	0.15 (0-0.43)	48	0.03 (0-0.21)	40

NOTE: The 40 proteins with AUC values of 0.7 or greater were ranked, as well as their sensitivity at 95% or 98% specificity. Values in parenthesis are the 95% CIs. Representative ROC curves for the six proteins with the most elevated levels in the sera of ovarian cancer patients versus healthy women are shown in Fig. 2B. A complete list of all 92 proteins is provided in Supplementary Table S4.

sample will be identified correctly) of each protein, we calculated the AUC based on ROC curves for each of the 92 proteins. In total, 40 proteins had an estimated AUC of at least 0.70 and were significantly associated with cancer status (Table 2). The AUC for CA125 was 0.97 (95% CI, 0.93–1.0), and the AUC for HE4 was 0.92 (95% CI, 0.87–0.96). ROC curves are shown (Fig. 2B) for the six proteins that had significantly elevated levels in the serum of women with ovarian cancer (Fig. 2A). Three additional proteins (KLK11, FR-alpha, and RSPO3) that were elevated in the ovarian cancer serum samples had AUC values >0.70. A list of the AUC values for all 92 proteins is provided in Supplementary Table S4.

The sensitivity at two levels of specificity (95% and 98%) was also calculated and is shown for the 40 proteins with AUC values >0.70 (Table 2). At 95% specificity, CA125 had a sensitivity of 95% (95% CI,

88.0%–100.0%), whereas HE4 had a sensitivity of 70% (95% CI, 52%–84%). At the stringent level of 98% specificity, CA125 had the highest sensitivity with a value of 93% (95% CI, 80%–98%), and HE4 ranked second with a sensitivity of 54% (95% CI, 36%–79%), whereas none of the other 90 proteins had a sensitivity above 50%. The sensitivity and specificity values for all 92 proteins are provided in Supplementary Table S4.

A multiprotein biomarker for late-stage ovarian cancer

Due to the low prevalence of ovarian cancer, a screening test must achieve a minimum specificity of 99.6% and a sensitivity of >75% for early stage disease to avoid an unacceptable level of false-positive results and achieve a positive predictive value of 10% (1). At the highly stringent level of 99.6% specificity, CA125 had a sensitivity of 85% (95% CI, 77%–98%) in the Oncology II plates. We hypothesized that combining additional

Downloaded from <http://aacrjournals.org/cancerpreventionresearch/article-pdf/12/3/171/2244383/171.pdf> by guest on 10 November 2024

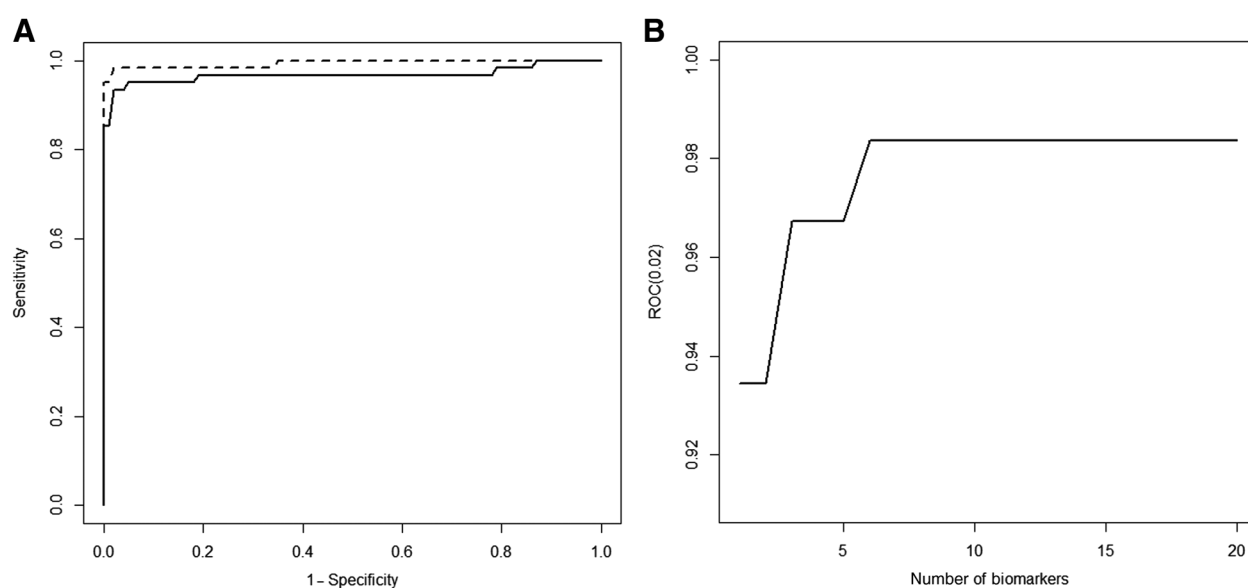


Figure 3.

ROC curves for CA125 alone and the six-protein classifier. **A**, ROC curves were graphed for discriminating advanced stage serous ovarian cancer versus healthy women for CA125 alone (solid line; AUC = 0.968) and the six-protein classifier (dashed line; AUC = 0.994) developed using supervised machine learning. **B**, ROC (0.02) values as a function of the number of proteins in the classifier. The sensitivity values for 98% specificity were graphed with the initial sensitivity of one protein (i.e., CA125 at 93.4%). The sensitivity increased as additional proteins were added, with a maximum plateau reached at 98.4% with six proteins total.

proteins with CA125 may improve the sensitivity even further.

We systematically tested all combinations of the 92 proteins using machine-learning techniques to develop a multiprotein biomarker for discriminating between ovarian cancer cases versus healthy women. A combined biomarker panel consisting of six proteins improved the AUC from 0.968 (95% CI, 0.926–0.998) for CA125 alone to 0.994 (95% CI, 0.989–1; Fig. 3A), and the sensitivity corresponding to 98% specificity increased from 93.4% (78.7%–98.4%) to 98.4% (93.4%–100%; Fig. 3B). The six proteins that were included in the multiprotein biomarker were CA125, FGF1, S100A4, EGF, ICOSLG, and MSLN. At the highly stringent level of 99.6% specificity, which is required for a screening test, this six-protein multiprotein classifier reached a sensitivity of 95.1%.

Validation of Proseek values by ELISA

Because the Proseek data indicated that the five proteins added to CA125 in the multiprotein model were present at lower levels in the serum of the ovarian cancer patients compared with healthy women, we validated the results for the two proteins (EGF and MSLN) with the most reliable commercially available kits. There is a controversy in the literature as to whether EGF is present at elevated or decreased levels in the sera of ovarian cancer patients relative to healthy women (29, 30). The Proseek Oncology II values showed decreased levels of EGF in the serum of ovarian cancer patients (Fig. 4A). Similarly, the 88 serum samples that we tested by ELISAs from Raybiotech (Fig. 4B)

and R&D Systems (Fig. 4C) showed that the ovarian cancer samples had decreased concentrations of EGF compared with healthy controls. Data from the Proseek plate exhibited a strong correlation with the protein values from both ELISA kits, with a Pearson correlation coefficient of 0.88 (95% CI, 0.82–0.92) for the Raybiotech kit and a Pearson correlation coefficient of 0.83 (95% CI, 0.75–0.89) for the R&D Systems Kit. Furthermore, there was an excellent correlation between the two ELISA kits ($R = 0.96$, $P < 0.01$).

MSLN has been described in the literature as a protein that is elevated in the serum of ovarian cancer patients (31), which is in contrast to the values that we observed from our Proseek Oncology II plate. When we tested 88 serum samples on an R&D Systems MSLN ELISA, we noted that MSLN levels were elevated in the ovarian cancer samples; the Proseek and ELISA values did not correlate. When PEA experiments were rerun specifically for MSLN using these same 88 serum samples, elevated levels of MSLN were observed in the cancer samples relative to the healthy controls. These results were subsequently confirmed using a second ELISA kit from Biologend, further confirming reports in the literature (31).

Discussion

Development of a blood-based assay for ovarian cancer detection could significantly improve the survival of patients if it identified cancers earlier. This is the first study to use Proseek Multiplex Oncology II plates to quantify 92

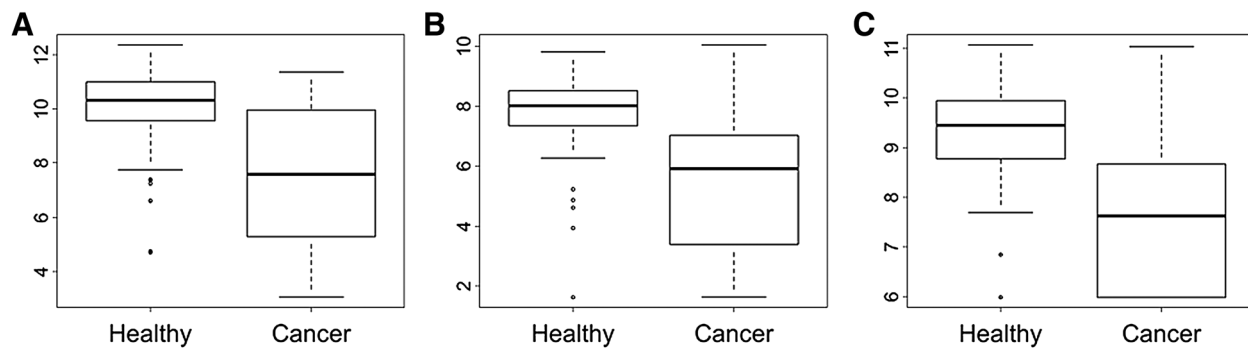


Figure 4.

Oncology II values for EGF correlate with protein values obtained by two different ELISA kits. Quantile boxplots showing expression of EGF in serum samples from 53 controls (Healthy) compared with 35 advanced stage serous ovarian cancer cases (Cancer). Outliers are defined as any value higher or lower than 1.5 multiplied by the interquartile range. EGF values obtained on: **A**, Oncology II plate; **B**, Raybiotech ELISA kit; and **C**, R&D Systems ELISA kit.

cancer-related proteins simultaneously using a large cohort of ovarian cancer serum samples. Oncology II plates were not designed specifically for monitoring the serum of ovarian cancer patients. The majority of the proteins on the Oncology II plate are secreted proteins that are elevated in the tissues or sera of other types of cancer. By conducting literature searches, we found that almost half (44/92) of the proteins have been reported to have elevated levels of expression in ovarian cancer; however, many of these 44 proteins were examined in ovarian cancer tissues, but not in serum. Importantly, several of the key proteins known to be elevated in the sera of ovarian cancer patients (e.g., CA125, HE4, MSLN, NECT4, MK, KLK11, KLK13, IL6, FOLR1; refs. 1, 22–25, 27, 32) are present on the Oncology II plate. In addition, other promising serum biomarkers for ovarian cancer are also included on the Oncology II plates, including proteins that are expressed in ovarian cancer tissues, as shown by IHC, or for which gene microarray data indicate upregulation of RNA. Relatively few of the proteins (~10%) were reported to have decreased expression in ovarian cancer compared with control samples. The other 40 proteins on the Oncology II plate have not yet been examined for their expression in ovarian cancer samples.

Similar to our previous analysis (13), using two different clustering methods (PCA and unsupervised hierarchical clustering), the healthy samples separated from the late-stage ovarian cancer samples, reinforcing the idea that a multiplex assay would be more effective at identifying ovarian cancer than any one protein measurement. The NPX data from the Oncology II plates for CA125 were strongly correlated with previously measured clinical laboratory values (correlation coefficient of 0.91, with a 95% CI of 0.88–0.94), providing further evidence that the PEA assay is comparable with existing technology for the analysis of serum samples.

Our analysis shows that many of the proteins on the Oncology II plate are associated with ovarian cancer.

Overall, we detected statistically significant ($P < 0.006$) differences between ovarian cancer samples and healthy controls for 52 of the 92 proteins. Many of these proteins have not been previously recognized as serum biomarkers for ovarian cancer. Twelve of these 52 proteins were present at significantly higher levels in the ovarian cancer serum samples compared with the healthy samples.

The ovarian cancer biomarker CA125 has primarily been used to monitor ovarian cancer recurrence and for differential diagnosis of pelvic masses, although using current technologies, it is not adequate for ovarian cancer screening. The high AUC value that we obtained (0.97) comparing sera from late-stage ovarian cancer patients with sera from healthy women confirms CA125's role as the best ovarian cancer biomarker to date. On the Oncology II plate, CA125 achieved 95% sensitivity at 95% specificity, and achieved 93% sensitivity at 98% specificity. We recognize that the 61 late-stage serum samples utilized in this preliminary study do not allow for meaningful comparison with CA125, because CA125 alone offers such high classification power in these samples. It was therefore difficult to improve upon the sensitivity, although we did succeed in increasing the sensitivity from 93.4% to 98.4% at 98% specificity using 5 additional proteins.

As expected, HE4 performed very closely to CA125, with an AUC value of 0.92. HE4 levels on the Oncology II plates were strongly correlated with ELISA values (correlation coefficient of 0.86, with a 95% CI of 0.8–0.91), providing further evidence that the PEA technology is comparable with existing technologies for testing serum. HE4 achieved 70% sensitivity at 95% specificity on the Oncology II plates when comparing sera from late-stage ovarian cancer cases with healthy women, but at 98% specificity, the sensitivity for HE4 decreased to 54%.

In agreement with our previous study on the Oncology Iv2 plate (13) and others (25, 28), we found high levels of both MK and IL6 in the sera of patients with ovarian cancer. FR-alpha, KLK11, and NECT4 also showed a significant

increase in their levels of expression in the sera of ovarian cancer patients as previously reported (24, 26, 27).

ADAMTS15, RSPO3, TNFRSF6B, and FCRLB, were among the 12 proteins that were significantly elevated in the sera of ovarian cancer patients compared with healthy controls (Table 1). ADAMTS15, RSPO3, and TNFRSF6B have been associated with ovarian cancer tissues (33–35) but have not been previously shown to be present at elevated levels in serum. In breast cancer tissues, increased ADAMTS15 expression is associated with prolonged relapse-free survival (36). FCRLB has not been previously reported to have an association with ovarian cancer.

Forty proteins on the Oncology II plate were present at significantly lower levels in the ovarian cancer serum samples compared with the healthy samples, as has been reported in the literature for some. For example, when SCF was tested on the Oncology Iv2 plate, we found that it was present at decreased levels in sera from early and late-stage serous ovarian cancer patients relative to healthy controls (13). Lawicki and colleagues (37) also demonstrated statistically lower SCF plasma levels in women with ovarian cancer versus benign ovarian tumors.

EGF is another protein that was found on the Oncology II plate to be present at significantly lower levels in the sera of ovarian cancer patients compared with healthy controls. When we tested serum samples using two different ELISA kits, we validated these findings. Conflicting reports are found in the literature regarding EGF levels in ovarian cancer (29, 30). This may be attributed to the antibodies that are used in the studies (recognizing intact protein vs. cleaved form; polyclonal vs. monoclonal antibody; or the various isoforms of EGF in the sample) or whether pre- or postsurgical serum was tested (30).

Based on the literature, some of these 40 proteins have been reported to be upregulated, not downregulated, in ovarian cancer tissues and/or sera. For example, S100A4 was reported to be elevated in ovarian cancer tissues (38), whereas we found it to be present at low levels in sera. Some differences in protein levels may be attributed to the biospecimen source that was used for the assay as well as the technique used to quantify the protein.

One possible explanation for the conflicting results could be due to preanalytical variables. Shen and colleagues (39) have shown that several proteins are affected by delays in centrifugation of plasma samples. Delays as short as 3 hours caused the release of EGF and 39 other proteins from cells. Their findings are critical to studies on biomarker discovery and validation, and emphasize the importance of strict adherence to protocols for processing biospecimens. In our study, all samples were processed using the same standard operating procedure, so there is no reason to suspect that differences observed in EGF levels were due to preanalytical variations.

Olink's criteria for including a protein on the Oncology II plate are based on upregulation of the RNA and/or protein

levels in some types of cancer according to the TCGA, The Human Protein Atlas, and the literature. However, the expression of each of these 92 proteins in the sera of ovarian cancer patients is largely unknown. For example, FGFBP1 is upregulated in several tumors, and it is associated with early stages of tumor formation where angiogenesis plays a critical role (40). Yet we found FGFBP1 to be present at lower levels in the sera of ovarian cancer patients than in the healthy controls.

MSLN was in this group of downregulated serum proteins even though RNA and protein levels are known to be elevated in ovarian cancer tissues (41) and serum levels are reported to be elevated in ovarian cancer patients (31). For this reason, PEA experiments were repeated specifically to study MSLN levels, and two different ELISAs were performed to quantify MSLN levels. In these experiments, MSLN levels were found to be elevated in the sera of the ovarian cancer patients relative to the healthy controls. Because MSLN binds to CA125 (42), it is possible that the multiplex PEA technology could be affected by CA125–MSLN complexes that form in ovarian cancer patients' sera (43). Using the earlier proximity ligation assay technology, Fredriksson and colleagues (44) found a 29-fold increase in CA125 and a 1.3-fold increase in MSLN between ovarian cancer and healthy plasma samples. If different antibodies are used, it may account for differences observed between assays.

Another factor that should be taken into account in such studies is that autoantibodies against proteins in the sera of ovarian cancer patients may interfere with the technology. This is a viable possibility, because neoantigens that are present early in the progression of cancer have been shown to elicit an immune response in patients. This has been shown for several ovarian cancer biomarkers, including CA125 (45), p53 (46), and MSLN (47).

Due to the low prevalence of ovarian cancer in the general population, a screening test must be both highly sensitive and specific for early stage disease to avoid an unacceptable level of false-positive results. A minimum specificity of 99.6% and a sensitivity of >75% are necessary to achieve a positive predictive value of 10% (1). In the UKCTOCS and NROSS studies, this has been achieved by measuring CA125 over time and performing transvaginal sonography only if the CA125 values rise. In this study, we generated a multiprotein classifier using six proteins (including CA125) that increased the sensitivity from 93.4% to 98.4%, while holding the specificity fixed at 98%. This 5% increase in sensitivity would have a significant effect on the number of women correctly identified when screening a large population. We acknowledge the potential limitation of our multiprotein classifier for ovarian cancer screening in asymptomatic women because only late-stage serous ovarian carcinoma serum samples were used in this study.

Interestingly, all five proteins that were added to CA125 in the multiprotein classifier were present at lower levels in

ovarian cancer sera relative to healthy controls when samples were run on the Oncology II plates. These five proteins had AUC values greater than 0.80, and their sensitivity at 95% specificity ranged from 48% to 97% when tested individually. Our results suggest that future biomarker discovery studies on unconventional platforms, such as the Proseek plates, may allow for discriminating cancer samples from healthy controls. In addition, these results suggest that future biomarker discovery studies should not exclusively focus on proteins that have increased levels in ovarian cancer sera. The utility of downregulated proteins as biomarkers will, however, depend critically on the precision of the assay and minimizing variation. It also remains to be determined whether significant lead time will be obtained with biomarkers that are often downregulated as an indirect effect of cancer growth on other tissues and organs in the host.

In this study, we focused on high-grade serous ovarian cancer, because it is the most prevalent and deadly subtype of ovarian cancer. Serum levels of CA125 are frequently higher in women with late-stage serous ovarian cancer compared with the other ovarian cancer subtypes, so it will be necessary to test sera from other ovarian cancer subtypes and most likely incorporate additional proteins into our next generation of multiprotein classifier, such as CA19.9 or REG4 which have previously been identified as biomarkers for nonserous ovarian cancer subtypes (48, 49).

Additional studies using a larger cohort of early stage ovarian cancer patients and asymptomatic patients (e.g., PLCO and UKCTOCS) in which serum samples were obtained several years prior to the diagnosis of high-grade serous ovarian cancer (6, 50) will allow for validation of these biomarkers. A customized plate comprised of ovarian cancer-specific proteins could enhance the ability of this technology to identify ovarian cancers at early stages. Proseek plates can be designed to detect a limited number of proteins that will achieve a high degree of sensitivity and specificity, and be used as a screening tool for detection of early stage ovarian cancer in the general population. Because the PEA technology combines the specificity of ELISA and other antibody-based methods with the sensitivity of the PCR, technologies that are already deployed in the clinic, it may prove to be readily translated into a screening test by virtue of its ability to perform multiplex biomarker detection and high-throughput quantification.

Disclosure of Potential Conflicts of Interest

M.A. Geller is a consultant/advisory board member for Voluntas. R.C. Bast Jr. has royalties from Fujirebio Diagnostics

References

1. Yang WL, Lu Z, Bast RC Jr. The role of biomarkers in the management of epithelial ovarian cancer. *Expert Rev Mol Diagn* 2017;17:577–91.

Inc. for discovery of CA125. No potential conflicts of interest were disclosed by the other authors.

Authors' Contributions

Conception and design: A.P.N. Skubitz, Q. Cao, J.S. Koopmeiners
Development of methodology: A.P.N. Skubitz, K.L.M. Boylan, T.K. Starr

Acquisition of data (provided animals, acquired and managed patients, provided facilities, etc.): A.P.N. Skubitz, K. Geschwind, M.A. Geller, J. Celestino, K.H. Lu

Analysis and interpretation of data (e.g., statistical analysis, biostatistics, computational analysis): A.P.N. Skubitz, K.L.M. Boylan, K. Geschwind, Q. Cao, T.K. Starr, J.S. Koopmeiners

Writing, review, and/or revision of the manuscript: A.P.N. Skubitz, K.L.M. Boylan, K. Geschwind, Q. Cao, T.K. Starr, M.A. Geller, R.C. Bast Jr., K.H. Lu, J.S. Koopmeiners

Administrative, technical, or material support (i.e., reporting or organizing data, constructing databases): A.P.N. Skubitz, K.L.M. Boylan, K. Geschwind, Q. Cao, J. Celestino, R.C. Bast Jr.

Study supervision: A.P.N. Skubitz, K.L.M. Boylan

Other (helped to identify and provided the samples for the analysis from the MD Anderson Gynecologic Cancer Tissue Bank): R.C. Bast Jr.

Acknowledgments

We thank Dr. Kenneth B. Beckman, Dinesha Walek, and Darrell Johnson in the University of Minnesota Genomics Center for helpful discussions and assistance with generating data on the Proseek Multiplex Oncology II plate. We thank Anna C. Rogers for performing the mesothelin ELISA. We thank Brad Nawa, Katarina Hornaeus, and Martin Lundberg of Olink for help with normalization of data and discussions on the use of the Proseek plates.

This work was supported by a grant from the Minnesota Ovarian Cancer Alliance (A.P.N. Skubitz), the Randy Shaver Cancer Research and Community Fund (T.K. Starr), Masonic Cancer Center (T.K. Starr), and the Masonic Cancer Center Support Grant for the Biostatistics Core (Q. Cao and J.S. Koopmeiners). Funding for R.C. Bast, Jr., K.H. Lu, and J. Celestino was provided by the Early Detection Research Network (5 U01 CA200462-02), the MD Anderson Ovarian SPOR (P50 CA83639 and P50 CA217685), National Cancer Institute, Department of Health and Human Services, the Cancer Prevention Research Institute of Texas (RP160145), Golfer's Against Cancer, the Mossy Foundation, the Roberson Endowment, National Foundation for Cancer Research, and generous donations from Stuart and Gaye Lynn Zarrow.

The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked *advertisement* in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

Received June 26, 2018; revised November 6, 2018; accepted January 25, 2019; published first February 1, 2019.

2. Nolen BM, Lokshin AE. Biomarker testing for ovarian cancer: clinical utility of multiplex assays. *Mol Diagn Thera* 2013;17: 139–46.

3. Menon U, Ryan A, Kalsi J, Gentry-Maharaj A, Dawney A, Habib M, et al. Risk algorithm using serial biomarker measurements doubles the number of screen-detected cancers compared with a single-threshold rule in the United Kingdom collaborative trial of ovarian cancer screening. *J Clin Oncol* 2015;33:2062–71.
4. Buys SS, Partridge E, Greene MH, Prorok PC, Reding D, Riley TL, et al. Ovarian cancer screening in the Prostate, Lung, Colorectal and Ovarian (PLCO) cancer screening trial: findings from the initial screen of a randomized trial. *Am J Obstet Gynecol* 2005;193:1630–9.
5. Lu KH, Skates S, Hernandez MA, Bedi D, Bevers T, Leeds L, et al. A 2-stage ovarian cancer screening strategy using the Risk of Ovarian Cancer Algorithm (ROCA) identifies early-stage incident cancers and demonstrates high positive predictive value. *Cancer* 2013;119:3454–61.
6. Jacobs IJ, Menon U, Ryan A, Gentry-Maharaj A, Burnell M, Kalsi JK, et al. Ovarian cancer screening and mortality in the UK Collaborative Trial of Ovarian Cancer Screening (UKCTOCS): a randomised controlled trial. *Lancet* 2016;387:945–56.
7. Yurkovetsky Z, Skates S, Lomakin A, Nolen B, Pulsipher T, Modugno F, et al. Development of a multimarker assay for early detection of ovarian cancer. *J Clin Oncol* 2010;28:2159–66.
8. Trabert B, Pinto L, Hartge P, Kemp T, Black A, Sherman ME, et al. Pre-diagnostic serum levels of inflammation markers and risk of ovarian cancer in the prostate, lung, colorectal and ovarian cancer (PLCO) screening trial. *Gynecol Oncol* 2014;135:297–304.
9. Assarsson E, Lundberg M, Holmquist G, Bjorkesten J, Thorsen SB, Ekman D, et al. Homogenous 96-plex PEA immunoassay exhibiting high sensitivity, specificity, and excellent scalability. *PLoS ONE* 2014;9:e95192.
10. Lundberg M, Eriksson A, Tran B, Assarsson E, Fredriksson S. Homogeneous antibody-based proximity extension assays provide sensitive and specific detection of low-abundant proteins in human blood. *Nucleic Acids Res* 2011;39:e102.
11. Thorsen SB, Lundberg M, Villablanca A, Christensen SL, Belling KC, Nielsen BS, et al. Detection of serological biomarkers by proximity extension assay for detection of colorectal neoplasias in symptomatic individuals. *J Transl Med* 2013;11:253.
12. Chen H, Zucknick M, Werner S, Knebel P, Brenner H. Head-to-Head comparison and evaluation of 92 plasma protein biomarkers for early detection of colorectal cancer in a true screening setting. *Clin Cancer Res* 2015;21:3318–26.
13. Boylan KLM, Geschwind K, Koopmeiners JS, Geller MA, Starr TK, Skubitz APN. A multiplex platform for the identification of ovarian cancer biomarkers. *Clin Proteomics* 2017;14:34.
14. de Hoon MJ, Imoto S, Nolan J, Miyano S. Open source clustering software. *Bioinformatics* 2004;20:1453–4.
15. Page RD. TreeView: an application to display phylogenetic trees on personal computers. *Computer Appl Biosci* 1996;12:357–8.
16. R Core Team. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2017.
17. Benjamini Y, Yekutieli D. The control of the false discovery rate in multiple testing under dependency. *Ann Stat* 2001;29:1165–88.
18. Pepe MS. The statistical evaluation of medical tests for classification and prediction. Oxford: Oxford University Press; 2004.
19. Efron B, Tibshirani RJ. An introduction to the bootstrap. Florida: Chapman and Hall/CRC; 1994.
20. Tibshirani R. Regression shrinkage and selection via the lasso. *J Royal Stat Soc Ser B* 1996;58:267–88.
21. Hastie T, Tibshirani R, Friedman J. The elements of statistical learning: Data mining, inference, and prediction. NY: Springer-Verlag; 2009.
22. Penson RT, Kronish K, Duan Z, Feller AJ, Stark P, Cook SE, et al. Cytokines IL-1beta, IL-2, IL-6, IL-8, MCP-1, GM-CSF and TNFalpha in patients with epithelial ovarian cancer and their relationship to treatment with paclitaxel. *Int J Gynecol Cancer* 2000;10:33–41.
23. Scorilas A, Borgono CA, Harbeck N, Dorn J, Schmalfeldt B, Schmitt M, et al. Human kallikrein 13 protein in ovarian cancer cytosols: a new favorable prognostic marker. *J Clin Oncol* 2004;22:678–85.
24. Diamandis EP, Okui A, Mitsui S, Luo LY, Soosaipillai A, Grass L, et al. Human kallikrein 11: a new biomarker of prostate and ovarian carcinoma. *Cancer Res* 2002;62:295–300.
25. Wu X, Zhi X, Ji M, Wang Q, Li Y, Xie J, et al. Midkine as a potential diagnostic marker in epithelial ovarian cancer for cisplatin/paclitaxel combination clinical therapy. *Am J Cancer Res* 2015;5:629–38.
26. Derycke MS, Pambuccian SE, Gilks CB, Kalloger SE, Ghidouche A, Lopez M, et al. Nectin 4 overexpression in ovarian cancer tissues and serum: potential role as a serum biomarker. *Am J Clin Pathol* 2010;134:835–45.
27. Kurosaki A, Hasegawa K, Kato T, Abe K, Hanaoka T, Miyara A, et al. Serum folate receptor alpha as a biomarker for ovarian cancer: Implications for diagnosis, prognosis and predicting its local tumor expression. *Int J Cancer* 2016;138:1994–2002.
28. Berek JS, Chung C, Kaldi K, Watson JM, Knox RM, Martinez-Maza O. Serum interleukin-6 levels correlate with disease status in patients with epithelial ovarian cancer. *Am J Obstet Gynecol* 1991;164:1038–42; discussion 42–3.
29. Balcan E, Demirkiran F, Aydin Y, Sanioglu C, Bese T, Arvas M, et al. Serum levels of epidermal growth factor, transforming growth factor, and c-erbB2 in ovarian cancer. *Int J Gynecol Cancer* 2012;22:1138–42.
30. Baron AT, Lafky JM, Boardman CH, Balasubramaniam S, Suman VJ, Podratz KC, et al. Serum sErbB1 and epidermal growth factor levels as tumor biomarkers in women with stage III or IV epithelial ovarian cancer. *Cancer Epidemiol Biomarkers Prev* 1999;8:129–37.
31. Madeira K, Dondossola ER, Farias BF, Simon CS, Alexandre MC, Silva BR, et al. Mesothelin as a biomarker for ovarian carcinoma: a meta-analysis. *Anais da Academia Brasileira de Ciencias* 2016;88:923–32.
32. Abdel-Azeez HA, Labib HA, Sharaf SM, Refai AN. HE4 and mesothelin: novel biomarkers of ovarian carcinoma in patients with pelvic masses. *Asian Pacific J Cancer Prev* 2010;11:111–6.
33. Boone JD, Arend RC, Johnston BE, Cooper SJ, Gilchrist SA, Oelschlager DK, et al. Targeting the Wnt/beta-catenin pathway in primary ovarian cancer with the porcupine inhibitor WNT974. *Lab Invest* 2016;96:249–59.
34. Jonsson JM, Bartuma K, Dominguez-Valentin M, Harbst K, Ketabi Z, Malander S, et al. Distinct gene expression profiles in ovarian cancer linked to Lynch syndrome. *Familial cancer* 2014;13:537–45.
35. Tucker LA, Zhang Q, Sheppard GS, Lou P, Jiang F, McKeegan E, et al. Ectopic expression of methionine aminopeptidase-2 causes cell transformation and stimulates proliferation. *Oncogene* 2008;27:3967–76.
36. Porter S, Span PN, Sweep FC, Tjan-Heijnen VC, Pennington CJ, Pedersen TX, et al. ADAMTS8 and ADAMTS15 expression predicts survival in human breast carcinoma. *Int J Cancer* 2006;118:1241–7.

37. Lawicki S, Gacuta-Szumarska E, Bedkowska GE, Szmitkowski M. Hematopoietic cytokines as tumor markers in gynecological malignancies. A multivariate analysis in epithelial ovarian cancer patients. *Growth Factors* 2012;30:357–66.
38. Kikuchi N, Horiuchi A, Osada R, Imai T, Wang C, Chen X, et al. Nuclear expression of S100A4 is associated with aggressive behavior of epithelial ovarian carcinoma: an important autocrine/paracrine factor in tumor progression. *Cancer Sci* 2006;97:1061–9.
39. Shen Q, Bjorkestén J, Galli J, Ekman D, Broberg J, Nordberg N, et al. Strong impact on plasma protein profiles by precentrifugation delay but not by repeated freeze-thaw cycles, as analyzed using multiplex proximity extension assays. *Clin Chem Lab Med* 2018;56:582–94.
40. Tassi E, Wellstein A. The angiogenic switch molecule, secreted FGF-binding protein, an indicator of early stages of pancreatic and colorectal adenocarcinoma. *Sem Oncol* 2006;33:S50–6.
41. Cheng WF, Huang CY, Chang MC, Hu YH, Chiang YC, Chen YL, et al. High mesothelin correlates with chemoresistance and poor survival in epithelial ovarian carcinoma. *Br J Cancer* 2009;100:1144–53.
42. Kaneko O, Gong L, Zhang J, Hansen JK, Hassan R, Lee B, et al. A binding domain on mesothelin for CA125/MUC16. *J Biol Chem* 2009;284:3739–49.
43. Coelho R, Marcos-Silva L, Ricardo S, Ponte F, Costa A, Lopes JM, et al. Peritoneal dissemination of ovarian cancer: role of MUC16-mesothelin interaction and implications for treatment. *Expert Rev Anticancer Ther* 2018;18:177–86.
44. Fredriksson S, Horecka J, Brustugun OT, Schlingemann J, Koong AC, Tibshirani R, et al. Multiplexed proximity ligation assays to profile putative plasma biomarkers relevant to pancreatic and ovarian cancer. *Clin Chem* 2008;54:582–9.
45. Fortner RT, Schock H, Le Cornet C, Husing A, Vitonis AF, Johnson TS, et al. Ovarian cancer early detection by circulating CA125 in the context of anti-CA125 autoantibody levels: Results from the EPIC cohort. *Int J Cancer* 2018;142:1355–60.
46. Yang WL, Gentry-Maharaj A, Simmons A, Ryan A, Fourkala EO, Lu Z, et al. Elevation of TP53 autoantibody before CA125 in preclinical invasive epithelial ovarian cancer. *Clin Cancer Res* 2017;23:5912–22.
47. Ho M, Hassan R, Zhang J, Wang QC, Onda M, Bera T, et al. Humoral immune response to mesothelin in mesothelioma and ovarian cancer patients. *Clin Cancer Res* 2005;11:3814–20.
48. Santotoribio JD, Garcia-de la Torre A, Canavate-Solano C, Arce-Matute F, Sanchez-del Pino MJ, Perez-Ramos S. Cancer antigens 19.9 and 125 as tumor markers in patients with mucinous ovarian tumors. *Eur J Gynaecol Oncol* 2016;37:26–9.
49. Lehtinen L, Vesterkvist P, Roering P, Korpela T, Hattara L, Kaipio K, et al. REG4 is highly expressed in mucinous ovarian cancer: a potential novel serum biomarker. *PLoS One* 2016;11:e0151590.
50. Black A, Huang WY, Wright P, Riley T, Mabie J, Mathew S, et al. PLCO: evolution of an epidemiologic resource and opportunities for future studies. *Rev Recent Clin Trials* 2015;10:238–45.

